Coal Smoke and the Costs of the Industrial Revolution*

W. Walker Hanlon
UCLA and NBER

June 30, 2016

Abstract

One of the longest-running debates in economic history is over the costs of the environmental degradation that accompanied the Industrial Revolution. Focusing on pollution from coal burning in the cities of 19th century England, this study brings together new data and a novel analysis strategy in order to estimate the magnitude of the effect of industrial pollution on local economic development from 1851-1911. My results show that the externality that coal use in some industries imposed on all other industries within a city had substantial costs, in terms of reduced city employment growth, and led to a slower overall rate of urbanization. In a city where coal use grew at a rate that was one standard deviation above the average, I estimate that city employment growth over a twenty year period was lower by 21-26 percentage points, equal to roughly one-half of average city employment growth. A counterfactual analysis suggests that plausible improvements in coal use efficiency from 1851-1911 could have increased the urbanization rate in England in 1911 from 34% to 38%, nearing modern levels. JEL Codes: N53, N13, R11, Q52

*I thank David Atkin, Dan Bogart, Leah Boustan, Karen Clay, Dora Costa, Neil Cummins, Pablo Fajgelbaum, Roger Foquet, Matt Kahn, Ed Kung, Miren Lafourcade, Adriana Lleras-Muney, Petra Moser, Alexander Rothenberg, Werner Troesken and seminar participants at Arizona, Arizona State, Boston University, Bristol, Carnegie-Mellon, LSE, NYU Stern, UC Davis, UC Irvine, UCLA, USC, Warwick, Wharton and the World Bank for their helpful comments. Reed Douglas, Qiyi Song and Vitaly Titov provided excellent research assistance. This paper has previously been circulated under the title "Endogenous City Disamenities: Lessons from Industrial Pollution in 19th Century Britain." I thank the Rosalinde and Arthur Gilbert Program in Real Estate, Finance and Urban Economics and the California Center for Population Research for generous funding. Author contact information: whanlon@econ.ucla.edu, 8283 Bunche Hall, UCLA, 405 Hilgard Ave., Los Angeles, CA 90095, Ph: 202-246-2458.

1 Introduction

The Industrial Revolution represented a watershed in economic growth, yet there is a long-running debate over the costs that industrialization imposed on the generations that lived through these events. In the early stages of the Industrial Revolution, factories were predominantly driven by clean water power, but by the mid-19th century, coal had become the dominant power source. As a result, by the 1840's, contemporary writers such as Engels (1845) described cities blackened with thick smoke and urban slums crowded with diseased and dying people.

How large were the external costs of Britain's reliance on dirty coal power in the 19th and early 20th century? Existing work on this topic, such as Williamson (1981b), suggests that these costs were not large. However, these findings have been criticized for drawing broad conclusions, which seem to contradict contemporary evidence, from relatively sparse data (Pollard (1981)).[1] As a result, the debate over the size of the negative externalities that accompanied the Industrial Revolution remains largely unresolved. This matters, both because of the central role that the Industrial Revolution played in economic history and because the experiences of early industrial countries can help inform the decisions of countries, such as China and India, that are currently struggling with the negative consequences of industrial growth.

This study documents the impact of industrial pollution on long-run local economic development in English cities from 1851-1911. I focus specifically on the effects of air pollution generated by coal burning, the most important type of industrial pollution during this period. In order to measure how industrial pollution affected local economic growth, I develop a new estimation approach that has three key features. One feature of my estimation strategy is that it allows me to separate the positive effect of industry growth on local employment growth, through job creation, from the negative effects that are generated when this growth occurs in heavily polluting

[1] See also (Williamson, 1981a, 1982).

industries. These negative pollution effects, which operate on all industries in a city, can occur either because pollution makes a location less attractive (the amenities channel), or because pollution makes workers and firms less productive (the productivity channel). The second feature of my estimation strategy is that it will capture the impact of pollution occurring through both of these channels. A third feature of my estimation strategy is that it can be implemented without the need for local wage and rent data, which are largely unavailable during the period I study. Instead, I use a model to show how data on quantities, in this case the quantity of employed workers, can be used in place of the more scarce data on prices (real wages in this case). As a result, my approach requires only panel data on city-industry employment, which I have constructed for every decade from 1851-1911 for 31 English cities

My results show that industrial coal use substantially reduced long-run employment growth in English cities during this period. Specifically, in English cities that experienced rapidly rising coal use, employment growth was systematically lower relative to the growth that we would have expected given the initial mix of industries in each city and national industry growth rates. The magnitude of these effects was large; based on my estimates, over a two-decade period, a city in which local industrial coal use grew at a rate that was one standard deviation above the national average would, as a consequence, have experience a reduction in employment growth of 21-26 percentage points, equal to about one-half of the average growth in employment across two-decade periods. These estimates reflect the *external* effect that coal use in some industries exerted on other sectors of the local economy. These findings are robust to the inclusion of a wide range of control variables including the initial level of coal use in each location, the city's proximity to coal reserves, city features such as weather, initial city innovation rates, changing local agglomeration forces operating through input-output connections or labor market pooling, changes in the share of high-skilled workers in the city, the share of exporting industries in the city, changes in the average local firm size, etc.

To assess the implications of these results for overall urbanization levels, I conduct a simple counterfactual to study the impact of more efficient coal use. This counterfactual is motivated by the 1871 Coal Commission Report, a detailed 1300 page study of coal use in Britain commissioned by Parliament. The report highlights substantial inefficiencies in industrial coal use and describes how simple low-cost improvements could have substantially reduced industrial coal use. However, these improvements were not adopted due to the combination of low coal prices, weak pollution regulation, and the fact that most of the impacts of pollution were external to firms. Guided by this report, I consider a counterfactual in which the growth of coal use from 1851-1911 was reduced by ten percent. My results suggest that the 31 analysis cities would have had an additional 1.5 million residents by 1911 and that their share of the English population would have been higher by four percentage points. Thus, my results suggest that had Britain adopted regulations to improve coal use efficiency, or had additional sources of clean water power been available, the nation would have been substantially more urbanized by the early 20th century.

To my knowledge this is the first study to document the effects of industrial pollution on local economic development over the long-run, though I build on previous work such as Kahn (1999).[2] This is possible, in part, because of the unique features offered by the historical setting that I consider. First, the cities covered by my study experienced high variation in the level of local pollution. This variation makes it possible to identify the impacts of pollution despite the fact that many other factors influence city growth. Second, population mobility was relatively high during this period, which means that city population and employment could respond to the

[2]Kahn (1999) studies the impact of a decline in local manufacturing on local pollution levels in rust-belt cities in the U.S., but does not estimate the impact of the pollution decline on local economic development. Another closely related paper is Chay & Greenstone (2005), which looks at the impact of pollution reductions resulting from the Clean Air Act on local housing values. Two other related papers are Banzhaf & Walsh (2008) and Bayer *et al.* (2009). The main difference between these previous contributions and the present paper is that I study long-run effects while focusing on local employment as the main outcome of interest.

effects of pollution.[3] Third, regulation was relatively limited. This includes both pollution regulation and regulations, such as zoning, that often constrain city growth in modern contexts. Fourth, detailed census data are available which allow me to track city-industry employment consistently across six decades in order to study long-run effects.

The analysis approach offered in this paper is derived from a standard Rosen-Roback model that is modified in two main ways. First, the model economy is composed of many industries that are heterogeneous in the intensity with which they use coal. Second, the overall level of coal use in a city can influence both local amenities and the productivity of local firms. The intuition behind the estimation strategy that emerges from the theory is straightforward; using panel data on city-industry employment, my analysis looks at whether employment growth in a particular city and industry systematically underperformed, relative the the national employment growth rate in that industry, in cities in which industrial coal use was increasing rapidly. To obtain plausibly exogenous variation, the change in local industrial coal use is based on the initial distribution of industries in a city interacted with the growth rates of those industries in all other locations and industry coal use intensity. This approach follows Bartik (1991), but differs from most existing work in that literature in that the estimation specification is derived from the theory.[4]

An important feature of the analysis approach that I propose is that it will capture the effect of local industrial coal use on both the level of city amenities, which influences the supply of workers in a city, as well as the impact on the productivity of local firms, which will influence the demand for workers. Previous work on this topic, such as Williamson (1981b), has focused only on the amenity channel. However, a

[3] See Long & Ferrie (2003) and Baines (1985).

[4] The only other study that I'm aware of which provides a micro-founded motivation for the Bartik instrument is Bartelme (2015). The main difference between Bartelme's paper and the model in this study is that in his study industries differ in a demand shift parameter while in this study they vary in their input shares. Having a theory that allows for variation in input shares is useful because many studies that use a Bartik instrument approach exploit industry-level variation in input use, for example, in the share of skilled to unskilled workers in an industry.

growing body of literature suggests that air pollution can have important effects on productivity.[5] The intuition behind my approach is that, regardless of whether coal use affects consumer amenities or firm productivity, the implications for employment are the same. Thus, focusing on employment as the outcome of interest allows me to capture both of these channels. In contrast, these two channels have opposing effects on the urban wage premium. Thus, if the productivity channel is important then a small urban wage premium (as found by Williamson (1981b)) can be still be consistent with large pollution costs. Moreover, the model makes it clear that when pollution affects productivity, the costs of urban pollution cannot be inferred from the urban wage premium alone.

Using a cross-section of local wage, rent and price data from 1905, I provide tentative evidence suggesting that workers were compensated with higher real wages in cities with more industrial coal use, consistent with a negative consumer amenity effect. However, the magnitude of the compensation was not large, suggesting that most of the effects I observe were driven by the impact of pollution on productivity. While these results should be viewed as tentative because they rely on simple cross-sectional regressions, the importance of the productivity channel is in-line with the existing historical literature, which emphasizes that the availability of employment was the primary driver of workers' location decisions.

This study reconciles the quantitative estimates of the costs of industrial pollution during the Industrial Revolution with the qualitative historical evidence describing the severity of the pollution problem during this period (e.g., Brimblecombe (1987), Mosley (2001),and Thorsheim (2006)) as well as with our current understanding of the substantial impacts that air pollution can have, even at the much lower concentrations experienced in modern developed economies. It also contributes to a small but growing set of quantitative studies examining the historical impact of coal-based

[5]See, e.g., Graff Zivin & Neidell (2012), Hanna & Oliva (2015), Chang *et al.* (2014), Chang *et al.* (2016), Isen *et al.* (Forthcoming), and Lavy *et al.* (2014).

air pollution, including Troesken & Clay (2011), which documents the evolution of polluted fogs in London in the 19th century, Barreca *et al.* (2014), Clay *et al.* (2015), and Clay *et al.* (2016), which study the health impacts of coal use in the 20th century, and Heblich *et al.* (2015), which shows how air pollution affected the sorting of residents across neighborhoods in 19th century British cities. Finally, this paper is related to Beach & Hanlon (2016), which documents the substantial mortality effects of industrial coal use in Britain in the mid-19th century.

This study is also related to a growing literature in urban economics focusing on the consumption value of cities (Glaeser *et al.* (2001)), and in particular, to work that considers endogenous amenities related to the composition of city residents (Rauch (1993), Moretti (2004), Diamond (2016)). In contrast, this paper highlights how the industrial composition of cities can also affect city amenities, through pollution, and introduces an analysis strategy that can be used to measure the impact of industrial pollution on local employment growth over the long-run.

In the next section I describe the empirical setting. Data and measurement are discussed in Section 3, followed by the theory, in Section 4. The main analysis is presented in Section 5, while Section 6 concludes.

2 Empirical setting

Landes (1998) describes the Industrial Revolution as composed of three elements: the replacement of human skill by machines, the introduction of engines to convert heat into work, and the substitution of mineral power sources – chiefly in the form of coal – for other power sources. One consequence of these changes was rapid growth in coal use by industry. British coal consumption averaged 65 million tons annually in 1852-1862 and rose to 181 million tons in the 1903-1912 period.[6] This amounted

[6]These figures are from the U.K. Department of Energy and Climate Change. For further details, see Appendix A.1.

to 4.3 tons per person in 1911.[7] Most of this coal – 60-65 percent – was burned by industry, and coal remained the dominant power source, by far, throughout this period.[8] Because some industries were particularly intensive users of coal, and these industries tended to agglomerate, industrial coal use could be highly geographically concentrated.[9] Also, before electricity transmission, power had to be generated on-site at factories, which were located in urban areas where they could be reached on foot by workers, increasing pollution exposure.

Figure 1 provides an illustration of the impact of industrial pollution in Sheffield, perhaps the most polluted of the northern industrial cities. These images come from 1920, after the end of the study period, but are likely to be similar to the conditions experienced during the late 19th and early 20th centuries. The left-hand image was taken on Sunday morning, when the factories were at rest, while the right-hand image was taken from the same vantage point on Monday at noon, when the factories were at work. Residential pollution would have been present at both times, so the contrast between these images illustrates the impact that industrial pollution had in the industrial cities of England.

Burning coal releases a variety of pollutants into the atmosphere, including suspended particles of soot and other matter, sulfur dioxide, and carbon dioxide. The release of suspended particles is particularly severe when combustion is inefficient, as it often was in the 19th century. These pollutants have a variety of negative effects on

[7] These figures are in imperial tons per year. For comparison, in 2012 the U.S. consumed about 2.5 tons of coal per person annually, China consumed about 2.7 tons per person, and Australia, one of the heaviest users, consumed around 5.8 tons per person. However, today most coal use occurs in electricity generation plants outside of urban centers.

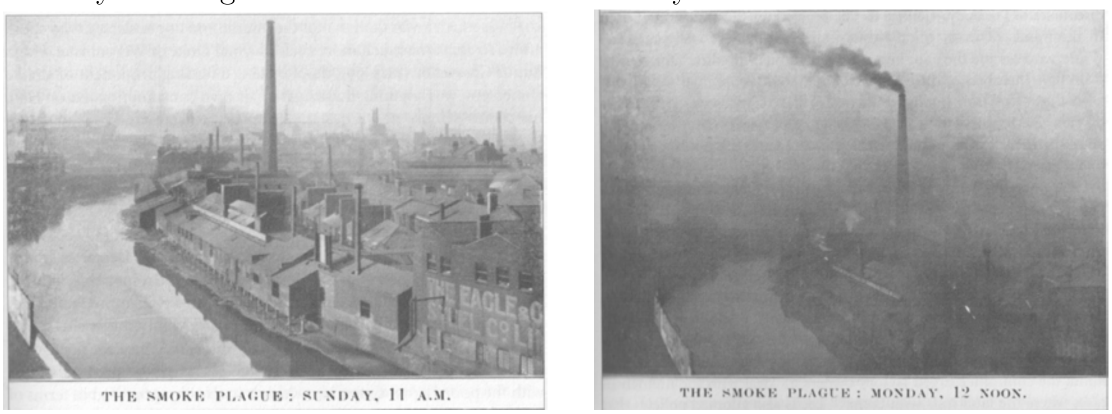
[8] Data from Mitchell (1988). Industry here includes both manufacturing and mining. In contrast, residential coal use accounted for only 17-25 percent of domestic consumption, but attracted more attention because it was particularly important in London. The remainder is composed of use by transportation and utilities. It is worth noting that residential coal use was more polluting, per ton burned, than industrial coal use. This is because it was burned less efficiently (at lower temperatures) and released at lower altitudes.

[9] These agglomeration patterns generally dated to the late 18th or early 19th century and were often due to geographic factors. For example, the location of the textile industry in the Northwest of England was driven by historical factors, such as the location of water power, that were of less importance in the second half of the 19th century (Crafts & Wolf (2014)).

the human system which have been documented in a large literature.[10] They affect people at all ages, with particularly severe impacts for the very young and very old.

Figure 1: An illustration from Sheffield in 1920

Sunday morning – the factories at rest   Monday at noon – the factories at work



The pictures above were taken from the same vantage point in Sheffield in 1920. While this is after the study period, the levels of pollution it reveals are likely similar to those experienced during the period I study. From William Blake Richmond, "The Smoke Plague of London", in *London of the Future,* Ashton Webb Ed., 1921. Reproduced from *Inventing Pollution*, by Peter Thorsheim (2006), Athens: Ohio University Press.

Several recent studies have documented the impact of direct pollution exposure on worker productivity. For example, Graff Zivin & Neidell (2012) show that ozone exposure reduced the productivity of agricultural workers. Using data from Mexico City, Hanna & Oliva (2015) show that air pollution can impact hours worked. Chang *et al.* (2016) shows that day-to-day variation in particulate pollution exposure lowered the productivity of pear packers. Their estimates suggest that the relatively small reductions in PM2.5 particulates achieved in the U.S. from 1999-2008 generated $16.5 billion in labor cost savings. Chang *et al.* (2014) uses evidence from call-center workers in China to show that the productivity effects of air pollution exposure extend even to white-collar jobs. In addition, early-life pollution exposure has been linked to a range of negative outcomes, including on cognitive ability (Lavy *et al.* (2014)) and

[10]See Rückerl *et al.* (2011) for a review of the epidemiological literature on this topic.

adult earnings (Isen *et al.* (Forthcoming)).

The pollution released by coal burning factories in 19th century Britain was widely recognized and discussed. For example, *The Times* (Feb. 7, 1882, p. 10)[11] wrote,

> *There was nothing more irritating than the unburnt carbon floating in the air; it fell on the air tubes of the human system, and formed a dark expectoration which was so injurious to the constitution; it gathered on the lungs and there accumulated.*

While pollution in London was more likely to be experienced by visitors and noted by the press, coal smoke pollution was particularly severe in the industrial cities of England. For example, describing a visit to Northwest England in 1890, Cannon Hardwicke Drummond Rawnsley wrote,

> *...chimneys, solid and square, were belching forth clouds of Erebean darkness and dirt...The heavens were black with smoke, and the smother of the mills, to one whose lungs were unaccustomed to breathing sulphurised air, made itself felt.*

While the health effects of air pollution were not fully understood, there was an appreciation for the link between coal-based air pollution and poor health. This is illustrated in the 1867 report of the Sanitary Association by Dr. M.A. Morgan on Manchester (reported by *The Times*, April 16, 1867), which stated that,

> *the chief solid impurity of the air of Manchester is coal smoke, and its mode of action on the human body is of two kinds – (1), as an irritant to the lungs, producing bronchitis or assisting in the production and maintenance of this disease...and (2), by its sulpheretted hydrogen reducing the tone of the system, and rendering it easily susceptible to zymotic [infectious] diseases.*

[11] Quoted from Troesken & Clay (2011). See that paper and Thorsheim (2006) for many other examples.

Further evidence was provided by the wave of deaths that occurred when anticyclone weather systems resulted in particularly dense fogs, such as in 1880, when the British Medical Journal estimated that over 1000 excess deaths occurred in London in one week.[12]

An important feature of this empirical setting is that there was a substantial amount of population mobility in England during this period, with large flows of population from rural areas as well as Ireland and Scotland into English cities.[13] This means that, when considering factors that influence city employment or population growth, the marginal mover that we should have in mind was someone outside of the cities who was faced with a decision about where to migrate. The search for work was the primary driver of these migration flows, though there is also some evidence that pollution levels affected location decisions, both within and across cities, as early as the 1840s.[14] For example, in the 1880s Robert Holland wrote that, "[t]he rich can leave the sordid city and make their homes in the beautiful country...the poor cannot do so. They must breath the stifling, smoky atmosphere..."[15]

Another important feature of this setting was the limited level of government regulation, including both pollution regulation and other regulations that would have affected city growth. While some steps were taken to regulate industrial pollution, these efforts often ran up against the *laissez faire* ideology that dominated British policy during this period as well as the political power of mill owners. New pollution

[12] For more information on the excess deaths that occurred in 1880 see the *British Medical Journal*, Feb. 14, 1880, p. 254. Beyond the health effects, coal smoke also had a myriad of other visible consequences. White cloths became stained and went out of style. Visibility was often so reduced that it caused traffic accidents. There is even evidence that pollution had evolutionary effects. Kettlewell (1955) describes how the Lepidoptera moths, originally white, evolved to take on a dark gray color in order to blend into the polluted forests near the northern industrial cities.

[13] Long & Ferrie (2003) describe this population as "highly mobile", while Baines (1985) suggests that internal migration accounted for roughly 40% of the population growth of British cities during this period. Only one city in the analysis database, Bath, did not experience substantial growth during the study period.

[14] Engels (1845), wrote, "These east and north-east sides of Manchester are the only ones on which the bourgeoisie has not built, because ten or eleven months of the year the west and south-west wind drives the smoke of all the factories hither, and that is for the working people alone to breath."

[15] Quoted from Thorsheim (2006), p. 44.

regulations were passed, including The Sanitary Act of 1866, The Public Health Act of 1875 , and The Public Health (London) Act of 1891. However, these acts allowed for substantial interpretation, contained important loopholes, and imposed relatively small fines.[16] As a result, historical evidence suggests that their effectiveness was limited, though they may have had more impact toward the end of the nineteenth century.[17] Other regulations affecting city growth, such as zoning laws, were also largely absent from this setting, which provides a particular clean opportunity for investigating the impact of pollution on city growth.[18]

Finally, Britain already had a well-developed transportation network by the beginning of the period studied here, including railroad connections to all of the analysis cities, as well as numerous canals and turnpikes. Thus, changes in internal transportation costs were modest during the study period, and evidence suggests that these changes had little impact on the location of industries (Crafts & Mulatu (2006)).

