Coal Smoke and the Costs of the Industrial Revolution*

W. Walker Hanlon
UCLA and NBER

PRELIMINARY

January 20, 2016

Abstract

There is a long-running but largely unresolved debate 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, I bring together new data and a novel analysis strategy in order provide new estimates of the magnitude of these effects on local economic development. My results show that increases in pollution due to industrial local coal use substantially reduced city employment growth and slowed the overall rate of urbanization in England from 1851-1911. While this pollution was associated with increased mortality, most of the impact of pollution on city growth was driven by the location decisions of firms and workers. I also present tentative evidence that, while pollution affected the amenity value of cities, most of the effects I observe were driven by the impact of pollution on the productivity of local workers and firms. Overall, these results show that local industrial pollution can come with large hidden economic costs, a finding that also has implications for the highly polluted industrial cities found in many emerging economies today.

*I thank Dan Bogart, Leah Boustan, Karen Clay, Dora Costa, Neil Cummins, Roger Foquet, Matt Kahn, Ed Kung, Miren Lafourcade, Adriana Lleras-Muney, Alexander Rothenberg, Werner Troesken and seminar participants at UC Davis, UC Irvine, Carnegie-Mellon, and UCLA 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.

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. Writers such as Engels (1845) describe cities blackened with thick smoke and urban slums crowded with diseased and dying people. Motivated by these contemporary observations, economic historians have sought to measure the external costs of industrial growth in 19th century British cities. Existing work on this topic, such as Williamson (1981b), suggests that the external costs of industrialization were not large, but these findings have been criticized for drawing broad conclusions, which seem to contradict much of the contemporary evidence, from relatively sparse data (Pollard (1981)).[1] As a result, at present this debate remains largely unresolved.

This study revisits this debate using new data and improved theoretical and econometric tools in order to provide new results that can bring us closer to resolution. I focus my attention on the impact of pollution generated by industrial coal burning in English cities from 1851-1911. Coal smoke pollution was the most widespread negative externality associated with industrialization during this period, and it is natural to focus on English cities because of the leading role that they played in the Industrial Revolution. The temporal scope of this study covers the second half of the 19th century, a period in which industrialization and urbanization had become wide-spread and the negative health effects of living in large cities peaked.

My results imply that the pollution generated by industrial coal use substantially affected the development of English cities during this period. Specifically, in English cities that experienced rapidly rising coal use, employment growth systematically underperformed 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

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

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 12-20 percentage points, equal to one-third to one-half of the average growth in employment across two-decade periods.

These results also imply that rising industrial coal use influenced the overall rate of urbanization in England during the second half of the 19th century. To assess the magnitude of these effects, I consider a counterfactual in which the growth rate of industrial coal use is reduced by ten percent without other economic effects. Under this counterfactual I estimate that population in the 31 analysis cities in 1911 would have been larger by about 1.5 million and the share of the English population in these cities, which rose from 30% in 1851 to 34% in 1911, would instead have reached 38% in 1911, bringing the share of population in these cities closer to modern levels.

Generating these results requires overcoming two major challenges. The first is a lack of direct pollution measures in the period I study.[2] To overcome this, I combine data from the Census tracking the local composition of industries in the largest English cities with information on the amount of coal used per worker in each industry. This allows me to construct estimates of local industrial coal use in 31 cities for each decade for 1851-1911. These data show that there was large variation across cities in the amount of industrial coal use, driven by the enormous variation in the coal use intensity of different industries, together with the uneven sorting of industries across locations.

Another major challenge is presented by the lack of detailed wage and rent data

[2]One consequence of the lack of pollution measures is that relatively little quantitative work has been done studying the impact of pollution in the 19th century. One of the few exceptions is Troesken & Clay (2011), which studies the evolution of polluted fogs in London over time. Fouquet (2012) discusses long-run pollution trends in Victorian Britain. More evidence documenting the impact of coal-based air pollution becomes available starting in the middle of the 20th century, including Barreca *et al.* (2014), Clay *et al.* (2015b), and Clay *et al.* (2015a). While quantitative studies on this topic are somewhat sparse, qualitative evidence is provided by a rich set of historical work, including Mosley (2001), Thorsheim (2006), and Brimblecombe (1987).

by location in the 19th century, which are generally needed to estimate the effects of local pollution. To deal with this, I develop a new estimation approach that can be implemented using only data on industry employment over time.[3] This estimation approach is derived from a standard spatial equilibrium model, in the Rosen-Roback tradition, which has been modified to incorporate many industries that are heterogeneous in the intensity with which they use coal as an input. The level of local industrial coal use affects employment growth, either by reducing the consumer amenity value experienced by city residents, or by affecting the productivity of local firms (through the of health workers, for example).[4] An important feature of my estimation strategy is that it will capture effects on city growth operating through either of these channels.