3 Data and measurement

The first key piece of data for this study is a measure of local industrial composition. These data come from the Census of Population, which reports the occupation of each person at each ten-year census interval from 1851-1911 for 31 of the largest cities in England.[19] The occupational categories reported in these data generally

[16]One example provided by Thorsheim (2006) is that the acts regulated "black smoke" and that defendants were able to avoid fines by claiming that their smoke was merely "dark brown."

[17]See, e.g., Thorsheim (2006) and Fouquet (2012).

[18]For example, no national zoning law existed in England until 1909. There were also very few place-based policies of the kind found in many modern economies, and little spatial redistribution of wealth through national taxes.

[19]The set of cities in the database includes all of the English cities for which city-level occupation data were reported by the Census for each decade from 1851-1911. These were the largest cities in England in 1851 with the exception of Plymouth, which is excluded because changes to the city border make it impossible to construct a consistent series for that city. Figure 3 in the Appendix includes a map of these cities. This study uses the most recent version of the database (v2.0) which was updated in March, 2016. The data, additional documentation, and descriptive statistics can be found at `http://www.econ.ucla.edu/whanlon` under Data Resources.

closely correspond to industries, such as cotton spinner or steel manufacturer.[20] To construct consistent series for 1851-1911, I combine the many occupational categories available in each census into a set of 26 broad industries, spanning nearly the entire private-sector economy.[21]

Because I am working with fairly aggregated industry categories, almost all industries are present in all cities.[22] However, the spread of industries across cities was far from even. For example, textile producers agglomerated in cities in Lancashire and Yorkshire, where they could account for as much as half of all private-sector employment. Cities such as Sheffield, Birmingham and Wolverhampton had a disproportionate share of metals industries, while ports such as Bristol and Liverpool had high shares of transportation and services.

The second necessary piece of information for this study is a measure of the coal intensity of each industry. This information is drawn from the first Census of Production, which was completed in 1907.[23] This Census collected detailed information on the amount of coal used in each industry, as well as industry employment, allowing me to construct a measure of coal use per worker in each industry.[24]

These data show that coal use intensity varied enormously across industries, a feature that plays a key role in this study. A table describing coal use intensity by industry is available in Appendix A.2.5. The most intensive industrial coal users, such as metal & machinery or earthenware & bricks, used coal to heat material up to high temperatures. These industries used more than forty tons per worker per year. Textiles, a moderate coal-using industry which consumed around ten tons per

[20]One unique feature of this data source is that it comes from a full census rather than a sample. This is helpful in reducing the influence of sampling and measurement error.

[21]A list of the industries included in the database is available in Appendix A.2.5.

[22]The exceptions are a few cities which have no employment in shipbuilding or mining. Observations with no city-industry employment are dropped from the analysis, leaving me with a slightly unbalanced panel.

[23]While these data come from near the end of the study period, this is the earliest available consistent source for this information.

[24]Coal and coke are combined in this study. Coke consumption was small relative to coal.

worker per year, generally used coal to power steam engines. Other industries, such as apparel or tobacco products, used very little coal, less than two tons per worker per year. This large variation in coal use intensity at the industry level, together with the tendency of industries to agglomerate in particular locations, resulted in substantial variation in the amount of industrial coal use at the city level.

I model industrial coal use in cities as determined by city-industry employment (L_{ict}), the coal use intensity of each industry (θ_i), and the national efficiency of coal use per worker, ρ_t:

$$COAL_{ct} = \rho_t \sum_i \left(L_{ict} * \theta_i \right) . \tag{1}$$

Estimates of θ_i for manufacturing industries are provided by the 1907 Census of Production, while Census of Population data provide city-industry employment. The ρ_t term can be calculated by comparing data on industrial coal use at the national level to the values obtained using data on θ_i and L_{it}.[25] In general, other industries, such as services, were not likely to be major coal users, so this measure should capture most industrial coal use.[26]

One assumption implicit in this approach is that *relative* coal use per worker across industries did not vary too much over time. While I cannot check for this over

[25]Specifically, I use the fact that $\ln(\rho_t) = \ln(COAL_t) - \ln\left(\sum_c \sum_i L_{ict} * \theta_i\right)$. In this equation, the $\sum_c \sum_i L_{ict} * \theta_i$ term can be calculated from the data, while national coal use in industry is available from Mitchell (1988). In practice, the inclusion of the ρ_t term will not affect the estimated coefficients because regressions are run on the log of coal use across locations within a period and the ρ_t will be absorbed by year effects. However, this term will affect the overall impact of coal use on city growth.

[26]An exception is local utilities, particularly gas, which was a major user of coal. Coal was used to make gas, which was then pumped to users in the city, where it was burned for light or heat. Despite the fact that local utilities used coal, I exclude local utility coal use from the pollution measure because gas providers may have reduced the amount of coal smoke residents were exposed to if the gas replaced more polluting forms of energy use in homes and offices. Another potential exception is transportation, particularly rail transportation, which used a substantial amount of coal. However, most of this coal would have been burned outside of stations, spreading it though the countryside. This makes it very difficult to determine the location of pollution related to coal use in the transportation sector. Thus, I also exclude transportation from the local coal use measure.

the study period, it is possible to check the extent to which industry coal use varies over time using data from the 1924 Census of Production, the next full production census after 1907. Comparing coal use per worker in industries in 1924 to the same values in 1907 provides an assessment of how rapidly these industry features could change. This analysis, described in Appendix A.2.6, shows two results. First, the relative coal use intensity across industries was quite stable over time.[27] This is comforting, particularly because the 1907-1924 period saw larger changes in the source of factory power, due to the introduction of electricity, than did the 1851-1907 period. Second, comparing 1907 and 1924 coal use per worker suggests that there was broad improvement in coal use efficiency over time which occurred relatively evenly across industries. This type of efficiency improvement will be captured in the ρ_t term.

Another assumption implicit in these coal use measures is that industry coal use does not vary too much across locations in response to variation in the relative level of wages or coal prices. Put another way, it is important that variation in city coal use due to local industry composition and differences in industry coal use intensity resulting from technological factors is substantially more important than the variation due to differences in the local prices of coal or other inputs. The enormous variation in coal use intensity across industries is important for making this a reasonable assumption.

One way to check both of these assumptions is to compare estimated levels of coal use calculated using the method described here to data on local coal use levels. While such data are generally unavailable, there is information on county-level coal use in the 1871 Coal Commission report. Comparing estimates of industrial coal use at the county level for 1871, based on the approach I have just described, to county-level coal use data from the 1871 report shows that my approach does a good job of replicating industrial coal use at the county level (the correlation is 0.912), particularly for more

[27] A regression of coal use per worker in 1924 values on coal use per worker in 1907 yields a coefficient of 1.021 with a s.e. of 0.061 and an R-squared of .949.

industrial and urbanized locations. The full analysis is available in Appendix A.2.7.

Estimates of industrial coal use per worker at the city level are described in Table 7 in Appendix A.2.4. These data show that there was substantial variation across cities in the expected level of coal use per worker, even among similarly sized cities. Sheffield, often cited as the prototypical polluted industrial city, emerges as the most intensive user of coal in the database, followed by other cities specializing in metals such as Birmingham and Wolverhampton. Textile manufacturing towns, such as Manchester and Leeds, show moderate levels, near the average. Commercial and trading cities, such as Liverpool and Bristol, as well as London, use industrial coal less intensively. Bath, a resort town, is the least polluted city in the database.

4 Theory

This section presents a spatial equilibrium model in the Rosen-Roback tradition, but modified in a few important ways in order to fit the empirical setting. The economy is made up of a fixed number of cities, indexed by c. These cities are small open economies that take goods prices as given. As is standard in spatial equilibrium models, workers and firms can move freely across cities and goods are freely traded. I begin by modeling the demand for labor in cities.

The economy is composed of many industries, indexed by i, each of which produce a homogeneous good. Each industry is composed of many perfectly competitive firms, indexed by f. Firms produce output using labor, a polluting input (coal), and a fixed local industry-specific resource.[28] The production function is,

$$y_{fict} = a_{ict} L_{fict}^{\alpha_i} C_{fict}^{\beta_i} R_{fict}^{1-\alpha_i-\beta_i} ,$$

[28] In Appendix A.3.3 I consider a model that also incorporates capital. This does not alter the basic estimating equation derived from the model, but it does influence how the estimation results are interpreted relative to the model parameters.

where L_{fict} is labor, C_{fict} is coal, R_{fict} is a local resource, and a_{ict} is the local productivity level in industry i. Let $\alpha_i, \beta_i \in [0,1)$ for all i, and $\alpha_i + \beta_i < 1$ for all i. Note that the production function parameters are allowed to vary at the industry level. This will result in industries employing different input mixes, with some using coal more intensively than others.

Local resources are fixed within each city and are industry-specific, with an available supply given by $\bar{R}_{ic}$.[29] These resources can be thought of as natural features or local endowments of entrepreneurial ability in a particular sector. They play an important role in the model; by introducing decreasing returns at the city-industry level, they allow multiple cities to be active in an industry even when productivity varies across cities, trade is costless, and markets are perfectly competitive.

Firms maximize profit subject to output prices p_{it}, the coal price ϕ_t, a city wage w_{ct}, and the price of local resources χ_{ict}. The firm's maximization problem in any particular period is,

$$\max_{L_{fict}, C_{fict}, R_{fict}} p_{it} a_{ict} L_{fict}^{\alpha_i} C_{fict}^{\beta_i} R_{fict}^{1-\alpha_i-\beta_i} - w_{ct} L_{fict} - \phi_t C_{fict} - \chi_{ict} R_{fict} \, .$$

Using the first order conditions from this problem, I obtain the following expression for the relationship between employment and coal use in each industry,

$$\frac{C_{ict}}{L_{ict}} = \left(\frac{\beta_i}{\alpha_i}\right)\left(\frac{1}{\phi_t}\right) w_{ct} \, . \tag{2}$$

This expression tells us that variation in the use of polluting inputs across industries will be governed in part by the industry-specific production function parameters α_i and β_i. The empirical analysis exploits the exogenous variation due to the β_i/α_i parameters, reflected by the θ_i term in Eq. 1, while abstracting from the variation

[29] This approach emphasizes natural advantages, which have been shown to be important determinants of industry location (Ellison & Glaeser (1999)). This type of approach has recently been used in papers by Kline & Moretti (2013), Kovak (2013) and Hanlon & Miscio (2014).

due to the endogenous w_{ct} term. The $(1/\phi_t)$ term in Eq. 2 implies that coal use per worker can vary over time in a way that is common to all industries, a feature that is reflected in the ρ_t term in Eq. 1.

Digressing briefly, it is worth emphasizing that the fact that the expression in Eq. 2 maps directly into the coal use values calculated using Eq. 1, and the fact that those coal use values do a good job of reproducing observed coal use levels in 1871 (see Appendix A.2.7), suggests that it is reasonable to apply the functional form used in the model across the study period. Put another way, if the model were a poor approximation of the world, then we would not expect coal use estimates based on the structure of the model to do a reasonable job of matching the observed data. Furthermore, the results in Appendix A.2.6 suggest that the patterns of change observed from 1907-1924 are also consistent with Eq. 2.

Using the first order conditions from the firm's maximization problem, and summing across all firms within an industry, I obtain the industry labor demand equation:

$$L_{ict} = \alpha_i^{\frac{1-\beta_i}{1-\alpha_i-\beta_i}} \left(a_{ict}p_{it}\right)^{\frac{1}{1-\alpha_i-\beta_i}} \left(\beta_i/\phi_t\right)^{\frac{\beta_i}{1-\alpha_i-\beta_i}} w_{ct}^{-\frac{1-\beta_i}{1-\alpha_i-\beta_i}} \bar{R}_{ic} \,. \tag{3}$$

Note that, in equilibrium, the sum of firm resource use must equal total city-industry resources, which are fixed at $\bar{R}_{ic}$.

One congestion force in the model is the limited supply of housing. The housing market itself is not a central focus of this paper, so I model housing in a reduced-form way,

$$\ln(r_{ct}) = \lambda \ln(L_{ct}) + \ln(\eta_c) \,, \tag{4}$$

where r_{ct} is the rental rate, L_{ct} is total city population, η_c represents fixed city-specific factors that influence construction costs, and $\lambda > 0$ is a parameter that determines the impact of increasing population on the housing price.[30]

Now, we turn to the supply of labor in a city. The model is populated by a continuum of homogeneous workers, each of which supply one unit of labor to the market. Workers consume a basket of goods with price P_t and housing. They also benefit from local amenities. The workers' indirect utility function is,

$$V_{ct} = \gamma \ln \left(\frac{w_{ct}}{P_t} \right) + (1-\gamma) \ln \left(\frac{w_{ct}}{r_{ct}} \right) + \ln(A_{ct}) \,.$$

where w_{ct} is the wage, A_{ct} is the amenity value, and the $\gamma \in (0,1)$ parameter determines the relative expenditure shares of housing and goods.

Workers are freely mobile across cities and have an outside option utility $\ln(v_t^*)$ in each period. In the empirical setting I consider, this can be thought of as either the utility of emigrating or the utility of living in the rural areas of the country. Given this, and using Eq. 4, the inverse labor supply equation for city c is,

$$w_{ct} = P_t^{\gamma} \, L_{ct}^{(1-\gamma)\lambda} \, \eta_c^{1-\gamma} \, A_{ct}^{-1} \, v_t^* \,. \tag{5}$$

In addition to workers, the model is also populated by capitalists who receive the rent from land and local resources. For simplicity, I assume that capitalists live and spend their income outside of the city.

Next, I want to incorporate the impact of local industrial pollution into the model. Coal pollution can impact the city by affecting both workers and firms. Focusing first on residents, I express the local amenity value as,

[30] This expression is similar to that used in previous work (e.g., Moretti (2011)) except that the elasticity of housing supply λ does not vary across cities. While this assumption is likely to be unrealistic in modern settings because of variation in zoning laws or other regulations, it is more reasonable in the empirical setting I consider. This is due in part to the lack of land-use regulations in the period I study and in part to the relatively homogeneous geography across English cities (compared to, say, U.S. cities).

$$A_{ct} = \delta_c \, C_{ct}^{-\psi} \epsilon_{ct}^{A} \,, \tag{6}$$

where C_{ct} is city coal use, δ_c represents a fixed city amenity, the ψ parameter determines the impact of local coal use on the amenity level, and ϵ_{ct}^{A} represents an idiosyncratic shock to the local amenity level.

Coal use can also affect the productivity of local firms. To build this channel into the model, I assume that local industry productivity can be separated into the impact of national changes in industry productivity, a_{it}, the impact of city-level coal use on firm productivity, $C_{ct}^{-\nu}$, where the parameter $\nu \geq 0$ determines the impact of local coal use on firm productivity, and an idiosyncratic shock to city-industry productivity, ϵ_{ict}^{P}. Thus, I have $a_{ict} = a_{it} C_{ct}^{-\nu} \epsilon_{ict}^{P}$.

Given the outside option utility, the national coal price, a set of national industry output prices, technology levels, and city industry resources, equilibrium in a city is defined as the set of local wages, resource prices, housing rent, and population, together with a set of industry employment and coal use levels, such that firms maximize profits, the local markets for resources clear, the housing market clears in each city, and city labor supply equals city labor demand.

For the empirical analysis, I need an expression that relates the growth in local industry employment to changes local industrial pollution. The starting point for this derivation is the industry labor demand expression given in Eq. 3 and the city labor supply expression in Eq. 5. Differencing these expressions over time, taking logs, and substituting out the wage terms, I obtain,

$$\begin{aligned}
\Delta \ln(L_{ict}) \quad = \quad & \left(\frac{-(1-\gamma)(1-\beta_i)\lambda}{1-\alpha_i-\beta_i} \right) \Delta \ln(L_{ct}) + \left(\frac{-\psi(1-\beta_i)-\nu}{1-\alpha_i-\beta_i} \right) \Delta \ln(C_{ct}) \\
- \quad & \left(\frac{1}{1-\alpha_i-\beta_i} \right) \Big[\beta_i \Delta \ln(\phi_t) + (1-\beta_i)\gamma \Delta \ln(P_t) + (1-\beta_i) \Delta \ln(v_t^*) \\
- \quad & \Delta \ln(a_{it} p_{it}) - \Delta \ln(\epsilon_{ict}^{P}) + (1-\beta_i) \Delta \ln(\epsilon_{ct}^{A}) \Big] \,.
\end{aligned} \tag{7}$$

Eq. 7 forms the basis for the main empirical specifications used in this paper. The $\Delta \ln(L_{ct})$ and $\Delta \ln(C_{ct})$ terms on the right-hand side of this equation capture, respectively, the impact of city congestion and of city coal use. The model suggests that both of these will negatively impact city-industry employment growth, though it is worth noting that the impact of city size may be positive if a city-size agglomeration force is included in the model.[31] In the middle row of Eq. 7 is a set of terms that vary only over time, but not across space. These will be absorbed by year effects in the empirical analysis. On the bottom row of Eq. 7, the first term reflects national industry-level demand or productivity shocks, the building blocks of the Bartik instrument. These can be absorbed by industry-time effects in the main analysis. The final two terms on the bottom row of Eq. 7 are the error terms. The structure of these terms makes it clear that I should allow for correlated errors across industries within the same location and time period in the empirical analysis.

The focus of the empirical analysis will be estimating the coefficient on the coal use and city-size terms in Eq. 7. As Eq. 7 shows, the impact of either coal use or congestion is determined by a combination of several model parameters. In the empirical analysis, I will estimate a single coefficient reflecting how, together, these parameters govern the relationship between either congestion or coal use and city growth, but I will not be able to identify the component parameters individually. For further discussion of this expression and its link to the coefficients estimated in the empirical analysis, see Appendix A.3.2.

In Appendix A.3.1 I provide some additional analysis relating the estimation approach suggested by Eq. 7 to the larger Bartik instrumentation literature and exploring the possibility of aggregating the estimating equation to the city level. This shows that, even when ignoring heterogeneity in the industry production function, it

[31] I have avoided adding a city-size agglomeration force to the model simply because it complicates the equilibrium conditions.

is not possible to derive a city-level estimation equation from the theory. This highlights the advantages of the industry-level estimation approach used in this paper and illustrates some of the potential issues that occur when mapping the city-level Bartik instruments used in many existing papers to microfoundations.

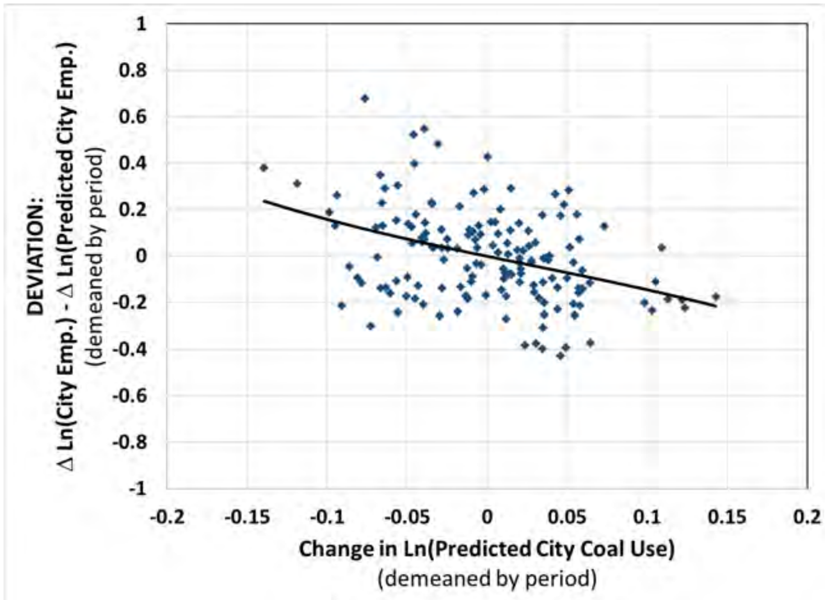
5 Analysis

This section begins with an analysis of the impact of coal use on local employment growth, first at the level of city-industries and then at the city level. These are the central results of the paper. Following that, I present a simple counterfactual that can help us think about the implications of coal use for overall urbanization levels. Finally, I provide some tentative evidence on the channels through which coal use may have affected city growth.

Before moving to the main analysis, I begin with some reduced-form empirical patterns at the city level, in Figure 2. The y-axis of this graph describes the difference between actual city employment growth, $\Delta \ln(L_{ct})$ and predicted city employment growth based on each city's initial industry composition and the growth rates of each industry in all other cities. This is taken across every two-decade period from 1851-1911 and then demeaned by period (similar patterns also appear for each period). The x-axis of Figure 2 describes the predicted change in city coal use, based on each city's initial industry composition, the growth rate of each industry in all other cities, and the coal use intensity of each industry. These values are taken over the same 20 year periods and also demeaned by year.

What we can see in Figure 2 is that, in cities and time periods where we predicted a greater increase in coal use, overall city employment growth is systematically lower than what we would expect given each city's initial industrial composition and national industry growth rates. This provides a first piece of reduced-form evidence suggesting that rising coal use may have lowered city employment growth.

Figure 2: Deviation vs. predicted change in city coal use



The y-axis is the difference between actual city employment growth over each two-decade period in city c and the predicted employment growth in that city-industry based on each city's initial employment by industry and employment growth in each industry in all other cities, summed across industries. The x-axis is the predicted change in city-level industrial coal use over the period, which is generated using the initial composition of city-industries interacted with national industry growth rates and measures of industry coal use per worker. The trend line is based on a third-order polynomial.

5.1 Coal use and city-industry employment growth

The starting point for the main analysis is Eq. 7. Converting this to a regression form, I have,

$$\Delta \ln(L_{ict}) = b_0 + b_1 \Delta \ln(C_{ct}) + b_2 \Delta \ln(L_{ct}) + \xi_{it} + e_{ict} \, , \qquad (8)$$

where the ξ_{it} is a set of industry-time effects which absorb the national-level factors in Eq. 7 as well as the industry-specific productivity and demand shocks, while e_{ict}

incorporates the idiosyncratic shocks to city amenities and city-industry productivity.

It is clear that a regression implementing Eq. 8 will suffer from serious identification issues. In particular, both the change in overall city employment and the change in city coal use will be endogenously affected by city-industry employment growth.

To deal with this, I replace these terms with predicted values. For overall city employment, let,

$$\Delta \ln(PrCityEMP_{ct}) = \ln\left(\sum_{j\neq i} L_{jct-\tau} * GR_{j-ct,t-\tau}\right) - \ln\left(\sum_{j\neq i} L_{jct-\tau}\right)$$

where $GR_{i-ct,t-\tau}$ is the growth rate of industry i in all cities other than c from $t-\tau$ to t. In this expression, τ determines the size of the time period over which differences are taken. I will explore differences ranging from one to three decades. This variable represents the expected growth in employment in all other local industries, given national industry growth rates and the initial industrial composition of the city. Note that, when studying industry i, that industry is dropped when constructing $\Delta \ln(PrCityEMP_{ct})$.[32] This helps us avoid endogeneity concerns, but ultimately it does not have a substantial impact on the results.

Next, to reflect the predicted change in city coal use, I define,

$$\Delta \ln(PredCoal_{ct}) = \ln\left(\sum_{j\neq i} L_{jct-\tau} * GR_{j-ct,t-\tau} * \theta_j\right) - \ln\left(\sum_{j\neq i} L_{jct-\tau} * \theta_j\right) .$$

where θ_j is coal use per worker in industry j. Note that industry i is dropped when calculating $\Delta \ln(PredCoal_{ct})$ in order to avoid endogeneity concerns.[33] It is important to note that the difference between $\Delta \ln(PredCoal_{ct})$ and $\Delta \ln(PrCityEMP_{ct})$ is due

[32]In practice this will cause $\Delta \ln(PrCityEMP_{ct})$ to also vary at the industry level, but, with a slight abuse of notation I do not include i in the subscript in order to make it clear that this variable is capturing a city-level effect.

[33]In practice this will cause $\Delta \ln(PrCityEMP_{ct})$ to also vary at the industry level, but, with a slight abuse of notation I do not include i in the subscript in order to make it clear that this variable is capturing a city-level effect.

only to variation in the coal intensity of industries, represented by θ_j.[34]

Putting these elements together, the main regression specification is,

$$\Delta \ln(L_{ict}) = b_0 + b_1 \Delta \ln(PredCoal_{ct}) + b_2 \Delta \ln(PrCityEMP_{ct}) + \xi_{it} + e_{ict}\,. \quad (9)$$

This specification addresses the most important identification concerns in Eq. 8, i.e., the endogenous affect of city-industry employment growth on city-level congestion and coal use. The main identification assumption in this analysis approach is that there is not some other industry feature that is both correlated with industry coal use intensity and affects local employment growth. This type of assumption is typical across the entire Bartik instrument literature. After presenting the main regression results, I present a variety of additional results including controls for the most likely channels through which the identification assumption might be violated.

The specification in Eq. 9 incorporates an assumption about the functional form for the effects of local coal use. In particular, it suggest that the impact of coal use is linear in logs. While empirical evidence on the form of this relationship is sparse, there are two pieces of evidence supporting this functional form. First, the scatterplot shown Figure 2 suggests that the relationship between predicted coal use and the extent to which city employment growth underperformed predicted city employment growth is close to linear in logs. Second, Beach & Hanlon (2016) provides evidence that the impact of coal use on mortality is linear in logs. To the extent that the mortality rate is a good indicator of the impact of coal use this suggests that the specification used here is reasonable.

Note that Eq. 9 abstracts from heterogeneous industry responses to changing

[34] Because of the way the $\Delta \ln(PrCityEMP_{ct})$ and $\Delta \ln(PredCoal_{ct})$ variables are constructed, there is likely to be a positive correlation between these variables. However, when taking differences the correlation between these variables is generally not too high. In particular, the results in Appendix A.4.2 show that for two-decade differences the correlation between these variables is 0.284 when all industries are included, and -0.0845 when focusing only on manufacturing industries. These should not be high enough to create substantial problems for the regression results.

levels of city pollution or city congestion forces, a feature suggested by the theory. While I begin the analysis by abstracting from heterogeneity in the response to coal use across industries, later I will also present results that explore these heterogeneous responses.

In relation to the theory, the estimated b_1 coefficient from Eq. 9 will reflect the impact of changes in local industrial coal use on city-industry employment growth, which will depend on how coal use affects the city amenity level, how coal use affects firm productivity, as well as the extent to which industries can respond to these effects by shifting employment away from polluted locations.[35] The theory suggests that this coefficient should be negative. Note that, because $\Delta \ln(PrCityEMP_{ct})$ is also included in the regression specification, the b_1 coefficient should be interpreted as the impact of a rise in local industrial coal use holding constant the overall local employment level, i.e., as an increase in the pollution intensity of local industry. Similarly, the b_2 coefficient should be interpreted as reflecting the impact of an increase in local employment holding fixed the level of local industrial coal use, i.e., a rise in completely clean employment.[36]

This estimation approach abstracts from variation in industry coal use intensity across cities. This is driven in part by data constraints, since city-specific industry coal use intensities are not observed. However, even if city-level industry coal use intensity was observed, I would probably not want to incorporate this into the explanatory variable because, as suggested by the theory, this value will be endogenous and dependent on local wage levels. Abstracting from spatial variation in industry coal use intensity avoids this endogeneity concern. Moreover, because of the very large variation in coal use intensity across industries, as well as the substantial vari-

[35] In the model, the ability of industries to shift production away from more polluted locations depends on the importance of fixed local resources in production. For further discussion of the link between the estimated coefficients and the theory, see Appendix A.3.2.

[36] In the theory, the sign of the b_2 coefficient is predicted to be negative. However, I have not included a city-size agglomeration force in the model. If this is included, then the sign of b_2 may be positive or negative.

ation in industry employment shares across cities, my measure is likely to capture most of the relevant variation in coal use across cities.

Estimation is done using pooled cross-sections of data (after taking differences), an approach that allows me to exploit as much of the available data as possible. This is vital because the key variation in this study occurs at the city level and only 31 cities are observed in the data. We may be concerned that industries within a city suffer from correlated standard errors and that standard errors may be correlated over time. To deal with this issue, I allow correlated standard errors across industries within the same city, following Conley (1999) and across time within the same city-industry, as in Newey & West (1987).[37]

I begin the analysis, in Table 1, by exploring results with differences taken over time periods ranging from one to three decades. The table includes results for all industries, in Columns 1-3, and for a set of manufacturing industries only, in Columns 4-6. I provide separate results for manufacturing industries only because these produce more tradable products and so are a better fit for the model, and also because some of the control variables that I will introduce later are available for only this set of industries.

Table 1 reveals several important patterns. The most important result for this study is that the coal use variable always has a negative impact on city-industry employment growth. This impact is clearer when we look over longer time differences, and becomes statistically significant for differences of two or three decades. The fact that the estimated impact grows over time suggests that it may take time for workers

[37]The theory suggests that errors may be correlated across industries within a city and time period as a result of the ϵ_{ct}^{A} term. For lag lengths over one there will mechanically be serial correlation in these regressions because the differences will overlap. Thus, it is important to allow for serial correlation at least equal to the lag length. An examination of alternative approaches to treating the standard errors shows that allowing correlated standard errors across industries within the same city has by far the largest impact on the standard errors. Once this type of correlation is allowed, extending the standard errors to allow correlation across industries in nearby cities (e.g., within 10km or 50km) does not lead to any substantial additional increase in the confidence intervals. To implement these standard errors, I follow Hsiang (2010).

and firms to react to the impact of rising industrial coal use. These delays are not surprising given that we are looking at reactions to changes in coal use levels, and that these reactions will involve changes in fixed capital investment patterns and costly migration decisions. In addition, if the effects of coal use are coming in part through productivity impacts on workers (as suggested by the analysis in Section 5.3) then it is likely to take time for these effects to emerge.

Table 1: Baseline city-industry regression results

	DV: Δ Log of city-industry employment					
	All industries			Manufacturing industries		
Difference:	One decade	Two decades	Three decades	One decade	Two decades	Three decades
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta Ln(PredCoal)$	-0.611	-1.987***	-3.016***	-0.444	-2.218***	-3.257***
	(0.621)	(0.732)	(0.803)	(0.685)	(0.632)	(0.813)
$\Delta Ln(PrCityEMP)$	-0.536	0.392	1.362*	-0.725	0.383	1.172*
	(0.586)	(0.757)	(0.826)	(0.528)	(0.553)	(0.692)
Constant	0.568***	1.284***	2.544***	0.338	1.907***	2.883***
	(0.158)	(0.199)	(0.373)	(0.217)	(0.332)	(0.481)
Ind.-time effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,809	4,012	3,208	2,773	2,312	1,849
R-squared	0.259	0.355	0.429	0.246	0.336	0.403

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across a number of decades equal to the lag length. All regressions use data covering each decade from 1851-1911. The regressions for all industries include 26 private sector industries spanning manufacturing, services, transport, and utilities. The results for manufacturing industries are based on 15 industries.

Table 1 also provides evidence of a negative shorter-run effect of employment growth in other city-industries that becomes positive over longer periods. This pattern is consistent with a city-size congestion force that weakens over time, together with positive city-size agglomeration benefits. This is reasonable if we think that there are some city features, such as infrastructure, that are difficult to adjust in the short-run but can be expanded in the long-run. Finally, it is worth noting that the R-squared values increase as we move to longer differences. This suggests that city-industry employment growth may be subject to idiosyncratic short-run shocks, but that longer-run growth patterns are more closely tied to predictable influences.

Later, I will discuss in more detail the magnitude of the coal use effects documented in Table 1, but before doing so it is useful to discuss some additional robustness results. Table 2 present the coefficient on the change in log coal use for a variety of robustness results (full results are in Appendix A.4.3). In the top panel, Columns 1-2 present results with a variety of city-level controls (these are listed in the table comments). Of the available controls, I find that cities with higher levels of initial innovation (based on patenting) and better access to coal reserves grew more rapidly, while larger cities and those with more rain or colder temperatures grew more slowly.[38] These patterns seem quite reasonable. Columns 3-4 present results from regressions including city fixed effects. These results make it clear that the patterns I document are not simply driven by a few slow-growing cities.[39] Columns 5-6 present results obtained while dropping London, the largest outlier city in the data. Finally, Columns 7-8 present results including as a control log employment in each city-industry at the beginning of each period.[40] Overall, my basic results do not appear to be sensitive to these alternative specifications.

In the middle panel of Table 2, I present results including a set of controls based on industry characteristics, which are available only for manufacturing industries. These controls directly address the main identification concern, i.e., that there may be some other industry characteristic that is correlated with coal use and affects city employment growth. The control variables that I have constructed are the share of

[38]One reason for including the coal proximity and weather variables is that they are the key factors determining residential coal use levels. Thus, controlling for these helps me to control for the effect of residential pollution.

[39]Further evidence on this is provided in Appendix Table 17, where I estimate the relationship between the lagged predicted change in city coal use and current city-industry employment growth. These results show that changes in city coal use in a previous period is not strongly related to city employment growth later on, and that including this variables does not substantially affect the estimated relationship between the predicted change in coal use in the current period and employment growth.

[40]The full results, in Appendix Table 16, show that initial city-industry employment is associated with slower subsequent city-industry growth.

(high skilled) salaried to (lower skilled) wage workers, average firm size, the share of output exported, the labor cost share, the female worker share and the youth worker share.[41] These reflect factors that are commonly cited by urban economics as affecting city growth. These industry characteristics are used to construct city-level changes using the exact same approach that was used to construct changes in city coal use using the industry coal per worker data. These variables are then included as controls in the regressions in Columns 9-15. Including these variables does not meaningfully affect my main results.[42]

In the bottom panel of Table 2, I include controls based on connections between industries, through input-output channels or labor force similarity. Recent work by Ellison *et al.* (2010) suggests that these may be an important channel for inter-industry agglomeration forces. The controls I use reflect, for each industry, the change in local employment in buyer industries, supplier industries, or industries employing workforces that are demographically or occupationally similar. The results in Columns 18-20 show that including these controls does not alter the main results. Finally, in Columns 21-22, I add controls for the initial level or the change in the rate of violence and industrial accidents in each city based on mortality data. This addresses concerns that workers in more coal-intensive industries could have brought other undesirable features, such as a propensity for crime, or that coal-using industries could have been more hazardous for workers.

[41]The data used to construct these controls are described in Appendix A.2.3. We may also be concerned that industries differ in the intensity with which they use land. Unfortunately, I have not been able to find a suitable measure of the land use intensity of industries in this period. However, the fact that I find that the impact of coal use is higher in industries with a greater labor cost share (see Appendix A.4.4), which is likely to mean a lower land cost share, suggests that land values are unlikely to be driving the results.

[42]Of the available controls, only industry labor cost share strongly predicts city growth. This likely reflects the relatively fast growth or services that took place during this period. Full results are available in Appendix Table 18.

Table 2: Robustness regression results with two-decade differences

	DV: Δ Log of city-industry employment							
	With additional controls		With city fixed effects		Dropping London		Initial ind. size controls	
	All ind.	Manuf.	All ind.	Manuf.	All ind.	Manuf.	All ind.	Manuf.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta Ln(PredCoal)$	-1.526**	-1.151*	-1.614***	-1.112*	-1.980***	-2.220***	-2.070***	-2.100***
	(0.696)	(0.622)	(0.586)	(0.614)	(0.740)	(0.670)	(0.737)	(0.637)
Observations	4,012	2,312	4,012	2,312	3,882	2,237	4,012	2,312
	Additional controls based on industry characteristics (manufacturing industries only)							
	Salaried worker shr.	Average firm size	Export shr.	Labor cost shr.	Female worker shr.	Youth worker shr.	All	
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	
$\Delta Ln(PredCoal)$	-2.197***	-2.300***	-2.217***	-2.688***	-2.206***	-1.857***	-2.181***	
	(0.634)	(0.676)	(0.673)	(0.633)	(0.639)	(0.657)	(0.705)	
Observations	2,312	2,312	2,312	2,312	2,312	2,312	2,312	
	Additional controls based on inter-industry connections						Controlling for violence	
	IO in	IO out	Demog. similarity	Occ. similarity	All		Initial level	Change
	(16)	(17)	(18)	(19)	(20)		(21)	(22)
$\Delta Ln(PredCoal)$	-2.135***	-2.149***	-2.149***	-2.179***	-2.093***		-2.076**	-2.140**
	(0.743)	(0.725)	(0.718)	(0.719)	(0.748)		(0.964)	(0.989)
Observations	3,549	3,549	3,549	3,549	3,549		2,411	2,411

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across a number of decades equal to the lag length. All regressions use data covering each decade from 1851-1911 and include the predicted change in city employment as well as industry-time effects. The additional controls in Columns 1-2 are the number of air frost days in each city, rainfall in each city, patents in the city from 1852-1858, log city population at the beginning of each period, the log of city coal use at the beginning of each period, carboniferous rock deposits within 50km and a seaport indicator. Columns 7-8 include controls for initial industry size. The controls in Columns 9-15 are city-level controls based on industry features constructed using the same approach used for city coal use. The controls in Columns 16-20 are for changes in industries sharing buyer or supplier linkages to the observation industry (IO in and IO out) or using demographically or occupationally similar labor forces. The violence controls are based on city-level mortality due to violence or accidents.

As an additional check, in Appendix Table 22 I estimate the impact of coal use separately for five main coal using industries. These results show similar estimated coal use impacts across the different industries. This is comforting, because it suggests that the results I'm obtaining are specifically related to the level of coal use, regardless of which industry it comes from.

I conduct two other exercises to assess the stability and statistical significance of the results. First, in Appendix A.4.7, I undertake a permutation exercise in in which I randomly reassigned the industry coal use per worker values across the 26 analysis industries 1000 times and then re-estimate results using the specification corresponding to Column 2 of Table 1. Comparing the estimated coal use coefficients from these placebo regressions to the coefficient obtained using the true data implies a confidence level of 99.1%. With the full set of city-level controls, the confidence level implied by the permutation test is 93.6%. Second, I re-run the results dropping each of the cities in the data using the specification in Column 2 of Table 1. This yields coefficients ranging from -1.30 to -2.29 with p-values ranging from 0.0018 to 0.0367.

Finally, in Appendix A.4.6 I estimate IV regressions in which the predicted change in local industrial coal use is used as an instrument for the change in local industrial coal use based on actual city-industry growth. The estimated coefficients on coal use obtained from these regressions range from -1.12 to -1.63.

Overall, these results consistently show a negative and statistically significant relationship between city coal use and city-industry employment growth, regardless of whether we are focused on all industries or just manufacturing industries. The magnitude of the estimated coefficients for two-decade differences range from -1.11 to over -2.5, with my preferred estimates, which include the full set of available controls, falling between -1.2 and -1.5. To interpret these estimates, it is useful to know that the average increase in log city coal use across all periods was 0.372 with a standard deviation of 0.176. Given these results, we should expect a city with an increase in coal use that is one standard deviation above the mean to have a reduction in city-industry employment growth of 21-26 percentage points over two decades. Average city-industry employment growth across all cities and periods was 43.7 percent and the standard deviation was 0.52. Thus, a one s.d. greater increase in city coal use would be expected to reduce city-industry employment growth by roughly one-half of either the average or the standard deviation of city-industry growth.

These results imply that rising coal use had a powerful effect on city employment growth. However, it is important to note that cities with rapidly rising coal use were also likely to be those with large shares of industries experiencing rapid growth at the national level, and therefore also likely to experience rapid growth at the city level. These positive fundamentals will obscure the effect of rising coal use in reducing city-industry employment growth. This feature means that it will be difficult to appreciate the impact of industrial pollution on city growth through casual observation alone, which explains why contemporary observers failed to appreciate the magnitude of the negative effect of pollution on local employment growth.

While the results described thus far estimate average effects of coal use across all industries, the theory suggests that these effects are likely to be heterogeneous. In particular, if coal pollution primarily affects workers (through either amenity or productivity channels), then we should expect these effects to be larger for more labor intensive industries. When I run regressions that include the interaction of the coal use variable with industry labor cost share this is what I find.[43] In particular, in the regression results shown in Appendix A.4.4, I observe negative and generally statistically significant coefficients on the interaction between the coal use and industry labor cost share variables.

Next, I shift my attention to estimating the impacts on overall city employment or population. Analyzing city-level results is useful because it allows me to look

[43] The labor cost share variable is the ratio of labor costs to total revenue. This variable is available only for manufacturing industries.

at alternative outcome variables, such as overall city population, and because these results incorporate a natural weighting of the importance of different industries. The city-level regression specification is,

$$\Delta \ln(L_{ct}) = a_0 + a_1 \Delta \ln(PrWorkpop_{ct}) + a_2 \Delta \ln(PrCoal_{ct}) + \xi_t + e_{ct} \,, \qquad (10)$$

where $\Delta \ln(L_{ct})$ is the change in actual city population (either the working or the total population), $\Delta \ln(PrWorkpop_{ct})$ is the predicted change in the working population of city c, $\Delta \ln(PrCoal_{ct})$ is the predicted change in log coal use in the city, and ξ_t is a full set of year effects. As before, predicted variables are generated using lagged city-industry employment patterns and industry growth rates in all other cities, with differences taken over two-decade periods.[44] Note that there is an important difference between the specification in Eq. 10 and the regressions based on Eq. 9: in Eq. 10, the $\Delta \ln(PrWorkpop_{ct})$ term will reflect both the positive direct impact of industry growth on overall city employment as well as any negative congestion effects generated by increasing population.

City-level results are presented in Table 3. Columns 1-2 present results obtained by aggregating the private-sector industries used in the main analysis to the city level. Columns 3-4 present results for the entire working population of the city.[45] Columns 5-6 present results for the total city population, including children, students, the retired, and other non-workers. We can see that rising city coal use is negatively related to growth in any of these populations, though the impact on private sector workers is stronger than on all workers, which in turn is stronger than the impact on overall population. This may reflect that government workers and others in similar occupations, as well as non-workers such as retirees or family members, may be less flexible in adjusting to changing levels of local pollution.

[44]For specifics on the construction of these explanatory variables, see Appendix A.4.8.

[45]This includes government workers, agricultural workers, casual laborers, etc.

When analyzing city-level outcomes, it is also possible to look at how the impact of coal use differs between men and women and across different age groups of the population. This analysis, available upon request, shows that the impact of coal use is similar for both genders. Similar coal use effects are also observed across age groups, though there is some evidence of slightly larger negative effects for the local population of children under five, a pattern that is consistent with the mortality results discussed next.