This new estimation strategy is one important contribution of this paper. The strategy that emerges from my theory is closely related to the widely-used Bartik or shift-share instrument (Bartik (1991)), but differs from most existing studies in this vein in that my estimation strategy is derived directly from the theory.[5] Grounding the estimation strategy in theory makes the assumptions underlying the estimation approach more transparent and is helpful for interpreting the estimated coefficients.

My main results raise several additional questions. For example, how much of the impact of coal pollution on city growth was due to increased mortality? To answer this question, I study the relationship between industrial coal use and mortality using panel data from 1851-1900. My results show that a one standard deviation increase

[3]The set of existing studies looking at the impact of pollution on population migration includes Banzhaf & Walsh (2008) and Bayer *et al.* (2009).

[4]While the amenity channel is the most straightforward, the impact of pollution on productivity has been highlighted in several recent papers, including Hanna & Oliva (2015) and Graff Zivin & Neidell (2012).

[5]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.

in local industrial coal use increased the mortality rate by two deaths per thousand, or about 10%. These effects occurred primarily among the very young or those over 35, and for adults the impacts were concentrated in respiratory diseases, the category most associated with air pollution. While these are substantial effects, they can explain only about 4.4% of the overall relationship between rising coal use and city population growth. This implies that most of the impacts I observe must be due to changes in the location decisions of workers and firms.

Another important question is whether the results I observe were driven by changes in the amenity value experienced by city residents or by changes in productivity. Using a cross-section of local wage, rent and price data from 1905, I provide tentative evidence showing 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 is not large, suggesting that most of the city growth results I observe were driven by the impact of pollution on productivity which affected workers' location decisions by changing the spatial distribution of job openings. These results, while tentative, fit recent studies providing evidence that even low levels of pollution can impact productivity (e.g., Graff Zivin & Neidell (2012) and Hanna & Oliva (2015)) as well as evidence that the availability of employment was the central motivation for migration decisions in the 19th century (Pollard (1981)).[6]

Could different policies have reduced the negative effects of pollution during this period without substantially impeding economic growth? In the last part of the paper, I draw on historical sources in order to shed some light on this question. In particular, I take advantage of a detailed study of coal use in Britain undertaken by a group of experts commissioned by Parliament and documented in the extensive 1871 Coal Commission Report. This report argues that, "Without traveling beyond known principles, it was thought that a considerable saving of fuel could be effected"

[6] Graff Zivin & Neidell (2013) provide a useful review of the literature on the impact of pollution on human capital and productivity.

(p. 104). In particular, they highlight that large savings of coal could have been achieved through simple improvements to how coal was fed into boilers and furnaces, by insulating boilers and steam engines to reduce lost heat, and by implementing available technologies for reusing hot air in furnaces. Most of these changes could have been achieved with relatively modest costs.

Why, then, were these improvements not implemented? Three factors appear to have been key. First, the costs of industrial pollution were largely external to firms, giving industrialists little direct incentive to use coal more efficiently. Second, cheap coal prices reduced the incentives of producers to invest in technologies that used coal more efficiently. Third, limited and weak government regulation, a product of the strong *laissez faire* ideology that dominated British politics during this period, as well as the political power of local industrialists, failed to provide strong incentives for investments in efficiency.

These findings have implications for highly polluted industrial cities in modern developing countries, where policymakers regularly face trade-offs between encouraging economic growth and protecting the local environment. In particular, my results show that allowing industrial pollution can impose large – but often more difficult to observe – economic costs on cities.

These findings are related to several strands of the existing literature. There is a natural link to the large literature studying the health and environmental effects of pollution, which is too large to review here.[7] This literature has shown that even low levels of pollution can have important economic and health costs. Relative to existing studies in this area, the novel contribution of this paper is to provide new data and a new estimation framework that allows me to study the long-run costs of industrial pollution. There is also a natural connection between this paper and recent work on endogenous city amenities. Existing work in this area considers endogenous

[7]Within this literature, my study is most similar to Kahn (1999) and Chay & Greenstone (2003). See Graff Zivin & Neidell (2013) for one recent review.

amenities related to the composition of city residents (Rauch (1993), Moretti (2004), Shapiro (2006), Diamond (Forthcoming)). In contrast, this paper highlights how the industrial composition of cities can also affect city amenities through pollution.