Table 3: City-level regression results

DV: Δ Log of city employment in analysis industries (two decade differences)						
	City employment in analysis industries		Total city working population		Total city population	
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \ln(PrWorkpop_{ct})$	0.955	0.433	0.756	0.0795	0.385	-0.229
	(0.666)	(0.726)	(0.664)	(0.724)	(0.624)	(0.730)
$\Delta \ln(PrCoal_{ct})$	-1.457**	-1.655**	-1.352**	-1.400**	-0.986	-1.055
	(0.657)	(0.670)	(0.650)	(0.665)	(0.633)	(0.686)
Time effects	Yes	Yes	Yes	Yes	Yes	Yes
Other controls		Yes		Yes		Yes
Observations	155	155	155	155	155	155
R-squared	0.067	0.202	0.084	0.208	0.099	0.213

*** p<0.01, ** p<0.05, * p<0.1. Standard errors allow serial correlation across two decades. The data cover 31 cities over each decade from 1851-1911, with differences taken over twenty-year periods. The additional controls included are the number of air frost days in each city, rainfall in each city, patents in the city from 1852-1858, log city population at the beginning of the period, and log city coal use at the beginning of the period. The full results show that rainfall and initial city size are negatively related to city growth, while patenting and the initial level of coal use are positively associated with city growth.

While investigating the health effects of pollution is not the main focus of this paper, it is natural at this point to wonder whether the mortality effects of coal-based pollution can potentially explain the impact of coal use on city size that I have estimated. To address this issue, I draw on the analysis of the impact of local industrial coal use on mortality in 19th century Britain provided by Beach & Hanlon (2016). That study uses detailed district-level data in order to show that local industrial coal use had a strong impact on mortality in Britain in 1851-1860. Moreover, because the coal use variable used in that study is generated using the same approach as the one in this study, it is easy to apply the mortality estimates from that paper here. When I do so, in Appendix A.4.10, the results suggest that the increase in mortality associated with increased coal use can explain only a small fraction, about 2%, of the overall impact of coal use on city size. Thus, most of the population effects of coal use must be due to differences in the location choices of workers and firms.

5.2 Implications for urbanization levels

Was there scope for environmental regulations to reduce the negative externalities of coal use documented above? If so, what impact might these improvements have had on the British urban system? In an attempt to answer these questions, this section provides a counterfactual analysis of the impact of improved coal use efficiency.

The counterfactual that I consider is motivated by rich historical source, the 1871 Coal Commission report.[46] This extremely detailed report, over 1300 pages long, aimed to understand all aspects of coal use in Britain. As part of this study, one committee was specifically assigned to, "inquire whether there is reason to believe that coal is wasted by carelessness or neglect of proper appliances for its economical consumption." This group, Committee B, interviewed some of the leading luminaries of the time, including Henry Bessemer, the inventor of (among many other things) the Bessemer process, and Charles William Siemens, the inventor of the regenerative furnace.

The main finding of Committee B was that there was evidence of wide-spread waste and inefficiency in the use of coal that could have been remedied at relatively small cost.[47] The committee highlighted two major areas where relatively low-cost

[46]The full title of this report is, *Report of the Commissioners Appointed to Inquire into the Several Matters Relating to Coal in the United Kingdom*.

[47]Perhaps we should not be surprised that 19th century producers failed to achieve efficiency in their coal use given that, even in the modern U.S., these is some evidence suggesting a widespread failure to adopt energy efficiency technologies with positive net present values. See, e.g., McKinsey

improvements could lead to substantial reductions in industrial coal use. The first was the procedures used for adding coal to boilers.[48] On this, the Committee writes,

> *The careless and wasteful manner of stoking in most of the coal-producing districts is not only a source of vast waste, but of extreme annoyance to all the surrounding neighborhood. Coal is piled upon the fire without any discretion, producing dense volumes of the blackest smoke, which is so much fuel actually thrown away; nor is the waste the worst part of it; vegetation is destroyed, or seriously injured, for miles, and that which acts so seriously on the plant cannot fail to be injurious to man.* (p.103)

Second, the committee argues that efficiency gains could have been achieved cheaply through insulating boilers and steam engines to limit heat loss, with savings estimated at 30 percent. They write,

> *...we feel called upon to notice the enormous waste of heat, and consequently wasteful consumption of fuel, in a very large majority of the steam boilers used in this country...through their being left to the influence of every change in the atmospheric conditions, quite exposed to winds, rains, and snows, when a slight covering of a non-conducting substance would, by protecting them, improve their steam producing power, and save a considerable quantity of coal.* (p. 103)

Having found that such improvements were available, the committee then asked, why were these efficiency-improving technologies not implemented by manufacturers? Their findings suggest three main explanations. First, coal was abundant and relatively inexpensive, and the committee found that, "in places where coal is cheap

& Co. (2009), "Unlocking Energy Efficiency in the U.S. Economy," and Alcott & Greenstone (2012).

[48] On p. 104, the report states that, "Imperfect combustion must be regarded as the first essential loss. The air is supplied so unskillfully that much passes into the chimney as hot air, carrying with it the vast quantity of unconsumed carbonaceous matter which we see escaping in black clouds from the top of the chimney. This imperfect combustion may be traced to the bad construction of the fireplaces, and to the reckless way in which coal is thrown into, and over, the mass of ignited matter in the fireplace."

and abundant, it is used with but little regard to economy, and that indeed in some localities the men actually boast of the quantity of coal which they have contrived to burn" (p. 129).[49] Second, pollution regulations were generally weak and ineffective, providing producers with little additional incentive for efficiency improvements (Thorsheim (2006), Fouquet (2012)). Third, coal pollution imposed city-level externalities, so that producers had little incentive to unilaterally reduce their coal consumption.[50]

Overall, the findings of the Coal Commission report suggest that, near the middle of my study period, efficiency gains in the range of 10-30% could have been achieved using existing technology at relatively low cost. Motivated by these findings, I use the model in order to consider a counterfactual in which the *growth* of coal use across the study period was reduced by ten percent without imposing additional economic costs.

The counterfactual is implemented by starting with the 1851 population of cities and then working forward, adding in the additional population that we would expect the cities to attract given a ten percent reduction in the growth of local industrial coal use in each period based on the estimates obtained above.[51] Note that, because the counterfactual relies on the structure of the model, it incorporates the countervailing congestion effects associated with increased population growth.

The results of this exercise for overall city population, shown in Table 4, suggest that the population of the 31 analysis cities in 1911 would have been larger by about 1.5 million under the counterfactual.[52] As a result, these cities would have included

[49] With the exception of a few short spikes, coal prices were generally low and stable across the study period (Table 6 in Appendix A.1). Clark & Jacks (2007) suggests that this may have been due to a relatively flat supply curve for coal in Britain during the 19th century.

[50] The fact that manufacturers made unilateral investments in chimneys suggests that they internalized at least some of the costs that direct exposure of their workers to coal smoke would have imposed. However, these chimneys merely served to disperse the coal smoke more broadly and manufacturers in the large cities that I investigate had little incentive to internalize these broader effects.

[51] To be specific, when running the counterfactual for total population I use the coefficient estimates from Column 5 of Table 3.

[52] In Appendix A.4.9 I explore counterfactuals for the working population of these cities using estimates based on either the city-industry or city-level regression results. The results based on city-

38% of the English population in 1911, compared to the 34% actually achieved in that year. Today the 31 largest urban areas in England account for just over 40% of the population. Thus, a reduction in the growth of coal use could have led British cities to approach modern urbanization levels much earlier.

These results are particularly interesting because of the strong link between urbanization and income, a pattern that has been observed across many countries and time periods (see Acemoglu *et al.* (2002) for some evidence on this relationship). As a result, urbanization is often used as a proxy for income in historical settings in which direct income measures are unavailable. Thus, the effects of coal use on urbanization implies that rising local coal use was also likely to have important welfare consequences (though a full analysis of the impact of rising local coal use on welfare is beyond the scope of this study).[53]

Table 4: Actual and counterfactual total population of the 31 analysis cities

	Actual		Counterfactual	
Year	Population	Share of English population	Population	Share of English population
1851	5,147,432	0.30	5,147,432	0.30
1881	8,445,658	0.34	9,186,470	0.37
1911	11,626,649	0.34	13,080,666	0.38

This simple counterfactual includes an important assumption about the elasticity of labor supply faced by cities. Each city faces an upward-sloping city labor supply

level estimates are quite similar to those obtained using the theoretically-consistent city-industry level regressions allowing heterogeneous effects of coal use across industry. Thus, the city-level results are likely to provide a good approximation for the true effect of coal use.

[53]In the model, the free mobility of workers implies that the local welfare effect of coal use will be absorbed by the owners of land and local resources, but these may be offset by the gains of land and resource owners in other locations. Welfare will also be affected if there are agglomeration benefits generated by urbanization, a factor that is beyond the scope of this study. In addition, pollution almost certainly had health effects for which local residents were not fully compensated. A full welfare analysis would need to account for these.

curve, and these curves can shift over time as a result of global forces shaping labor supply. However, given global labor supply conditions, which determine the reservation utility in each period, the supply curve for workers to any particular city is not affected by the growth of the other analysis cities. While this is a strong assumption, it is not unreasonable in the setting I consider because English cities were part of a large international labor market where they competed with locations as distant as Australia, Argentina and the U.S. for workers, particularly workers from Ireland.

5.3 Consumer disamenities or productivity effects?

In the model, coal use can affect city growth through either consumer amenities or firm productivity. To separate these channels, we need location-specific wage, rent, and price data. While such data are generally unavailable, they are provided for a cross-section of 51 cities in 1905 from a report produced by the Board of Trade.[54] While these data are limited, and therefore the results of this section should be interpreted with caution, they can provide some suggestive evidence on the channels that may be generating the effects documented above.

To begin, I use the model to derive a standard expression relating the quality-of-life in cities to local amenities. Starting with the indirect utility function and substituting in Eq. 6, I obtain,

$$[\gamma \ln(P_t) + (1-\gamma)\ln(r_{ct})] - \ln(w_{ct}) = \ln(\delta_c) - \psi \ln(C_{ct}) - v_t^* \,. \tag{11}$$

The left-hand side of this equation is the difference between local costs, weighted by expenditure shares, and the local wage, a standard measure of local quality-of-life.[55]

[54] The Board of Trade data cover slightly more than 51 cities, but I am only able to use cities where city-industry data are also available, since those data are needed in order to calculate city coal use.

[55] Albouy (2012) suggests adjusting the standard approach to (1) include the local cost of goods other than housing, (2) include non-wage income, and (3) account for federal income taxes and deductions. Non-wage income and income taxes are not a major concern in my empirical setting. I

Estimating this equation allows me to obtain the parameter ψ, which determines how local coal use affects city employment growth through the amenity channel.[56]

These regressions are run using wage data for skilled builders and skilled engineers, occupations that are found in most or all of the cities.[57] The cost data include both rental rates and the local prices of goods, which the Board of Trade combined based on the expected share of expenditures going towards housing.

Table 5 presents the results. Columns 1-3 use the wages of skilled builders while Columns 4-6 are based on skilled engineer's wages, which are available for a smaller set of cities. Each column includes the log of city coal use as an explanatory variable, while additional control variables are added in Columns 2-3 and 5-6.[58] In all specifications, city coal use is negatively related to the amenity value of the city, and this relationship is statistically significant in most of the results.[59]

The results in Table 5 indicate that coal use had a negative impact on the quality-of-life in British cities in 1905. However, the magnitude of the estimates suggest that this effect was not large. In Appendix A.4.11 I describe how these estimates, together with the results from the main analysis, can be used to analyze the relative importance of the amenities and productivity channels. These calculations show that,

incorporate the first adjustment he recommends into my analysis by using Board of Trade cost of living estimates which include both housing and local goods prices.

[56]This is essentially the same data and estimating approach used in Williamson (1981b), though he uses different data to infer local pollution levels. This highlights the fact that his approach will identify only the amenity channel.

[57]Skilled occupations are used because skilled workers were likely to be more mobile across cities, so these wage data are more likely to reflect city amenities, and because the wives of skilled workers were less likely to work, so the wage of skilled male workers will better reflect household income than the wage of unskilled workers. This issue was raised by Pollard (1981) in his critique of Williamson (1981b), who focused instead on unskilled wages. Further details on the Board of Trade data are presented in Appendix A.2.2.

[58]Spatial correlation is potentially a concern in these regressions. To deal with this, I have explored allowing spatial correlation of standard errors for cities within 50km of each other, following Conley (1999). I find that this delivers smaller confidence intervals, and therefore more statistically significant results, than those obtained using robust standard errors. To be conservative, Table 5 reports the larger robust standard errors.

[59]Further analysis shows that these effects are driven by a combination of lower rents and goods prices in more polluted cities together with small and generally statistically insignificant increases in wages.

for plausible values of the production function parameters, the impact of coal use on city employment growth through the channel of consumer amenities is much smaller than the impact through productivity effects.[60]

Table 5: Comparing quality-of-life measures to city coal use

	DV: QOL_c **for Skilled Builder**			**DV:** QOL_c **for Skilled Engineer**		
	(1)	(2)	(3)	(4)	(5)	(6)
$Ln(COAL_c)$	-0.0172*	-0.0504**	-0.0454**	-0.0294***	-0.0452**	-0.0378*
	(0.00946)	(0.0203)	(0.0195)	(0.0108)	(0.0174)	(0.0194)
$Ln(POP_c)$		0.0421**	0.0329		0.0185	0.0129
		(0.0208)	(0.0208)		(0.0187)	(0.0208)
Controls			Yes			Yes
Observations	51	51	51	47	47	47
R-squared	0.053	0.133	0.204	0.139	0.153	0.183

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. The QOL measure is constructed using data for 1905 from the Board of Trade. $COAL_c$ is calculated using industry coal interacted with city's industrial composition in 1901. CityPop is the population of the city in 1901. Note that wage data for skilled engineers is available for fewer cities than wage data for skilled builders. Included controls: air frost days and rainfall.

6 Conclusion

There has long been debate over the magnitude of the external costs of the Industrial Revolution and the "dark satanic mills" that it brought to English cities. By bringing together new data and a novel estimation approach, this paper moves us closer to resolving this debate. My results show that local industrial pollution related to coal use came with large costs for local economic growth. My findings contrast sharply with previous work (e.g., Williamson (1981b)), which suggested that these costs were "trivial," but are in line with both the historical evidence on the severity of the pollution experienced by cities in the 19th century as well as our current understanding of the substantial impacts that such pollution can bring.

The problems of industrialization and pollution experienced by 19th century En-

[60]I consider plausible values of the production function parameters because, given the available data, it is not possible to directly estimate all of the necessary parameters.

glish cities are echoed today in the industrial cities in the developing world. Policymakers in places such as China and India face important questions about whether to encourage industrial growth or to protect the local environment. Often, the economic benefits of industrial growth are directly observable, while the costs imposed by pollution are less tangible. This study provides the first rigorous estimates of the long-run local economic impacts that can accompany industrial pollution. While the relationship between industrialization and pollution has surely changed over the past century, the magnitude of the effects I document provide a warning against ignoring the economic consequences of local pollution. At the same time, this study develops tools that can potentially be applied in order to measure the consequences of industrial pollution in other relatively data-sparse settings.

References

Acemoglu, Daron, Johnson, Simon, & Robinson, James A. 2002. Reversal of Fortune: Geography and Institutions in the Making of the Modern World Income Distribution. *Quarterly Journal of Economics*, **117**(4), 1231 – 1294.

Albouy, David. 2012 (May). *Are Big Cities Bad Places to Live? Estimating Quality of Life across Metropolitan Areas.* Working paper.

Alcott, H., & Greenstone, M. 2012. Is There an Energy Efficiency Gap? *Journal of Economic Perspectives*, **26**(1), 3–28.

Baines, Dudley. 1985. *Migration in a Mature Economy.* Cambridge, UK: Cambridge University Press.

Banzhaf, H. Spencer, & Walsh, Randall P. 2008. Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism. *The American Economic Review*, **98**(3), pp. 843–863.

Barreca, Alan, Clay, Karen, & Tarr, Joel. 2014 (February). *Coal, Smoke, and Death: Bituminous Coal and American Home Heating.* NBER Working Paper No. 19881.

Bartelme, Dominick. 2015 (October). *Trade Costs and Economic Geography: Evidence from the U.S.* Mimeo.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Bayer, Patrick, Keohane, Nathaniel, & Timmins, Christopher. 2009. Migration and Hedonic Valuation: The Case of Air Quality. *Journal of Environmental Economics and Management*, **58**(1), 1 – 14.

Beach, B., & Hanlon, W.W. 2016 (June). *Coal Smoke and Mortality in an Early industrial Economy.* Mimeo.

Bowley, A.L. 1937. *Wages and Income in the United Kingdom Since 1860.* London: Cambridge University Press.

Brimblecombe, Peter. 1987. *The Big Smoke: A History of Air Pollution in London Since Medieval Times.* Methuen.

Chang, T., Graff Ziven, J., Gross, T., & Neidell, M. 2014 (February). *Particulate Pollution and the Productivity of Pear Packers.* NBER Working Paper No. 19944.

Chang, T., Graff Ziven, J., Gross, T., & Neidell, M. 2016 (June). *The Effect of Pollution on Worker Productivity: Evidence from Call-Center Workers in China.* NBER Working Paper No. 22328.

Chay, Kenneth Y., & Greenstone, Michael. 2005. Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy*, **113**(2), pp. 376–424.

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

Clay, K, Lewis, J., & Severnini, E. 2015 (October). *Pollution, Infectious Disease, and Mortality: Evidence from the 1918 Spanish Influenza Pandemic*. NBER Working Paper No. 21635.

Clay, Karen, Lewis, Joshua, & Severnini, Edson. 2016 (March). *Canary in a Coal Mine: Impact of Mid-20th Century Air Pollution Induced by Coal-Fired Power Generation on Infant Mortality and Property Values*. Working paper.

Conley, Timothy G. 1999. GMM Estimation with Cross Sectional Dependence. *Journal of Econometrics*, **92**(1), 1 – 45.

Crafts, Nicholas, & Mulatu, Abay. 2006. How Did the Location of Industry Respond to Falling Transport Costs in Britain before World War I? *The Journal of Economic History*, **66**(3), pp. 575–607.

Crafts, Nicholas, & Wolf, Nikolaus. 2014. The Location of the UK Cotton Textiles Industry in 1838: a Quantitative Analysis. *Journal of Economic History*, **74**(4), 1103–1139.

Diamond, Rebecca. 2016. The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000. *American Economic Review*, **106**(3).

Ellison, G, & Glaeser, EL. 1999. The geographic concentration of industry: Does natural advantage explain agglomeration? *American Economic Review*, **89**(2), 311–316.

Ellison, G., Glaeser, E., & Kerr, W. 2010. What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*, **100**(3), pp. 1195–1213.

Engels, Friedrich. 1845. *The Condition of the Working Class in England*. Penguin Books.

Fernihough, Alan, & O'Rourke, Kevin. 2014 (January). *Coal and the European Industrial Revolution*. NBER Working Paper No. 19802.

Fouquet, R. 2008. *Heat, Power and Light: Revolitions in Energy Services*. Cheltenham, UK: Edward Elgar Publications.

Fouquet, R. 2012. The Demand for Environmental Quality in Driving Transitions to Low-Polluting Energy Sources. *Energy Policy*, **50**, 138–149.

Glaeser, Edward L., Kolko, Jed, & Saiz, Albert. 2001. Consumer city. *Journal of Economic Geography*, **1**(1), 27–50.

Graff Zivin, Joshua, & Neidell, Matthew. 2012. The Impact of Pollution on Worker Productivity. *American Economic Review*, **102**(7), 3652–73.

Hanlon, W.W., & Miscio, A. 2014 (December). *Agglomeration: A Dynamic Approach*. NBER Working Paper No. 20728.

Hanna, Rema, & Oliva, Paulina. 2015. The Effect of Pollution on Labor Supply: Evidence from a

Natural Experiment in Mexico City. *Journal of Public Economics*, **122**, 68–79.

Heblich, Stephan, Trew, Andrew, & Zylberberg, Yanos. 2015 (June). *East Side Story: Historical Pollution and Neighborhood Segregation*. Working paper.

Hsiang, Solomon M. 2010. Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *Proceedings of the National Academy of Sciences*, **107**(35), 15367–15372.

Isen, Adam, Rossin-Slater, Maya, & Walker, W Reed. Forthcoming. Every Breath You Take – Every Dollar You Make: The Long-Term Consequences of the Clean Air Act of 1970. *Journal of Political Economy*.

Kahn, Matthew E. 1999. The Silver Lining of Rust Belt Manufacturing Decline. *Journal of Urban Economics*, **46**(3), 360 – 376.

Kettlewell, H Bernard D. 1955. Selection Experiments on Industrial Melanism in the Lepidoptera. *Heredity*, **9**, 323–342.

Kline, Patrick, & Moretti, Enrico. 2013. Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority. *Quarterly Journal of Economics*.

Kovak, Brian K. 2013. Regional Effects of Trade Reform: What is the Correct Measure of Liberalization? *American Economic Review*, **103**(5), 1960–1976.

Landes, David S. 1998. *The Wealth and Poverty of Nations*. New York, NY: W.W. Norton & Company.

Lavy, Victor, Ebenstein, Avraham, & Roth, Sefi. 2014 (October). *The Impact of Short Term Exposure to Ambient Air Pollution on Cognitive Performance and Human Capital Formation*. NBER Working Paper No. 20648.

Lee, Jamie. 2015 (January). *Measuring Agglomeration: Products, People and Ideas in U.S. Manufacturing, 1880-1990*. Mimeo.

Long, Jason, & Ferrie, Joseph. 2003. Labour Mobility. *In:* Mokyr, Joel (ed), *Oxford Encyclopedia of Economic History*. New York: Oxford University Press.

Mitchell, B.R. 1984. *Economic Development of the British Coal Industry 1800-1914*. Cambridge, U.K.: Cambridge University Press.

Mitchell, Brian R. 1988. *British Historical Statistics*. Cambridge, UK: Cambridge University Press.

Moretti, Enrico. 2004. Workers' Education, Spillovers, and Productivity: Evidence from Plant-level Production Functions. *American Economic Review*, **94**(3), 656–690.

Moretti, Enrico. 2011. Local Labor Markets. *Pages 1237–1313 of: Handbook of Labor Economics*, vol. 4. Elsevier.

Mosley, Stephen. 2001. *The Chimney of the World*. Cambridge, UK: The White Horse Press.

Newey, Whitney K., & West, Kenneth D. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, **55**(3), pp. 703–708.

Pollard, Sidney. 1981. Sheffield and Sweet Auburn–Amenities and Living Standards in the British Industrial Revolution: A Comment. *The Journal of Economic History*, **41**(4), pp. 902–904.

Rauch, JE. 1993. Productivity Gains from Geographic Concentration of Human Capital - Evidence from the Cities. *Journal of Urban Economics*, **34**(3), 380–400.

Rückerl, R, Schneider, A, Breitner, S, Cyrys, J, & Peters, A. 2011. Health Effects of Particulate Air Pollution: A Review of Epidemiological Evidence. *Inhalation Toxicology*, **23**(10), 555–592.

Thomas, Mark. 1987. *An Input-Output Approach to the British Economy, 1890-1914*. Ph.D. thesis, Oxford University.

Thorsheim, Peter. 2006. *Inventing Pollution*. Athens, Ohio: Ohio University Press.

Troesken, W, & Clay, K. 2011. Did Frederick Brodie Discover the World's First Environmental Kuznets Curve? Coal Smoke and the Rise and Fall of the London Fog. *In:* Libecap, G., & Steckel, R. H. (eds), *The Economics of Climate Change: Adaptations Past and Present*. University of Chicago Press.

Williamson, Jeffrey G. 1981a. Some Myths Die Hard–Urban Disamenities One More Time: A Reply. *The Journal of Economic History*, **41**(4), pp. 905–907.

Williamson, Jeffrey G. 1981b. Urban Disamenities, Dark Satanic Mills, and the British Standard of Living Debate. *The Journal of Economic History*, **41**(1), pp. 75–83.

Williamson, Jeffrey G. 1982. Was the Industrial Revolution Worth It? Disamenities and Death in 19th Century British Towns. *Explorations in Economic History*, **19**(3), 221 – 245.

A Appendix

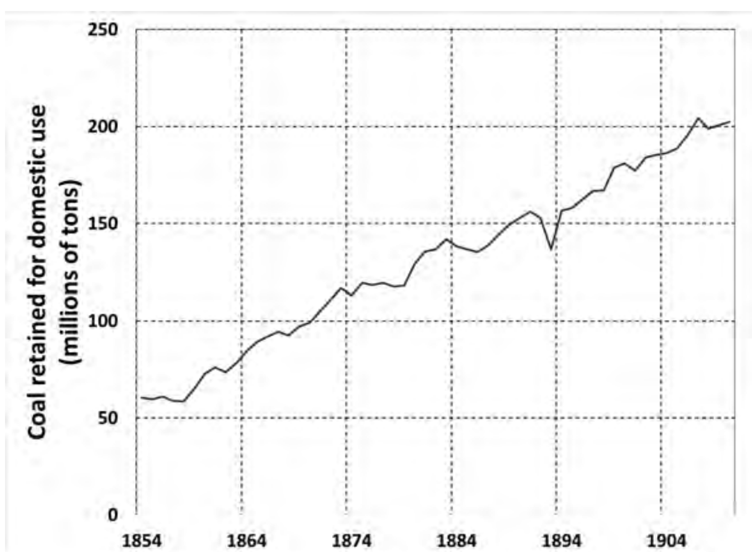
A.1 Empirical setting appendix

Figure 3 presents a map of the cities included in the analysis. We can see that the cities are drawn from across the country, though there is a concentration of cities in the Northwest region, the industrial heartland of England.

Figure 3: English cities included in the study

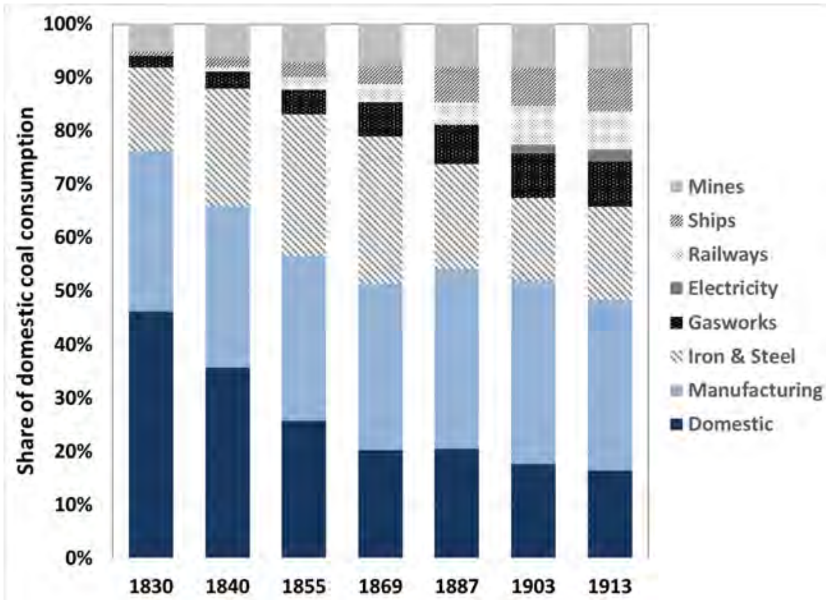
The next set of figures illustrate the increase in coal use across the study period. Figure 4 shows the steady rise in British coal consumption across the study period using data from Mitchell (1988). Figure 5 breaks this down by the different categories of users. This figure shows that the uses captured in my industrial coal use measure, which includes manufacturing, iron & steel, and mining, cover the majority of total coal consumption. In contrast, residential use accounts for just 17-25% of coal consumption during the period I study, a fraction that was declining over time. Figure 6 describes the price of coal at the major exporting ports. There are a couple of important points to take away from this figure. First, except for a few short spikes, the price of coal was largely stable across the study period. Second, prices were quite similar in different parts of the country, reflecting the low cost of transportation in England during this period.

Figure 4: British coal consumption, 1854-1910



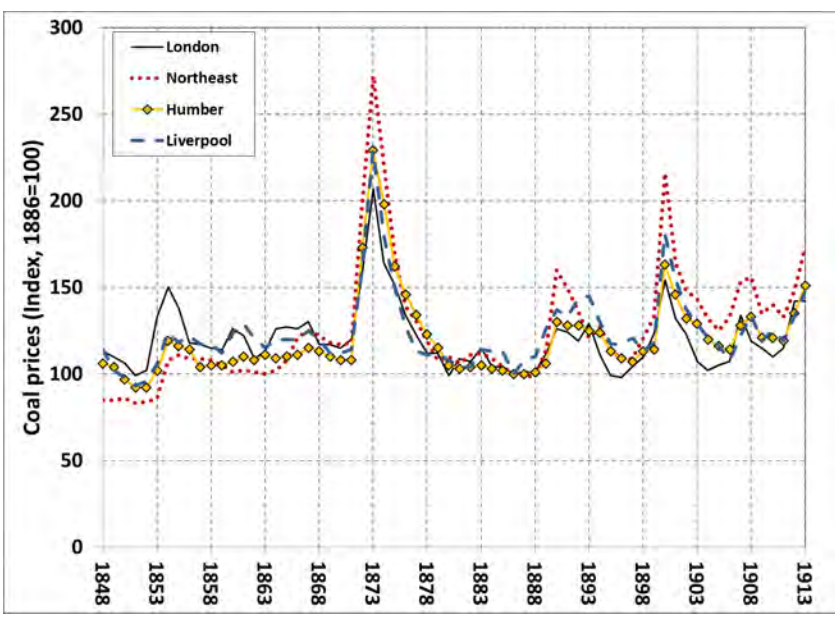
Data from Mitchell (1988).

Figure 5: British coal consumption by use, 1830-1913



Data from Mitchell (1988).

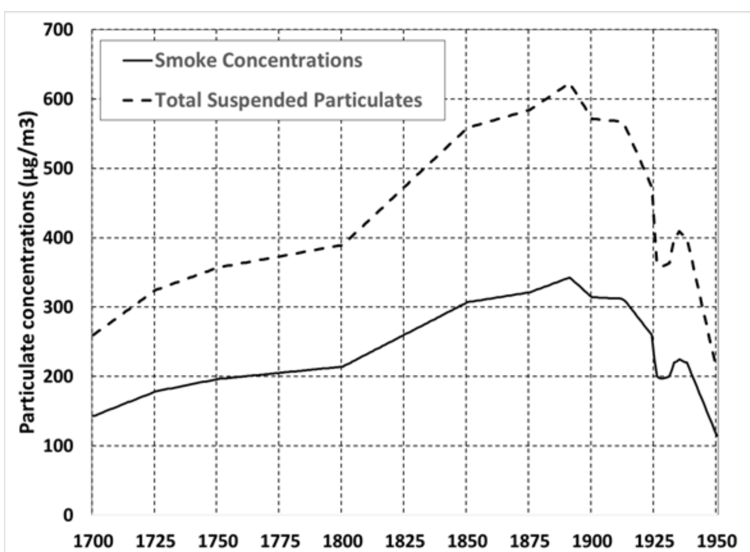
Figure 6: Coal prices at the major exporting ports



Data from Mitchell (1984).

Figure 7 presents estimates of smoke levels in London from 1700-1850 from Brimblecombe (1987) which have been translated into TSP levels by Fouquet (2008). These estimated smoke levels are based on data on the inflows of coal into London for consumption. While these estimates are rough, they suggest that British cities experienced very high levels of pollution. As a point of comparison, the WHO guidelines for 24 hour mean PM10 pollution exposure are 50 $\mu g/m^3$. Heavily polluted modern industrial , such as Karachi or Delhi, regularly face levels from 250-300 $\mu g/m^3$ according to the WHO.

Figure 7: Smoke and particulate estimates for London, 1700-1950



Smoke estimates from Brimblecombe (1987). These are converted into TSP levels by Fouquet (2008).

Table 6 presents a summary of major air pollution regulations and other related events in Britain during the study period.

Table 6: History of British air pollution regulation, 1851-1911

1853-6	Smoke abatement acts relating to the Metropolitan area
1866	The Sanitary Act empowered sanitary authorities to take action in cases of smoke nuisances
1875	The Public Health Act containing a smoke abatement section on which legislation to the present day has been based
1881	Smoke abatement exhibition at South Kensington organized by the Public Health and Kyrle Societies
1882	Formation of the National smoke abatement institution
1891	The Public Health (London) Act
1899	Formation of the Coal Smoke Abatement Society
1909	Sheffield smoke abatement exhibition, at which was set up the Smoke Abatement League of Great Britain (mainly for the provinces and centered later in Manchester and Glasgow).

Source: The Glasgow Herald (Sept. 24, 1958)

A.2 Data appendix

This appendix provides additional details on the new data sets used in this study, beginning with the data gathered from the 1907 Census of Production. I do not review the construction of the Census of Population data, which is described in detail at `http://www.econ.ucla.edu/whanlon` under Data Resources.

A.2.1 Census of Production data

The 1907 Census of Production, Britain's first industrial census, provides the earliest comprehensive look at the characteristics of British industries. For the purposes of this paper, the most important piece of information provided by the Census of Production is the amount of coal and coke burned in each industry. Figure 8 shows an example of what these data look like for the iron and steel industries.

Figure 8: An example of the Census of Production fuel use data

Trade.	Net Output of Firms Furnishing Particulars.		Fuel consumed by Firms Furnishing Particulars.	
	Amount.	Percentage of Total Net Output of the Trade.	Coal.	Coke.
	£		Tons.	Tons.
Iron and Steel Trades (Smelting, Rolling and Founding).	12,539,000	41·7	3,728,524	162,006
Tinplate Trade	1,681,000	83·7	708,896	52
Wrought Iron and Steel Tube Trade	985,000	45·0	243,062	13,519
Wire Trades	1,637,000	77·2	187,956	15,223
Anchor, Chain, Nail, Bolt, Screw and Rivet Trades.	1,258,000	54·4	110,147	28,655
Galvanized Sheet, Hardware, Hollow-ware, Tinned and Japanned Goods and Bedstead Trades.	4,347,000	66·5	226,668	70,520
Engineering Trades (including Electrical Engineering).	32,632,000	64·6	1,400,171	468,503
Royal Ordnance Factories	1,452,000	100·0	95,991	10,156
Naval Ordnance Factories	77,000	100·0	1,874	200
Shipbuilding Yards and Marine Engineering Trades :—				
Private Firms	14,142,000	76·3	606,317	90,099
Government Yards and Lighthouse Authorities.	2,470,000	99·2	113,075	10,741
Cycle and Motor Trades...	3,904,000	66·2	36,982	8,967
Cutlery Trade	491,000	45·4	15,603	3,318
Tool and Implement Trades	1,278,000	61·1	109,815	35,259
Blacksmithing Trade	1,169,000	79·1	52,655	16,251
Needle, Pin, Fish-hook, and Button Trades ...	418,000	49·4	14,679	915
Lock and Safe Trades	467,000	72·3	8,328	2,457
Small Arms Trades	162,000	30·1	3,801	588
Heating, Lighting, Ventilating, and Sanitary Engineering Trades.	903,000	57·6	8,801	11,335
Railway Carriage and Wagon Trades	3,189,000	89·5	300,144	80,888
Railways (Construction, Repair, and Maintenance of Permanent Way, Rolling Stock, Plant, &c.).	17,082,000	99·9	1,013,708	161,867
Total	102,283,000	66·8	8,987,197	1,191,519

To construct coal use per worker in each industry, I begin by adding together coal and coke used in each industry. Next, I inflate that value to reflect the fact that only a fraction of firms in the industry furnished particulars to the census office. I then match the industries listed in the Census of Production to the broader industry categories available in the Census of Population data and sum across each of the Census of Population categories. Finally, I divide by the number of workers in the industry, which is also reported in the Census of Production.

It is necessary to make an additional modification for one industry, "Chemicals, coal tar products, drugs and perfumery", which was one component of the broader "Chemical and allied trades" category. The adjustment is necessary due to the fact that a large amount of coal was used by that industry to produce coal-based products such as coal tar. Since this coal wasn't burned, I don't want to count it toward industry coal use. Unfortunately, the Census does not separately report the amount of coal used for products such as coal tar and the amount burned for energy. To separate these amounts, I use the horsepower of engines in the industry, which is reported in the Census. I then calculate the amount of coal used per horsepower in all of the other branches of the "Chemicals and allied products" sector and then multiply the number of horsepower used in the "Chemicals, coal tar products, drugs and perfumery" by this value to obtain an estimate of the amount of coal burned in that sub-sector. The result of this adjustment is a reduction of about one-third in the amount of coal use per worker in the Chemical & Drug sector.

A.2.2 Board of Trade data

This study also takes advantage of data from a 1908 report from the Labour Department of the British Board of Trade, which reports data primarily gathered in 1905. The goal of this report was to document the conditions of the working class in the various major towns of Britain, including the rents and prices they faced for common goods such as bread, meat and butter, and the wages they earned.

The first piece of data provided by these reports are rental rates. The rental data were "obtained from officials of the local authorities, from the surveyors of taxes, or from the house owners and agents in the various towns...A considerable number of houses in each town were visited, partly for purposes of verification and supplementary inquiry, and partly that some account might be given of the character of the houses and accommodation afforded." All rents were then converted to an index, with London as the base, by comparing the rent of the most predominant dwelling type in a town to the rental rate for that dwelling type in London. It is worth noting that these index numbers reflect the cost of housing relative to a similar accommodation in London, not the amount spent by a worker on housing relative to a similar worker in London.

Price data for the towns were obtained by surveying "representative tradesmen in possession of a working-class custom," as well as co-operative societies and larger multi-branch retail firms. The prices were quoted for October 1905. The center of the price ranges for each item in a town is then used. To weight the items, the Board of Trade used information from an inquiry into the expenditures of working-class families in 1904. These data were obtained from 1,944 surveys filled out by workmen throughout the country. Together, these data allow the construction of index numbers describing the price level of goods commonly purchased by workers in each city. The Board of Trade also constructed a combined index of prices and rents in which prices were given a weight of 4 and rents a weight of 1.

Wage data are also available from these reports. These data come from four trades which were present in many towns: construction, engineering, printing and furnishing. Of these, I focus on the construction and engineering trades, where data are available for more towns than the printing and furnishing trades. For the construction and engineering trades, separate wage data were collected for skilled workers and unskilled laborers. The wage data are weekly wage rates and may be affected by variation in the standard number of hours worked across locations.

A.2.3 Constructing additional control variables

One threat to identification in this study is the possibility that there may be other industry features that vary across industries in a way that is correlated with industry coal use and affects overall city size. One way to help guard against this concern is to construct additional control variables based on other potentially important characteristics that vary across industries. For the purposes of this study it is possible to construct additional controls for several potentially important factors:

Salaried workers: Work by Rauch (1993), Moretti (2004), and more recently Diamond (2016) suggests that the presence of high-skilled workers may impact overall city growth. To control for this potential effect, I use data from the 1907 Census of Production which divides workers into wage earners and salaried workers. This gives me the share of salaried employment by industry, which I interact with overall industry employment information in order to obtain estimates of the share of salaried workers in the city.

Firm size: The 1851 Census includes information gathered from business owners on the number of workers that they employ. This information is available by industry. Using this, I construct a variable reflecting the firm size experienced by the average worker in each industry in that year. I can then interact this with city-industry employment in order to get a population-weighted average firm size in each city.

Labor cost share: Labor cost shares were constructed using information from the 1907 Census of Production and from Bowley (1937). For each industry, the Census of Production provides the gross and net output value as well as employment by gender. To calculate total labor cost share in each industry I use wage data from Bowley (1937), which reports the average wage for different industry groups in 1906, separated into male and female wages. Multiplying these by the number of male and female workers in each industry from the 1907 Census of Production gives total labor cost in each industry.

Export shares: The share of industry output sold to export is estimated using information from the 1907 Input-Output table constructed by Thomas (1987). This table includes both total industry sales as well as industry export sales, which together give me the share of industry sales that are exported.

Industry female and youth labor shares: The share of female workers in each industry and workers under 20 in each industry are based on Census of Population data for 1851, which reports industry occupation by gender and divided into those over and under 20.

Rainfall and Air-frost data: The data on rainfall and air frost days comes from modern data collected by the Met weather service for a thirty-year period. An air frost day is defined as a day in which the air temperature drops below the freezing point of water at a height of one meter above ground.

City patenting data: The data on patenting at the city level are from 1852-1858. These data come from a compilation done by the Patent Office and included among the patent abstract records at the British Library's Business and Intellectual Property Section. I am not aware of a source that lists patent counts by location after 1858.

Proximity to Carboniferous geological strata: The data on the location of the carboniferous geological strata comes from the British Geological Survey. The proximity of each city to the carboniferous strata was constructed using GIS. In the analysis presented in this appendix, I use the share of bedrock within 50km of the city that is made up of carboniferous strata. I have also explored alternative windows, such as 10km and 100km.

Input-output connections: The input-output data used in this study were constructed by Thomas (1987) using data from the 1907 Census of Manufactures.

Industry demographic similarity: The demographic similarity of the workforces of any pair of industries is based on data from the Census of Population from 1851. These data divide industry employment into male and female workers and those over or under 20. The demographic similarity measure for a pair of industries is simply the correlation between the two industries in the share of the workforce that is in each of these four bins.

Industry occupational similarity: The occupational similarity of any pair of industries is based on the correlation in the vector of employment shares for each occupation. Industry occupation data is built on U.S. Census data for 1880 (the British census does not simultaneously measure occupation and industry until later).

A.2.4 City coal use intensity data

Table 7: Industrial coal use per private-sector worker for analysis cities (tons/year)

City	1851	1861	1871	1881	1891	1901	1911	Avg.	Growth
BATH	1.7	2.2	2.5	2.7	2.7	2.5	2.8	2.4	0.40
BRIGHTON	2.1	2.4	3.0	3.2	3.1	2.9	2.8	2.8	0.35
NORTHAMPTON	2.4	2.9	3.5	3.0	2.8	2.8	2.9	2.9	0.22
PORTSMOUTH	2.7	3.6	4.0	5.0	4.7	4.8	4.6	4.2	0.56
LIVERPOOL	2.7	3.4	4.1	4.4	4.1	4.1	4.1	3.8	0.42
LONDON	2.7	3.3	3.8	3.8	3.9	3.7	3.4	3.5	0.30
LEICESTER	2.8	3.8	4.6	4.0	3.5	4.2	4.9	4.0	0.43
SOUTHAMPTON	2.9	3.2	4.2	4.7	3.8	3.1	3.1	3.6	0.22
HULL	3.2	4.4	6.1	5.7	5.7	5.6	6.1	5.3	0.67
BRISTOL	3.2	4.1	4.7	4.7	4.5	4.6	4.8	4.4	0.36
NORWICH	3.3	4.0	4.8	5.1	4.6	4.2	4.1	4.3	0.29
NOTTINGHAM	3.8	4.9	6.0	7.4	7.2	6.9	6.9	6.1	0.63
IPSWICH	3.9	4.8	5.9	6.1	5.8	5.9	6.8	5.6	0.43
HUDDERSFIELD	4.6	5.6	7.0	7.5	7.7	7.4	7.3	6.7	0.47
BLACKBURN	5.1	6.8	7.7	8.4	8.2	7.8	7.9	7.4	0.46
MANCHESTER	5.1	6.3	7.0	7.3	7.4	7.1	6.9	6.7	0.32
SUNDERLAND	5.1	6.4	9.0	8.9	8.1	8.1	8.1	7.7	0.51
PRESTON	5.2	6.7	7.7	8.2	7.7	7.2	7.1	7.1	0.38
HALIFAX	5.4	6.6	8.4	9.4	9.6	9.3	9.6	8.3	0.55
STOCKPORT	5.4	6.5	7.1	7.1	6.3	5.8	6.6	6.4	0.19
SOUTH_SHIELDS	5.6	5.6	7.9	7.7	7.5	7.8	8.4	7.2	0.28
DERBY	5.7	7.0	9.0	10.2	9.4	8.6	8.5	8.3	0.47
NEWCASTLE	5.7	7.7	9.1	8.2	9.0	8.7	8.5	8.1	0.42
BRADFORD	6.0	7.2	7.9	8.2	7.7	7.3	7.1	7.4	0.23
LEEDS	6.3	8.6	10.1	10.0	9.1	8.6	8.7	8.8	0.39
BOLTON	6.5	8.9	9.7	10.4	10.3	9.7	9.6	9.3	0.44
OLDHAM	6.7	9.4	9.9	11.2	11.8	11.2	10.8	10.1	0.51
BIRMINGHAM	9.6	11.6	12.8	14.0	13.6	12.3	11.5	12.2	0.28
GATESHEAD	10.3	13.0	15.2	14.3	13.3	11.6	11.1	12.7	0.23
WOLVERHAMPTON	11.4	13.2	16.1	15.8	14.2	12.0	11.6	13.5	0.18
SHEFFIELD	12.4	14.8	17.6	17.2	16.5	15.3	15.5	15.6	0.26
Average	5.1	6.4	7.6	7.9	7.5	7.1	7.2		
Std. Dev.	2.6	3.2	3.7	3.7	3.5	3.1	3.0		

Author's calculations based on city-industry employment data from the Census of Population and industry coal use per worker data from the Census of Manufactures, as described in Section 3.

A.2.5 Industry coal use intensity data

Table 8: Industry coal use per worker and industry employment in 1851

Industry	**Coal/ worker**	**Workers in 1851** National	Analysis cities
Earthenware, bricks, etc.	48.9	83,353	19,580
Metal and engine manufacturing*	43.7	431,411	167,052
Chemical and drug manufacturing	40.1	35,655	11,501
Mining	28.9	328,062	18,413
Oil, soap, etc. production	20.7	17,063	12,188
Brewing and beverage production	19.4	27,527	8,179
Leather, hair goods production	12.1	57,097	26,737
Food processing	12.0	302,259	113,610
Textile production	10.1	968,412	315,646
Paper and publishing	9.7	66,622	42,578
Shipbuilding	6.1	26,840	14,498
Wood furniture, etc., production	5.4	136,794	69,648
Vehicle production	2.6	15,574	9,021
Instruments, jewelry, etc.	2.0	43,818	31,048
Apparel	1.6	873,835	328,669
Tobacco products	1.1	3,915	3,298

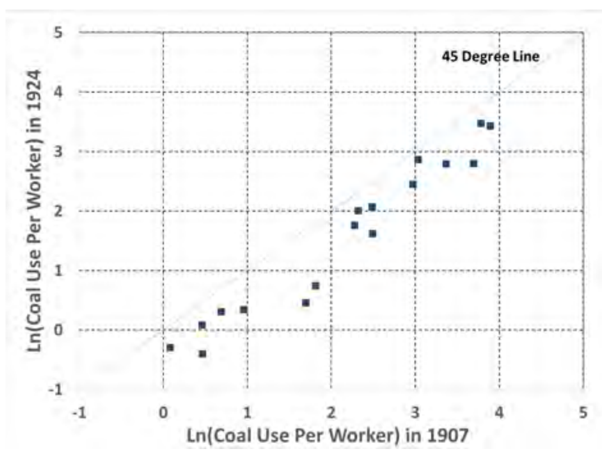
*Metal and engine manufacturing includes iron and steel smelting. Coal per worker is in tons per year. These values come from the 1907 Census of Production. The number of workers in each industry in 1851 come from the Census of Population Occupation reports.

A.2.6 Analyzing the change in relative industry coal intensity over time

The assess the stability of relative coal use intensity across industries, I can compared data from the 1907 Census of Production to the next Census, in 1924. Figure 9 provides a scatterplot of industry coal use per worker for each industry in 1907 and 1924 as well as corresponding regression results. Each point in the figure corresponds to one industry. This figure shows that there was very little change in the *relative* coal intensity of industries from 1907 to 1924. This is reflected in the coefficient on coal use per worker in 1907, which is very close to, and statistically indistinguishable from, one. We can also see that all of the points are below the 45 degree line, which appears in the regression results as a negative constant term. This suggests that there were changes in coal use per worker that were similar, in percentage terms, across all industries during this period. This type of change is reflected in the ρ_t term in Eq. 1.

The stability in relative industry coal use intensity described by Figure 9 is a particularly strong result because we would expect industry coal use to change more slowly in the 1851-1907 period than in the 1907-1924 period due to the adoption of electrical power by some manufacturing industries during the latter period. The shift to electricity had the potential to substantially affect industry coal use, whereas in the 1851-1911 period coal was the dominant energy source for industries and there were few alternatives.

Figure 9: Comparing industry coal use in 1907 and 1924



	DV: Coal per worker in 1924
Coal per worker in 1907	1.021***
	(0.0612)
Constant	-0.623***
	(0.151)
Observations	17
R-squared	0.949

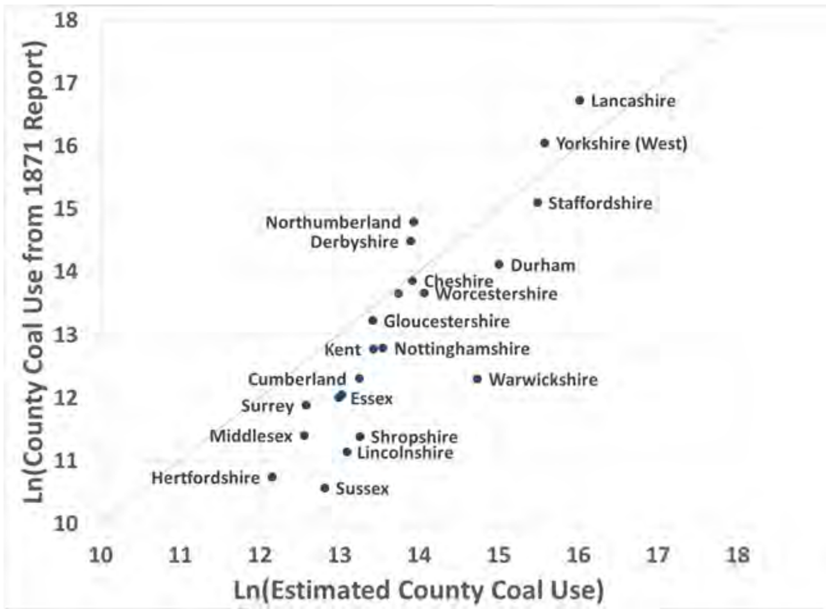
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

A.2.7 Comparing to 1871 county-level coal use

As an additional check of the coal use measure I have constructed, I compare county-level industrial coal use calculated using my methodology to estimates for 1871 based on data from the 1871 Coal Commission Report. That report, which was prompted by fears of a coal shortage in the early 1870s, included a survey of industrial coal use in a selection of English counties. Within each county, circulars were sent to firms asking them about their coal use. Using the resulting reports, and adjusting for the number of circulars returned in each county, I am able to calculated industrial coal use levels in the counties surveyed, though these figures will be imperfect because only major industrial establishments were surveyed. I then compare these estimates to results obtained by applying my methodology to county-level industrial employment data from the 1871 Census of Population combined with industry coal use intensity measures from the 1907 Census of Production.

Figure 10 describes the results for the set of available counties. In this graph, the y-axis describes county-level coal use constructed from the 1871 Coal Commission report while the x-axis gives the county coal use estimated using the methodology introduced in this paper. In general, the points lie close to the 45 degree line, suggesting that my methodology does a reasonable job of matching the estimates obtained using the data from the Coal Commission report. The methodology used in this paper does particularly well for the larger and more industrial counties. The greatest differences occur in the more rural counties with low levels of coal use, where my methodology overestimates industrial coal use relative the the figures from the 1871 Coal Commission report. However, these are also the counties where the figures from the Coal Commission report are most likely to understate county coal use because smaller industrial establishments, which were omitted from the Coal Commission report, are likely to form a more important coal user in less industrialized counties. Overall, these results provide additional evidence that the methodology used to calculate industrial coal use in this paper delivers reasonable results.

Figure 10: Comparing county industrial coal use in 1871



A.3 Theory appendix

A.3.1 Further theory results and aggregating to the city level

To gain further intuition, and to move closer to the Bartik-instrument approach used in previous studies in this literature, it is useful to substitute out the a_{it} and p_{it} terms in Eq. 7. To do so, I sum employment in an industry across all cities and then take time differences, to obtain,

$$\frac{L_{it}}{L_{it-1}} = \left(\frac{a_{it}p_{it}}{a_{it-1}p_{it-1}}\right)^{\frac{1}{1-\alpha_i-\beta_i}} \left(\frac{\phi_t}{\phi_{t-1}}\right)^{\frac{-\beta_i}{1-\alpha_i-\beta_i}} \Omega_{it} \,, \tag{12}$$

where Ω_{it} reflects how changes in city wage levels interact with the national distribution of industries across locations (determined by local industry-specific resources) to affect national industry growth rates, which I will refer to as the adjustment factor.[61] Then, substituting Eq. 12 into Eq. 7, I obtain:

$$\begin{aligned} \Delta \ln(L_{ict}) &= \Delta \ln(L_{it}) + \left(\frac{-(1-\gamma)(1-\beta_i)\lambda}{1-\alpha_i-\beta_i} \right) \Delta \ln(L_{ct}) \\ &+ \left(\frac{-\psi(1-\beta_i)-\nu}{1-\alpha_i-\beta_i} \right) \Delta \ln(C_{ct}) - \left(\frac{1}{1-\alpha_i-\beta_i} \right) \Big[(1-\beta_i)\gamma \Delta \ln(P_t) \\ &+ (1-\beta_i)\Delta \ln(v_t^*) - \Delta \ln(\epsilon_{ict}^P) + (1-\beta_i)\Delta \ln(\epsilon_{ct}^A) \Big] - \ln(\Omega_{it}) \,. \end{aligned} \tag{13}$$

This expression suggests that changes in city pollution levels ($\Delta \ln(C_{ct})$) or city congestion forces ($\Delta \ln(L_{ct})$) will cause systematic deviations between city-industry employment growth ($\Delta \ln(L_{ict})$) and the national employment growth in that industry ($\Delta \ln(L_{it})$). Thus, Eq. 13 highlights the basic intuition behind my empirical strategy.

Typically, studies using a Bartik instrument approach aggregate national industry-level shocks to obtain city-level effects. However, the vast majority of the studies in this literature do not micro-found the Bartik instrument that they use, particularly when the instrument relies on heterogeneity in industry inputs (e.g., variation in industry-level employment of skilled vs. unskilled workers). Next, I explore the extent to which my theoretical framework can be aggregated in order to motivate a city-level analysis. This exercise serves to highlight some of the issues faced in connecting existing reduced-form Bartik instrument studies to microfoundations. It also clarifies the advantages of the industry-level analysis used in most of this paper.

In order to have any hope of aggregating to the city level, we have to begin by sacrificing industry production function heterogeneity. It is still possible to incorpo-

[61] Specifically, $\Omega_{it} = \left(\sum_c w_{ct}^{\frac{\beta_i - 1}{1-\alpha_i-\beta_i}} C_{ct}^{\frac{-\nu}{1-\alpha_i-\beta_i}} \epsilon_{ict}^{P\frac{1}{1-\alpha_i-\beta_i}} \bar{R}_{ic} \right) \left(\sum_c w_{ct-1}^{\frac{\beta_i - 1}{1-\alpha_i-\beta_i}} C_{ct-1}^{\frac{-\nu}{1-\alpha_i-\beta_i}} \epsilon_{ict-1}^{P\frac{1}{1-\alpha_i-\beta_i}} \bar{R}_{ic} \right)^{-1}$. This adjustment factor reflects the extent to which national industry growth rates fail to correctly reflect the technology and demand shifts, represented by the $a_{it}p_{it}$ terms, because of changes in the wages or coal use levels occurring in different cities in which industry i is present. Note that when summing across all cities, the adjustment factor will not vary at the city level.

rate some industry-level heterogeneity in the form of industry demand shifter, as in Bartelme (2015), but I have not found a way to incorporate heterogeneity in input shares.

After setting $\alpha_i = \alpha$ and $\beta_i = \beta$, Eq. 13 can be summed to the city level to obtain:

$$\begin{aligned} \Delta \ln(L_{ct}) &= \left(\frac{-\psi(1-\beta)-\nu}{\sigma}\right)\Delta \ln(C_{ct}) + \left(\frac{-(1-\beta)\gamma}{\sigma}\right)\Delta \ln(P_t) + \left(\frac{-(1-\beta)}{\sigma}\right)\Delta \ln(v_t^*) \quad (14) \\ &+ \left(\frac{-(1-\beta)}{1-\alpha-\beta}\right)\Delta \ln(\epsilon_{ct}^A) + \left(\frac{1-\alpha-\beta}{\sigma}\right)\ln\left[\sum_i \frac{L_{ict-1}}{L_{ct-1}}\frac{L_{it}}{L_{it-1}}\left(\frac{\epsilon_{ict}^P}{\epsilon_{ict-1}^P}\right)^{\frac{1}{1-\alpha-\beta}}\frac{1}{\Omega_{it}}\right]. \end{aligned}$$

where $\sigma = 1 - \alpha - \beta - (1-\gamma)(\beta - 1)\lambda > 0$. The key thing to note in Eq. 14 is the last term on the right-hand side, which includes each industry's initial share of city employment interacted with the national industry growth rate: the building blocks of the Bartik instrument. Thus, this expression suggests that endogenous disamenities such as coal-based pollution can cause city-level employment growth to systematically diverge from what we would expect based on the initial mix of industries in a city and national industry growth rates. However, Eq. 14 also highlights that when this Bartik-style approach is applied at the city level it is at best an approximation due to the presence of the Ω_{it} term, which cannot be directly mapped to the theory. This helps explain why previous studies in the literature, with the exception of Bartelme (2015), have not offered a micro-founded Bartik estimation strategy.

This exercise illustrates the advantages of the industry-level analysis used in this paper. Running the analysis at the industry level makes it possible to derive the estimating equation directly from the theory.

A.3.2 Discussion – linking the theory and empirical results

The coefficient estimated in the main regression results (e.g., Table 1) corresponds to the exponent on the coal term in Equations 7 and 13, i.e.,

$$\frac{-\psi(1-\beta_i)-\nu}{1-\alpha_i-\beta_i}.$$

This subsection examines this expression in some detail. To begin, consider the denominator in this expression: $1-\alpha_i-\beta_i$. This value corresponds to the exponent on the local city-industry resources in the production function and therefore the share of firm costs spent on these fixed local resources. Thus, this parameter determines the importance of fixed local factors in production, which in turn determines the ease with which production can be relocated across locations. The more important are local resources, the more difficult it is to relocate production across locations. As a result, the larger is the $1-\alpha_i-\beta_i$ term, the smaller will be the response of local employment to local coal use.

Next, consider the left-hand term in the numerator, $-\psi(1-\beta_i)$. Note that $1-\beta_i = (1-\alpha_i-\beta_i)+\alpha_i$. Thus, holding fixed the importance of local resources in production, the $1-\beta_i$ term is directly related to α_i, which determines the importance of labor in production. This term is telling us that the effect of coal use on employment in a particular industry through the amenities channel will be directly linked to the importance of workers in production in that industry. This makes sense because the impact of changing amenities operates entirely through workers.

Next, consider the second term in the numerator, ν. This reflects the impact of coal use on employment through the productivity channel. Unlike the amenities term, this productivity term is not multiplied by $1-\beta_i$. This is because of the way that I have modeled the productivity effects of coal use, and in particular, the fact that local coal use affects total factor productivity, rather than specifically affecting workers. This is the simplest way to model the productivity channel and it will be realistic if coal use has effects on the productivity of other inputs. Alternatively, a more sophisticated model might focus on the effect of coal use on labor-augmenting technology only.

The coefficient estimates obtained from the empirical analysis will reflect the combined impact of all of these forces. In particular, the coefficient estimates will incorporate (1) the impact of coal use on productivity, given by the ν term, (2) the impact of coal use working through local amenities, which depends on how much coal affects amenities, given by the ψ term, and the importance of labor in the production function, reflected by the $1-\beta_i$ term, and (3) the ability of the industry to respond to these forces by shifting production across locations, which will depend on the importance of city-industry resources in production.

Finally, note that workers are paid their marginal product in the model, so that if workers become less productive because of the impact of coal use, firms will have a natural tendency to pay them less. However, in spatial equilibrium, firms in a polluted city cannot just pay less productive workers less, or else the workers will choose to go to a different city. Instead, the marginal product of workers must be increased so that their wages are consistent with spatial equilibrium. This is achieved by some workers leaving the city, so that the ratio of workers to local resources falls, which increases the marginal product of workers in order to bring the local wage back to spatial equilibrium.

A.3.3 Extension – adding capital to the model

In this appendix I consider a simple extension to the theory that incorporates capital into the model. To do so, I modify the production function to be,

$$y_{fict} = a_{ict} L_{fict}^{\alpha_i} C_{fict}^{\beta_i} K_{fict}^{\iota_i} R_{fict}^{1-\alpha_i-\beta_i-\iota_i},$$

where K_{fict} is the amount of capital used by the firm. Capital is mobile across locations and the price of capital, s_t, can vary over time.

Solving this model through, I obtain a modified version of Eq. 13:

$$\Delta \ln(L_{ict}) = \left(\frac{-(1-\gamma)(1-\beta_i-\iota_i)\lambda}{1-\alpha_i-\beta_i-\iota_i}\right)\Delta \ln(L_{ct}) + \left(\frac{-\psi(1-\beta_i-\iota_i)-\nu}{1-\alpha_i-\beta_i-\iota_i}\right)\Delta \ln(C_{ct}) \tag{15}$$
$$- \left(\frac{1}{1-\alpha_i-\beta_i-\iota_i}\right)\Big[\beta_i\Delta \ln(\phi_t) + \iota_i\Delta \ln(s_t) + (1-\beta_i-\iota_i)\gamma\Delta \ln(P_t)$$
$$- \Delta \ln(a_{it}p_{it}) - \Delta \ln(\epsilon^P_{ict}) + (1-\beta_i-\iota_i)\Delta \ln(v^*_t) + (1-\beta_i-\iota_i)\Delta \ln(\epsilon^A_{ct})\Big].$$

As this expression makes clear, adding capital to the model (at least in this simple way) does not alter the basic estimating equation. The main effect is to change somewhat the interpretation of the estimated coefficient in terms of the model parameters. To gain some intuition here, suppose that the exponent on the local resources term in the production function $(1-\alpha_i-\beta_i-\iota_i)$ does not change as a result of the inclusion of capital into the model, so that the denominator of the coefficient on the coal use term is unchanged. In this case, we can see that the impact of adding capital to the model is to affect the impact of consumer amenities on employment. In particular, the impact of consumer amenities on employment growth, which was originally determined by $-\psi(1-\beta_i)$ is now determined by $-\psi(1-\beta_i-\iota_i)$. This implies that the impact of rising coal use on local employment in industry i will be smaller when the labor share of expenditure in industry i is smaller. However, the overall implications of the model are essentially unchanged.

How will the growth in capital in a city industry respond to increasing city coal use? To see this, I follow the same procedure used for labor to solve for the change in capital across a period:

$$\frac{K_{ict}}{K_{ict-1}} = \Bigg[\left(\frac{s_t}{s_{t-1}}\right)^{-(1-\alpha_i-\beta_i)}\left(\frac{\phi_t}{\phi_{t-1}}\right)^{-\beta_i}\left(\frac{P_t}{P_{t-1}}\right)^{-\alpha_i\gamma}\left(\frac{L_{ct}}{L_{ct-1}}\right)^{-\alpha_i\lambda(1-\gamma)}$$
$$\left(\frac{C_{ct}}{C_{ct-1}}\right)^{-\psi\alpha_i-\nu}\left(\frac{v_t*}{v_{t-1}*}\right)^{-\alpha_i}\left(\frac{p_{it}a_{it}}{p_{it-1}a_{it-1}}\right)\left(\frac{\epsilon^A_{ct}}{\epsilon^A_{ct-1}}\right)^{\alpha_i}\left(\frac{\epsilon^P_{ict}}{\epsilon^P_{ict-1}}\right)\Bigg]^{\frac{1}{1-\alpha_i-\beta_i-\iota_i}}$$

This expression tells us that the growth of capital in a city-industry will also be reduced as a result of the growth in city coal use (a similar pattern will be observed for city-industry coal use). The exponent on the coal use term shows that this will occur both as a result of reduced firm productivity (the ν term) and through the consumer disamenity (the ψ term), with the impact of the consumer disamenity effect dependent on the importance of labor in the industry's production function.

It is also interesting to look at how the change in capital used in a city-industry compares to the change in labor used. To explore this, I derive:

$$\frac{K_{ict}/K_{ict-1}}{L_{ict}/L_{ict-1}} = \left(\frac{s_t}{s_{t-1}}\right)^{-1} \left(\frac{P_t}{P_{t-1}}\right)^{\gamma} \left(\frac{v_t*}{v_{t-1}*}\right) \left(\frac{\epsilon_{ct}^A}{\epsilon_{ct-1}^A}\right) \left(\frac{C_{ct}}{C_{ct-1}}\right)^{\psi}$$

This expression suggests that firms will become more capital intensive in cities in which coal using is growing more rapidly (similarly, they will also become more coal-intensive). In the current model, this effect occurs only through the consumer disamenity effect, because I have modeled the productivity effect such that it will have a symmetric effect on capital and labor.

A.4 Analysis appendix

A.4.1 Summary statistics for analysis variables

Table 9 presents summary statistics for the main analysis variables used in the industry-level analysis when all private-sector industries are included, using two-decade differences. Table 10 presents summary statistics when only manufacturing industries are included, also using two-decade differences. Table 11 presents summary statistics for the city-level analysis.

Table 9: Summary statistics for variables used in the main city-industry analysis (two decade differences)

Variable	Mean	Std. Dev.	Min.	Max.
$\Delta Ln(L_{ict})$	0.437	0.52	-5.032	3.689
$\Delta Ln(PrEMP_{ict})$	0.369	0.256	-0.151	1.251
$\Delta Ln(PrCityEMP)$	0.271	0.057	0.086	0.478
$\Delta Ln(PredCoal)$	0.372	0.176	0.136	0.852
Ln(City Patenting)	4.312	1.509	0	8.875
City Air-frost Days	39.633	9.941	22.7	56
City Rainfall	0.805	0.19	0.557	1.294
N =		4012		

Table 10: Summary statistics for analysis of manufacturing industries only

Variable	Mean	Std. Dev.	Min.	Max.
$\Delta Ln(L_{ict})$	0.393	0.523	-2.73	3.689
$\Delta Ln(PrEMP_{ict})$	0.326	0.244	-0.151	1.251
$\Delta Ln(PrCityEMP)$	0.251	0.078	0.035	0.512
$\Delta Ln(PredCoal)$	0.369	0.184	0.121	0.853
Ln(City Patenting)	4.313	1.509	0	8.875
City Air-frost Days	39.615	9.94	22.7	56
City Rainfall	0.805	0.189	0.557	1.294
N =		2312		

Table 11: Summary statistics for city-level analysis variables

Variable	Mean	Std. Dev.	Min.	Max.
Δ Emp., Analysis Industries	0.333	0.181	-0.112	0.921
Δ Emp., All Workers	0.321	0.18	-0.128	0.915
Δ Total Population	0.334	0.179	-0.039	0.915
$\Delta Ln(PrCityEMP)$	0.27	0.056	0.106	0.41
$\Delta Ln(PredCoal)$	0.372	0.176	0.169	0.838
N		155		

Note that there are slight differences between the city-level coal variable summary statistics based on the city-industry data and those based on the city data. In theory these should be the same. The sources of the differences are due primarily to the fact that there are a small number of missing city-industries observations due to zeros in the city-industry level database, which means that not all cities have the same number of observations.

A.4.2 Correlation between the key right-hand side variables

Because of the way that the $\Delta Ln(PrCityEMP)$ and $\Delta Ln(PredCoal)$ variables are constructed, it is natural that these will be correlated. Table 12 examines these correlations for different time differences. We can see that these variables show fairly high correlations in levels. However, when we look in changes the correlation drop substantially, particularly when focusing on manufacturing industries only. Because these variables appear in the regressions in changes, this suggests that the results are not being driven by a strong correlation between these variables.

Table 12: Correlations between the $Ln(PrCityEMP)$ and $Ln(PredCoal)$ variables

	All industries		
	One decade differences	Two decade differences	Three decade differences
Levels	0.8914	0.8911	0.8928
Changes	0.4793	0.2843	0.2128
	Manufacturing industries only		
	One decade differences	Two decade differences	Three decade differences
Levels	0.9252	0.9270	0.9289
Changes	0.2169	-0.0854	-0.1125

A.4.3 Additional robustness tables: City-industry analysis

Table 13 explores the robustness of my main results to the inclusion of a variety of city-level control variables, focusing on results for all industries using two-decade differences. Column 1 adds in geographic controls for air-frost days and rainfall, two important features of the British climate, as well as the level of patenting in the city in the 1850s to capture the innovative potential of the local economy.[62] In Column 2, I explore the impact of adding a control for changes in city borders. This control is omitted from most of the analysis because city border changes were an endogenous response to city growth, but it is still comforting to see that the results hold even when I control for these border changes. In Column 3, I add in controls for city size and city coal use at the beginning of each difference period. These results show that larger cities grew more slowly on average, while those with more initial coal use grew more rapidly, perhaps reflecting better access to coal deposits. Column 4 adds in an additional control for the share of the bedrock within 50km of the city that is composed of carboniferous (coal bearing) geological strata. This provides an exogenous measure of each city's access to coal deposits. I have also experimented with using the share within 10km or 100km and these deliver similar results. Column 5 adds in an indicator variable for whether a city is a major seaport. This does not substantially alter the results. I have also calculated results in which I include variables for the tonnage of shipping through the port in 1865 or the number of vessels and these deliver similar results. Finally, Column 6 presents results with London excluded from the data. In general, the signs on the city controls are as expected. For example, the positive effect of access to coal reserves is consistent with work by Fernihough & O'Rourke (2014). However, the inclusion of these variables does not change the baseline results.

[62] Data on patenting rates by location are not available after 1858.

Table 13: City-industry regression results with city-level controls

DV: Δ Log of city-industry employment (two decade differences)						
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta Ln(PredCoal)$	-2.455***	-2.194***	-1.794***	-1.513**	-1.526**	-1.980***
	(0.781)	(0.746)	(0.683)	(0.691)	(0.696)	(0.740)
$\Delta Ln(PredCityPop)$	0.243	0.334	-0.644	-0.740	-0.791	0.449
	(0.700)	(0.702)	(0.626)	(0.595)	(0.604)	(0.771)
Ln(City Patenting)	0.00657	0.00514	0.0220*	0.0258**	0.0252**	
	(0.00891)	(0.00843)	(0.0118)	(0.0115)	(0.0116)	
City Air-frost Days	-0.00376**	-0.00387**	-0.00353**	-0.00302**	-0.00386**	
	(0.00153)	(0.00157)	(0.00148)	(0.00150)	(0.00185)	
City Rainfall	-0.0691	-0.0634	-0.150	-0.251**	-0.251**	
	(0.108)	(0.102)	(0.108)	(0.112)	(0.110)	
Ln(Initial city pop.)			-0.145***	-0.0990**	-0.0953**	
			(0.0301)	(0.0388)	(0.0395)	
Ln(Initial coal use)			0.114***	0.0626	0.0630	
			(0.0311)	(0.0438)	(0.0441)	
Border Chg. Flag		0.105***				
		(0.0288)				
Carb. access (50km)				0.131*	0.127*	
				(0.0673)	(0.0684)	
Seaport flag					-0.0247	
					(0.0380)	
Constant	2.169***	1.961***	2.124***	2.085***	2.112***	1.612***
	(0.435)	(0.422)	(0.399)	(0.384)	(0.389)	(0.314)
Ind-time effects	Yes	Yes	Yes	Yes	Yes	Yes
Dropping London						Yes
Observations	4,012	4,012	4,012	4,012	4,012	3,882
R-squared	0.361	0.369	0.373	0.375	0.375	0.351

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across up to two decades. All regressions use data covering each decade from 1851-1911. City patenting uses data from 1852-1858. Air-frost days are days when the air temperature drops below freezing. Air-frost and rainfall data are from the Met. Initial city population and initial city coal use are based on the initial year for each differenced period. The border change flag indicates whether the city border changed in a period. Carb. access (50km) is the share of the bedrock within 50km of the city that is composed of carboniferous (coal bearing) geological strata. The seaport flag is an indicator for whether the city was a major seaport, as identified based on trade data from 1865.

An alternative to including the set of city-level controls is to instead include a full set of city fixed effects. Results including city fixed effects are described in Table 14.

Table 14: City-industry regression results with city fixed effects

	DV: Δ Log of city-industry employment			
	All industries		**Manufacturing only**	
Difference:	Two decades	Three decades	Two decades	Three decades
	(1)	(2)	(3)	(4)
$\Delta Ln(PredCoal)$	-1.614***	-2.311***	-1.112*	-1.911***
	(0.586)	(0.699)	(0.614)	(0.648)
$\Delta Ln(PrCityEMP)$	-0.153	0.849	-0.0729	0.714
	(0.527)	(0.568)	(0.537)	(0.662)
Ind-time effects	Yes	Yes	Yes	Yes
Observations	4,012	3,208	2,312	1,849
R-squared	0.333	0.422	0.314	0.393

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across up to two decades. Lagged predicted coal use is the predicted change in city level coal use from the two previous decades. All regressions use data covering each decade from 1851-1911, but because of the need to use changes over the previous two decades as a right-hand side variable, the outcomes variables are limited to the 1871-1911 period.

An alternative approach to the main regression specification is suggested by Eq. 13 from the theory, which expresses the change in log employment in industry i and city c relative to the change in log employment in industry i across all cities. I can implement this alternative estimation strategy using,

$$\Delta \ln(L_{ict}) = b_0 + b_1 \Delta \ln(PrEMP_{i-ct}) + b_2 \Delta \ln(PrCityEMP_{ct}) + b_3 \Delta \ln(PrCoal_{ct}) + \xi_t + e_{ict} \,, \quad (16)$$

where $\Delta \ln(PrEMP_{ict})$, is the growth in employment that we would predict given employment growth in industry i in all cities other than c, i.e.,

$$\Delta \ln(PrEMP_{i-ct}) = \ln\left(L_{ict-\tau} * GR_{i-ct,t-\tau}\right) - \ln\left(L_{ict-\tau}\right) \,.$$

Note that in Eq. 16, the full set of industry-time effects has been replaced by a set of year effects so that the $\Delta\ln(PrEMP_{i-ct})$ term can be included.

It is important to note that, unlike the specification used in the main results, the specification shown in Eq. 16 only approximates the expression suggested by the theory (Eq. 13). This is because the adjustment factor Ω_{it} is included in the error term when regressions are based on 16. In contrast, the adjustment factor is absorbed by the industry-time effects when regressions are based on Eq. 9. Thus, comparing results based on Eq. 16 to those based on Eq. 9 reveals the bias generated when the adjustment factor is left in the error term.

Table 15 presents results based on this alternative specification for lag lengths ranging from one to three decades. We can see that these results are qualitatively similar to those presented in Table 1, but the effect of coal use is consistently smaller than the effect estimated in the main text. This suggests that failing to account for the adjustment factor is biasing the results in Table 15 towards zero.

Table 15: Results using the regression specification in Eq. 16

	DV: Δ Log of city-industry employment					
	All industries			Manufacturing industries		
Difference:	One decade	Two decades	Three decades	One decade	Two decades	Three decades
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta Ln(PredCoal)$	-0.453	-1.491**	-2.278***	-0.169	-1.326***	-1.768***
	(0.515)	(0.622)	(0.689)	(0.501)	(0.452)	(0.510)
$\Delta Ln(PrCityEMP)$	-0.749	-0.163	0.579	-0.973**	-0.380	-0.0951
	(0.500)	(0.661)	(0.729)	(0.396)	(0.417)	(0.453)
$\Delta Ln(PrEMP_{ict})$	0.925***	1.000***	1.054***	0.889***	1.004***	1.041***
	(0.0409)	(0.0314)	(0.0309)	(0.0585)	(0.0473)	(0.0451)
Constant	0.301**	1.111***	1.687***	0.248	1.077***	1.578***
	(0.137)	(0.263)	(0.318)	(0.152)	(0.236)	(0.302)
Time effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,809	4,012	3,208	2,773	2,312	1,849
R-squared	0.175	0.273	0.341	0.177	0.261	0.320

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across a number of decades equal to the lag length. All regressions use data covering each decade from 1851-1911. The regressions for all industries include 26 private sector industries spanning manufacturing, services, transport, and utilities. The results for manufacturing industries only are based on 15 industries.

Another factor that may influence city-industry growth rates is the initial size of the city-industry at the beginning of a period. If this were somehow correlated with industry coal use, then this could potentially bias the results. To explore this issue, Table 16 presents results that include as a control the initial size of the city-industry (in logs) at the beginning of each period. These results show that, on average, initially smaller industries were able to achieve more rapid growth. However, controlling for this factor does not reduce the estimated impact of coal use on industry growth.

Table 16: City-industry regression results with initial industry size controls

	DV: Δ Log of city-industry employment			
	All industries		**Manufacturing only**	
Difference:	Two decades	Three decades	Two decades	Three decades
	(1)	(2)	(3)	(4)
$\Delta Ln(PredCoal)$	-2.070***	-3.213***	-2.100***	-3.376***
	(0.737)	(0.790)	(0.637)	(0.812)
$\Delta Ln(PredCityEmp)$	1.002***	2.238***	0.641	1.667**
	(0.785)	(0.828)	(0.548)	(0.702)
$Ln(L_{ict-\tau})$	-0.0402***	-0.0695***	-0.0378***	-0.0550***
	(0.00914)	(0.0124)	(0.0110)	(0.0141)
Constant	1.737***	2.757***	1.940***	3.077***
	(0.305)	(0.370)	(0.329)	(0.481)
Ind-time effects	Yes	Yes	Yes	Yes
Observations	4,012	3,208	2,312	1,849

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across up to two decades. Lagged predicted coal use is the predicted change in city level coal use from the two previous decades. All regressions use data covering each decade from 1851-1911, but because of the need to use changes over the previous two decades as a right-hand side variable, the outcomes variables are limited to the 1871-1911 period.

Table 17 considers results that include the change in local coal use in the two decades before each observation as an explanatory variable. These results can help address concerns that there may be cities that have rapidly rising coal use and slow employment growth across all periods. These results show that there is no clear relationship between coal use in the previous two decades and city-industry employment growth, regardless of whether the current predicted change in city coal use is included in the regression. Moreover, including lagged coal use has little impact on the estimated coefficient on the relationship between the predicted change in coal use in the current two-decade period and city-industry employment growth (though with fewer observations the standard errors are larger).

Table 17: City-industry regression results with lagged changes in coal use

DV: Δ Log of city-industry employment (two decade differences)				
	All industries		**Manufacturing only**	
	(1)	(2)	(3)	(4)
$\Delta Ln(PredCoal)$		-2.485**		-2.152**
		(1.001)		(0.842)
$\Delta Ln(PrCityEMP)$	-0.944	0.676	-1.953***	-0.646
	(0.733)	(1.045)	(0.436)	(0.607)
Lagged $\Delta Ln(PredCoal)$	-0.484	0.0261	0.377	0.596
	(0.448)	(0.464)	(0.417)	(0.451)
Constant	1.038***	0.980***	0.293	0.437*
	(0.179)	(0.176)	(0.234)	(0.239)
Ind.-time effects	Yes	Yes	Yes	Yes
Observations	2,400	2,400	1,382	1,382
R-squared	0.328	0.342	0.355	0.360

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across up to two decades. Lagged predicted coal use is the predicted change in city level coal use from the two previous decades. All regressions use data covering each decade from 1851-1911, but because of the need to use changes over the previous two decades as a right-hand side variable, the outcomes variables are limited to the 1871-1911 period.

Table 18 presents results obtained while including additional control variables based on several available industry characteristics: the share of salaried to wage workers in an industry, average firm size, the share of output exported, the ratio of labor costs to revenue, the share of industry workers that were female, and the share of industry workers that were under 20. Each of these controls is constructed using the same approach that was used to construct the predicted change in local industrial coal use. The data used to construct these variables are described in Appendix A.2.3.

Table 18 makes it clear that the main results are robust to the inclusion of these controls. Of the available controls, only changes in the labor intensity of local production appears to have any meaningful relationship to local employment growth. Given previous results, it is somewhat surprising to see that changes in the share of skilled workers in the city had little impact on overall city employment growth. This suggests that worker skills may have been somewhat less important in the historical setting I consider than they are in modern cities.

Table 18: Results including controls based on other industry characteristics

DV: Δ Log of city-industry employment (two decade differences)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$\Delta Ln(PredCoal)$	-2.197***	-2.300***	-2.217***	-2.688***	-2.206***	-1.857***	-2.181***
	(0.634)	(0.676)	(0.673)	(0.633)	(0.639)	(0.657)	(0.705)
$\Delta Ln(PredCityEmp)$	0.386	0.467	0.383	0.712	0.404	0.310	1.829**
	(0.561)	(0.641)	(0.564)	(0.540)	(0.540)	(0.557)	(0.818)
$\Delta Ln(Salariedwkr.shr.)$	0.0935						-1.217
	(1.152)						(1.549)
$\Delta Ln(Avg.firmsize)$		0.181					1.511**
		(0.613)					(0.754)
$\Delta Ln(Exportsshr.)$			-0.00380				1.651
			(1.089)				(1.283)
$\Delta Ln(Laborcostshr.)$				10.74**			30.41***
				(5.284)			(7.155)
$\Delta Ln(Femaleemp.shr.)$					0.123		1.144
					(0.586)		(0.796)
$\Delta Ln(Youthemp.shr.)$						-2.683	-20.44***
						(2.551)	(4.341)
Constant	1.891***	1.947***	1.906***	2.146***	1.903***	1.670***	1.587***
	(0.351)	(0.351)	(0.357)	(0.335)	(0.334)	(0.364)	(0.366)
Ind. time effects	Yes	Yes	Yes	Yes	Yes	Yes	
Observations	2,312	2,312	2,312	2,312	2,312	2,312	2,312
R-squared	0.336	0.336	0.336	0.338	0.336	0.337	0.349

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across up to two decades. These results are for manufacturing industries only, since the controls are only available for these industries. All regressions use data covering each decade from 1851-1911 with differences taken over two decade periods. Details of the construction of these control variables are available in Appendix A.2.3.

Another potential concern in the analysis is that changes in city coal use may be correlated with changes in local agglomeration forces. To address this issue, for each industry i in city c, I include controls for the change in employment in industries in that city, weighted by the amount that other industries buy from industry i, the amount that other industries supply to industry i, the demographic similarity of the other industry's workforces to the workforce of industry i, and the occupational similarity of the other industry's workforces to the workforce of industry i. These controls are labeled IOout, IOin, DEM, and OCC, respectively. The data used to construct these controls are described in detail in Appendix A.2.3.

Results obtained while including these controls are shown in Table 19. These results are for a set of 23 industries for which the connections matrices are available, with differences taken over two decades (similar results but with stronger coal use effects are obtained when taking three-decade differences). These results show that the basic relationship between rising local coal use and city-industry growth continues to be negative and statistically significant when these controls are included.

It may seems surprising that industries do not appear to benefit from employment growth among their buyer and supplier industries. This is likely due to the fact that growth in local buyers or suppliers comes with two offsetting forces. While it means more local customers or suppliers, it also means greater congestion in the city. This may explain why previous studies, such as Lee (2015), do not find strong evidence of static agglomeration forces during this period. It is important to recognize that these static agglomeration forces differ from the dynamic agglomeration forces studied by Hanlon & Miscio (2014). I have also calculated additional results in which I include controls for the dynamic agglomeration forces documented in that study. The main results are also robust to the inclusion of these controls.

Table 19: City-industry regression results with industry connections controls

DV: Δ Log of city-industry employment (two decade differences)					
	(1)	(2)	(3)	(4)	(5)
$\Delta Ln(PredCoal)$	-2.135***	-2.149***	-2.149***	-2.179***	-2.093***
	(0.743)	(0.725)	(0.718)	(0.719)	(0.748)
$\Delta Ln(PredCityEmp)$	0.723	0.726	0.716	0.691	0.762
	(0.747)	(0.753)	(0.748)	(0.756)	(0.757)
$\Delta Ln(IOin)$	-0.184				-0.196
	(0.273)				(0.282)
$\Delta Ln(IOout)$		-0.113			-0.0985
		(0.331)			(0.343)
$\Delta Ln(DEM)$			-0.0194**		-0.0193**
			(0.00801)		(0.00806)
$\Delta Ln(OCC)$				0.00987	0.0106
				(0.0658)	(0.0662)
Constant	1.794***	1.754***	1.735***	1.754***	1.788***
	(0.297)	(0.311)	(0.308)	(0.310)	(0.298)
Ind-time effects	Yes	Yes	Yes	Yes	Yes
Observations	3,549	3,549	3,549	3,549	3,549
R-squared	0.260	0.260	0.260	0.262	0.262

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across up to two decades. All regressions use data covering each decade from 1851-1911. $\Delta Ln(IOin)$ indicates the change in city employment in supplier industries, weighted by the share of industry i''s inputs from that industry. $\Delta Ln(IOout)$ indicates the change in city employment in buyer industries, weighted by the share of industry i's output that go to each industry. $\Delta Ln(DEM)$ indicates the change in employment in other city industries weighted by the correlation between the demographics (age and gender) of the workforce of that industry and the workforce of industry i. $\Delta Ln(OCC)$ indicated the change in employment in other city industries weighted by the correlation between the occupations employed in that industry and the occupations employed in industry i.

We may also be worried that the workers who choose employment in heavily coal-using industries are different in important ways than workers in other industries. For example, we may worry that they accept exposure to higher levels of pollution because they place a lower value on life, which may also make them more prone to violence or crime. That, in turn, could affect growth in other local industries. To help control for this potential issue, I exploit mortality data giving the rate of deaths due to violence. This violence measure will reflect both violent crime and accidents (including industrial accidents). Because it includes industrial accidents, and because these accidents were more common in mechanized industries, violence is related to the level of local coal use. It has also been suggested that coal smoke may have increased violence and crime because it reduced visibility. If the violence controls capture some of this effect then that will generate a downward bias in the estimated impact of pollution on cities.

Table 20 looks at the impact of including these violence controls in regressions taken over two-decade differences. Columns 1-3 look at all industries while Columns 4-6 focus on manufacturing industries only. Because the violence data are only available through 1900, the study period here is shorter than that used in the main regression results. Thus, for comparability, Columns 1 and 4 include results estimated without including controls for violence. In Columns 2 and 5 I include controls for the level of violence in the first decade of each period. More violent locations tend to experience slower growth in city-industry employment. In Columns 3 and 6, I instead control for the change in deaths due to violence across each period. Here I observe a positive relationship between the change in violence and city-industry growth. This likely reflects the fact that the violence measure includes industrial accidents, which will be increasing in growing cities.

Table 20: City-industry regression results with violence controls

DV: Δ Log of city-industry employment (two decade differences)						
	All industries			Manufacturing industries		
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta Ln(PredCoal)$	-2.114**	-2.076**	-2.140**	-3.558***	-3.624***	-3.626***
	(0.962)	(0.964)	(0.989)	(0.795)	(0.805)	(0.846)
$\Delta Ln(PredCityEmp)$	0.548	0.626	0.589	1.842**	2.110***	1.920**
	(0.974)	(1.030)	(1.016)	(0.731)	(0.772)	(0.763)
Initial violence rate		-0.0630			-0.122*	
		(0.0646)			(0.0637)	
Δ Violence rate			0.108			0.0733
			(0.167)			(0.195)
Constant	1.670***	1.673***	1.675***	2.469***	2.549***	2.496***
	(0.398)	(0.404)	(0.403)	(0.408)	(0.423)	(0.427)
Ind-time effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,411	2,411	2,411	1,388	1,388	1,388
R-squared	0.392	0.393	0.393	0.345	0.348	0.346

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across up to two decades. All regressions use data covering each decade from 1851-1900, a period in which by-cause mortality data are consistently available. The initial violence rate variable is the age-standardized mortality rate due to violence in the districts corresponding to each city in the decade starting at year $t - 20$. The Δ violence rate variable is the change in the age-standardized mortality rate due to violence in the districts corresponding to the city between the decade starting at year $t - 20$ and the decade starting at year t.

A.4.4 Heterogeneous effects of coal use

The results presented in the main text hide substantial underlying heterogeneity. To dig into some of the sources of these heterogeneous effects, I look at how the impact of coal use varies depending on the importance of labor in production. For manufacturing industries, I am able to calculate the ratio of labor costs to output value. This allows me to look at whether coal use has a stronger impact on industries where labor input costs are larger. Table 21 presents results obtained when I include the interaction between the labor cost share and the city size and city coal use variables. In Columns 1-2, I include the interaction between the coal use and city size variables with each industry's labor cost share. Instead of including industry-time effects, these regressions include a control for the change in industry employment in all other cities. This allows me to include the industry labor cost share as a separate control. Consistent with the theoretical predictions, these results suggest that the impact of local industrial coal use was stronger for more labor-intensive industries. Columns 3-4 present a similar set of results, but with industry-time fixed effects. Columns 5-6 include additional interactions with industry coal use intensity, in order to show that these effects are not driven by variation in industry coal use that is correlated with industry labor cost shares. There is also some evidence that more coal-intensive industries were less affected by either rising local coal use or increasing city size, relative to less coal-intensive industries.

Table 21: Heterogeneous effects in more labor intensive industries

DV: Δ Log of city-industry employment (two decade differences)						
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta Ln(L_{i-ct})$	0.984***	0.975***				
	(0.0471)	(0.0485)				
$\Delta Ln(PredCoal)$	-1.135**	-0.633	-0.812	0.0638	-0.776	0.0686
	(0.468)	(0.445)	(1.045)	(1.076)	(1.117)	(1.139)
$\Delta Ln(PredCoal)$*Labor Shr.	-1.171**	-1.177**	-7.023	-7.331*	-7.182	-7.585*
	(0.525)	(0.528)	(4.400)	(4.403)	(4.446)	(4.446)
$\Delta Ln(PredCoal)$*Coal Use					-0.00313	-0.000247
					(0.0236)	(0.0237)
$\Delta Ln(PrCityEmp)$	-0.541	-1.724***	-0.797	-2.326**	-1.001	-2.538***
	(0.461)	(0.531)	(0.881)	(0.952)	(0.912)	(0.963)
$\Delta Ln(PrCityEmp)$*Labor Shr.	0.933	0.905	5.920	6.154*	5.967	6.260*
	(1.278)	(1.282)	(3.688)	(3.682)	(3.701)	(3.689)
$\Delta Ln(PrCityEMP)$*Coal Use					0.0148	0.0144
					(0.0210)	(0.0210)
Industry Labor Cost Shr.	0.347	0.359				
	(0.366)	(0.368)				
Time effects	Yes	Yes				
Ind-time effects			Yes	Yes	Yes	Yes
Additional controls		Yes		Yes		Yes
Observations	2,312	2,312	2,312	2,312	2,312	2,312
R-squared	0.263	0.275	0.270	0.278		

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across up to two decades. All regressions use data covering each decade from 1851-1911 with differences taken over two decade periods. These data cover only the set of manufacturing industries for which labor cost share data are available. The additional controls included are the number of air frost days in each city, rainfall in each city, patents in the city from 1852-1858, log city population at the beginning of the period, and log city coal use at the beginning of the period.

A.4.5 The effect of coal use by major polluting industries

While the main analysis focuses on coal use as a city level variable, it is also possible to look at the impact of coal use associated with particular industries. This is useful because it can help us assess whether the results I observe are being driven by one particular coal-using industry, which would be a cause for concern. To implement this check, I calculate city level coal use coming from each of the four most intensive coal using industries – Earthenware & Bricks, Metal & Machinery, Mining, and Chemicals – as well as Textiles, which is a major coal user due to the very large size of that industry.

The results, in Table 22, show that the impact of coal use is fairly similar across most of the major coal-using industries. The fact that the estimated impact of coal use is similar across industries provides some evidence that these impacts are not picking up the influence of other pollutants. This is because we would expect the release of other pollutants that are correlated with coal use to vary substantially across industries.

The main exception appears to be Chemicals, where the impact of coal use per ton appears to be lower than for the others. This may be due in part to measurement error, exacerbated by the fact that Chemicals is the smallest of the industries studied here. Also, within the Chemicals industry coal was often used as an input into other products, such as tar, rather than being burned for fuel. As a result, when constructing coal use intensity values for this sector, it is necessary to make an adjustment for coal used by the industry that was not burned. A discussion of how this adjustment was done is available in Appendix A.2.1. However, despite this adjustment, it may be the case that the amount of coal use associated with the chemicals industry is greater than the actual amount burned by chemical companies. This would cause downward bias in the estimated impact of coal use in that sector, which may explain why the impact of coal use in this industry appears to be lower than the impact in all of the other industries studied in Table 22.

Table 22: Impact of coal use by using industry

DV: Δ Log of city-industry employment (two decade differences)		
	(1)	(2)
$\Delta Ln(PredCoal)$ – Mining	-2.296***	-2.091***
	(0.356)	(0.314)
$\Delta Ln(PredCoal)$ – Metals & Machinery	-2.258*	-1.990
	(1.281)	(1.340)
$\Delta Ln(PredCoal)$ – Textiles	-2.108***	-2.564***
	(0.631)	(0.667)
$\Delta Ln(PredCoal)$ – Earthenware & Bricks	-2.115**	-2.495***
	(0.914)	(0.957)
$\Delta Ln(PredCoal)$ – Chemicals	-0.470	-0.424
	(0.305)	(0.322)
$\Delta Ln(PredCityEmp)$	-0.768***	-1.711***
	(0.289)	(0.327)
Year effects	Yes	Yes
Additional controls		Yes
Observations	4,012	4,012
R-squared	0.375	0.392

*** p<0.01, ** p<0.05, * p<0.1. Standard errors, in parenthesis, allow correlation across industries within a city in a period and serial correlation within a city-industry across up to two decades. All regressions use data covering each decade from 1851-1911 with differences taken over two decade periods. The additional controls included are the number of air frost days in each city, rainfall in each city, patents in the city from 1852-1858, log city population at the beginning of the period, and log city coal use at the beginning of the period.

A.4.6 Instrumental variables regressions

While the main analysis uses predicted values for the key explanatory variables, it is also possible to use these predicted values as instruments for pollution levels based on observed changes in city-industry employment. However, obtaining sufficiently strong instruments requires a slightly different estimation approach that focuses on changes in the local intensity of coal use per worker. This is necessary because city-industry employment growth is impacted both by congestion forces related to growing city population and by changes in city amenities related to local pollution, but the growth in city population and the growth in local pollution will also influence each other. Focusing on the *intensity* of local coal use, using coal use per worker, helps get around this issue because, by putting local population in the denominator, it washes out congestion forces that impact all industries (including polluting industries) in a similar way. Put another way, the predicted coal use values provide a good instrument for changes in the intensity of local coal use, but have more difficulty predicting changes in the level of coal use.

Thus, for IV regressions I consider the following specification,

$$\Delta \ln(L_{ict}) = a_0 + a_1 \Delta \ln(PrCityEMP_{ct}) + a_2 \Delta \ln(CoalPW_{ct}) + \xi_{it} + e_{ict} ,$$

where $CoalPW_{ct}$ reflects the amount of coal used per private sector worker in the city. The first stage is,

$$\Delta \ln(CoalPW_{ct}) = b_0 + b_1 \ln(PrCityEMP_{ct}) + b_3 \Delta \ln(PredCoal_{ct}) + \xi_{it} + \epsilon_{ict}$$

where $\Delta \ln(PredCoal_{ct})$ is the excluded instrument.

It is worth noting that changing the key dependent variable from the log of coal use to the log of coal use per worker will not affect the estimated coefficient on the coal use term in the main regression specification. The only impact will be on the coefficient on the predicted city employment term as well as the interpretation of the estimated coefficient on $PrCityEMP_{ct}$. In particular, when I include the log of coal use as a right hand side variable, the estimated coefficient on $PrCityEMP_{ct}$ represents the impact that we would expect an increase in employment in a completely clean industry in a city to have. In contrast, when I use instead the log of coal use per worker as an explanatory variable, the estimated coefficient on $PrCityEMP_{ct}$ represents the impact that we would expect from an increase in overall employment, holding the intensity of coal use in the city constant. Because increasing overall employment while holding the intensity of coal use constant implies an increase in the overall level of coal use in the city, we should expect the coefficient on $PrCityEMP_{ct}$ to be more negative when coal per worker is used as an explanatory variable rather than coal use.

Table 23 presents the IV results. I focus here on results based on manufacturing industries. IV regressions that include all industries often do not have strong enough first-stages to allow us to draw clear conclusions, reflecting the fact, in non-manufacturing industries (which are less likely to be traded), national industry growth rates do not do as good of a job predicting actual city-industry employment growth.

These results are estimated while clustering standard errors by city-industry, to allow serial correlation, and by city-time, to allow correlated standard errors across industries within the same city in the same year. This type of clustering is somewhat more restrictive than the approach used in the main text, but is easier to implement in IV regressions.

Columns 1-2 present first-stage regression results. These show that predicted coal use is a strong predictor of actual city coal use intensity (coal per worker), with F-statistics above ten in both specifications. Columns 3-4 present the IV results. These provide evidence of a negative relationship between local industrial coal use intensity and city-industry employment growth. In general the estimated effects are somewhat smaller than those presented in the main text. Note that the coefficients on the $\Delta Ln(PredCityEMP)$ term are substantially more negative, but it is important to recognize that the interpretation of these coefficients has changed.

Table 23: IV results for manufacturing industries and two decade differences

	First-stage results DV: $\Delta \ln(CoalPW_{ct})$		IV results DV: Δ Log of city-industry emp.	
	(1)	(2)	(3)	(4)
$\Delta \ln(CoalPW_{ct})$			-1.634**	-1.115*
			(0.743)	(0.655)
$\Delta Ln(PredCoal)$	1.357***	1.264***		
	(0.412)	(0.390)		
$\Delta Ln(PredCityEMP)$	-1.424***	-1.508***	-1.944***	-2.775***
	(0.376)	(0.405)	(0.362)	(0.579)
Ind-time effects	Yes	Yes	Yes	Yes
Other controls		Yes		Yes
F-stat on excluded inst.	10.88	10.48		
Observations	2,312	2,312	2,312	2,312

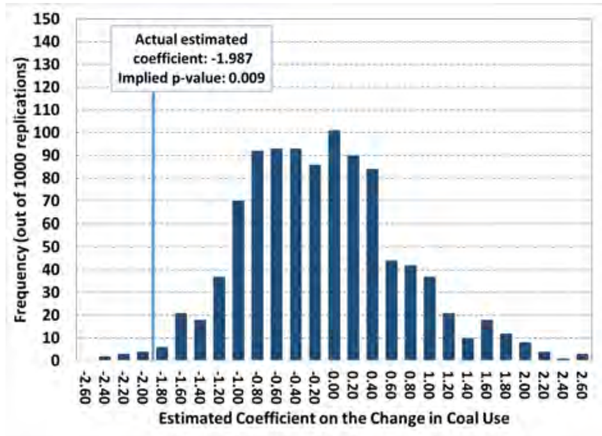
*** p<0.01, ** p<0.05, * p<0.1. Standard errors are clustered by city-year and city-industry. All regressions use data covering each decade from 1851-1911 with differences taken over two decade periods. The additional controls included are the number of air frost days in each city, rainfall in each city, patents in the city from 1852-1858, log city population at the beginning of each period, and the log of city coal use at the beginning of each period.

A.4.7 Permutation results

As an alternative approach to constructing confidence intervals, I have calculated results using the baseline specification including all analysis industries and taking two-decade differences (Column 2 of Table 1) but with industry coal use per worker values randomly reassigned across the 26 industries in the database. These results were calculated for 1000 random matches of coal use per worker values to industries. A histogram of the coefficients on the change in local industrial coal use term obtained from these 1000 permutations is presented in Figure 11.

The coefficient obtained when applying the same approach to the true data, -1.987, is also indicated on the histogram. Comparing this value to the set of values obtained when randomly allocating the industry coal use levels provides an alternative way of assessing the statistical significance of my results. In particular, only 0.9 percent of the randomly generated coefficients are more negative than the coefficient obtained

Figure 11: Histogram



using the true data, suggesting a confidence level of 99.1%.

A.4.8 Specifics of the explanatory variables in the city-level analysis

The key explanatory variables used in the city-level analysis are:

$$\Delta \ln(PrWorkpop_{ct}) = \ln\left(\sum_i L_{ict-20} * GR_{i-ct,t-20}\right) - \ln\left(\sum_i L_{ict-20}\right)$$

$$\Delta \ln(PrCoal_{ct}) = \ln\left(\sum_i L_{ict-20} * GR_{i-ct,t-20} * \theta_i\right) - \ln\left(\sum_i L_{ict-20} * \theta_i\right)$$

where $GR_{i-ct,t-20}$ is the growth rate of industry i in all cities other than c over the two-decade period.

A.4.9 Counterfactual city working populations

Table 24 describes how a ten percent reduction in the growth of coal use across the 1851-1991 period impacts the change in private-sector employment in the 31 analysis cities. I use three alternative approaches to estimating the impact of coal use. The first counterfactual uses a specification matching Column 2 of Table 1 but with each observation weighted by initial city-industry employment. The next counterfactual is based on estimates done at the city-industry level with industry-specific coal-use coefficients, also weighted by city-industry employment in 1851. The last is based on the city-level estimates shown in Column 1 of Table 3. The counterfactual estimates in Table 24 suggest that slowing the growth of coal use by just 10% would have led to substantial increases in employment in the analysis cities. The results suggest that these cities could have employed between 400,000 and one million additional workers.

A useful point to take away from these results is that the counterfactuals estimated at the city-industry level allowing for heterogeneity in the effect of coal use across industries and including industry-year controls (which will deal with the adjustment factor) are similar to the results based on the city-level estimates. This suggests that studies using a Bartik-style instrumentation approach at the city-level and abstracting from industry heterogeneity are likely to provide a reasonable approximation to the theoretically-consistent estimates done at the city-industry level including a full set of industry-time controls.

Table 24: Actual and counterfactual working population of the 31 analysis cities

		Counterfactuals		
	Actual population of 31 analysis cities	Baseline city-industry estimates	City-industry estimates with heterogeneous effect of coal use by industry	City-level estimates
1851	2,111,293	2,111,293	2,111,293	2,111,293
1881	3,274,995	3,455,590	3,744,670	3,708,641
1911	4,963,286	5,350,400	6,077,705	5,907,747
Growth: (1851-1911)	135.1%	153.4%	187.9%	179.8%

See text for details.

A.4.10 Mortality analysis

What portion of the effect of coal pollution on city growth can we attribute to increased mortality in more polluted cities? To address this question, I provide a back-of-the-envelope calculation using results from Beach & Hanlon (2016). Since I'm looking at mortality, which occurs across all city residents, I focus specifically on the impact of coal use on overall city population growth documented in Column 6 of Table 3.

For this exercise, I consider two hypothetical cities of 100,000 people, shown in the first row of Table 25. The first hypothetical city, in Column 1, grows at the rate observed aggregating all cities in my data (0.333 per decade). The second city, in Column 2, grows at the same average rate less the effect of an increase in industrial coal use that is one s.d. above the mean for the 1881 decade. From the summary statistics in Table 11, a one s.d. greater increase in industrial coal use is equal to 0.176. Multiplying this by the estimated impact of rising industrial coal use on total city population in Column 6 of Table 3 (-1.055), the decadal growth in this city should be 0.172. Thus, a one s.d. increase in city industrial coal use reduces the city population growth rate by about half.

Given these estimates, the city experiencing a more rapid rise in industrial coal use would end up with 940 fewer people after one year of growth. This difference is shown in Column 3. Next, I need to estimate the number of additional deaths that would have occurred in the more polluted city as a result of rising coal use. The results from Beach & Hanlon (2016) suggest that a one log-point increase in local industrial coal use resulted in roughly one additional death per thousand. Thus, an increase in log coal use of 0.176 is expected to generate 18 additional deaths over one year in a city of 100,000 people. As a result, rising mortality can explain about 2% of the reduction in overall population, which implies that the remaining 98% are attributable to changing migration decisions.

Table 25: The role of mortality and migration

	Benchmark city	City with one s.d. greater increase in coal use	Difference
Initial population	100,000	100,000	
Decadal population growth:	0.333	0.147	
Compound annual growth:	1.0168	1.0074	
Population after one year	101,679	100,739	940
Estimated impact of coal use on mortality (per 1000):			1
Estimated additional deaths in coal intensive city:			18
Fraction of population growth diff. explained by mortality:			0.019

A.4.11 Examining the channels

The estimated effect of coal use on real wages in each city, reported in Table 5 in the main text, corresponds to the ψ parameter in the model. Using these, together with assumptions about the production function parameters, it is possible to calculate the ν parameter, which will then allow us to think about the relative strength of the amenity and productivity channels implied by my estimates. In particular, abstracting from heterogeneity in the production function parameters, the relationship between the ψ and ν parameters is determined by,

$$\frac{-\psi(1-\beta)-\nu}{1-\alpha-\beta} = X \,, \tag{17}$$

where X is the estimated coefficient on the relationship between coal use and city growth. From this equation, we can calculate ν given our estimates of X and ψ for different assumptions on the production function parameters.

In Table 26, I calculate the ν parameter for a variety of plausible values of the production function parameters. I consider both the highest and lowest estimates of ψ from Table 5 and estimates of the coal use effect on city-industry employment growth of -1.2 and -1.5. We can see that in almost all cases the ν parameter is larger than the estimated ψ, in some cases by an order of magnitude. Moreover, note that the true effect of the amenity channel in this model, relative to the productivity channel, depends not on ψ but on $\psi(1-\beta)$. Thus, these results provide tentative evidence that the productivity channel was likely to have been more important in generating the impact of coal use on city employment than the amenity channel.

Table 26: Calculating ν for a variety of production function parameters

Estimated ψ	Labor and fixed factors share (1 - β)	Fixed factors share (1 - α - β)	Productivity effect parameter (ν) using coal effect of -1.2	Productivity effect parameter (ν) using coal effect of -1.5
Lowest:				
0.0172	0.30	0.05	0.065	0.080
	0.30	0.10	0.125	0.155
	0.50	0.05	0.069	0.084
	0.50	0.10	0.129	0.159
	0.50	0.20	0.249	0.309
	0.70	0.10	0.132	0.162
	0.70	0.20	0.252	0.312
	0.70	0.40	0.492	0.612
Highest:				
0.0504	0.30	0.05	0.075	0.090
	0.30	0.10	0.135	0.165
	0.50	0.05	0.085	0.100
	0.50	0.10	0.145	0.175
	0.50	0.20	0.265	0.325
	0.70	0.10	0.155	0.185
	0.70	0.20	0.275	0.335
	0.70	0.40	0.515	0.635