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 provides a discussion of the potential for efficiency gains during this period. Section 7 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.[8] This amounted to 4.3 tons per person in 1911.[9] Most of this coal – 60-65 percent – was burned by industry.[10] Because some industries were particularly intensive users of coal, and these industries tended to agglomerate, industrial coal use could be highly geographically concentrated. 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

[8]These figures are from the U.K. Department of Energy and Climate Change.

[9]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.

[10]This includes mining, which is included in the industrial coal use measures used in this study. Data from Mitchell (1988). 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.

by workers.

The pollution released by coal burning factories 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]

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.

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. While the search for work was the primary driver of these flows, there is also some evidence that pollution levels affected location decisions, both within and across cities, as early as the 1840s. 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.*

Similarly, 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..."[13]

In addition to coping with rising pollution through relocation, there were also efforts to apply regulation. However, 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 regulations were passed, including The Sanitary

[12] *British Medical Journal* (Feb. 14, 1880, p. 254).

[13] Quoted from Thorsheim (2006).

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.[14] As a result, historical evidence suggests that their effectiveness was limited, though they may have had more impact toward the end of the century.[15]

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.[16] The occupational categories reported in these data closely correspond to industries, such as cotton spinner or steel manufacturer.[17] 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.[18]

One feature that emerges from these data is that the spread of industries across cities was far from even. For example, textile producers agglomerated in cities in Lancashire and Yorkshire, where they often accounted for one-quarter to one-third of private-sector employment. Cities such as Sheffield, Birmingham and Wolverhampton had a disproportionate share of metals industries, while ports such as Bristol and

[14]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 smokestacks emitted only dark brown smoke.

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

[16]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. Figure 2 in the Appendix includes a map of these cities.

[17]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.

[18]For further details about these data see Hanlon & Miscio (2014) and the online appendix to that paper.

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.[19] 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.[20]

These data show that coal use intensity varied enormously across industries, a feature that plays a key role in this study.[21] 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 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.[22]

This study will also draw on several useful auxiliary data sets. Data on mortality patterns in the cities I study, averaged across each decade, are available from the Registrar General's reports for 1851-1900.[23] These data include the age at which mortality occurred as well as the cause of death. I will also draw on a cross-section of data on local wages, rents and prices prepared by the Board of Trade. These data cover 51 cities in 1905.[24]

Next, I discuss how the data are used to generate a measure of local industrial

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

[20]Coal and coke are combined in this study. In practice, coke consumption is small relative to coal.

[21]A table describing coal use intensity by industry is available in Appendix A.2.5.

[22]Data describing the variation in industrial coal use across cities are available in Appendix A.2.6.

[23]Further details are available in Appendix A.2.2.

[24]For more information on the Board of Trade data, see Appendix A.2.3.

pollution in each location and period. 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:

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

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 does not vary too much over time. While I cannot check for this over 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.2.7, shows two results. First, the

[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.

relative coal use intensity across industries was remarkably stable over time. 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 impact of variation in coal use driven by differences in the relative prices of labor and coal. This is a reasonable assumption given the massive variation in industry coal use intensity.

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 often 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 to county-level coal use reported in the 1871 report, I find that my approach does a good job of replicating industrial coal use at the county level, particularly for more industrial and urbanized locations.[27]

Estimates of industrial coal use per worker at the city level are described in Table 8 in Appendix A.2.6. 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

[27] The full analysis is available in Appendix A.2.8.

Manchester and Leeds, show moderate levels, near the average. Commercial and trading cities, such as Liverpool and Bristol, as well as London, use coal less intensively. Bath 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.[28] 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 which I call coal, and a fixed local industry-specific resource.[29] The production function is,

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

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 avail-

[28]Within this tradition, I draw specifically on the recent model in Moretti (2011).

[29]In Appendix A.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.

able supply given by $\bar{R}_{ic}$.[30] 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 allowing 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}.[31] 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,

$$\frac{C_{fict}}{L_{fict}} = \left(\frac{\beta_i}{\alpha_i}\right) \frac{w_{ct}}{\phi_t} \,.$$

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. This is the variation exploited in the empirical portion of the study.

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

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

[30] This approach has recently been used in papers by Kline & Moretti (2013), Kovak (2013) and Hanlon & Miscio (2014).

[31] Cities are treated as small open economies, so output prices and the coal price are treated as exogenous.

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

The main 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,

$$\ln(r_{ct}) = \lambda \ln(L_{ct}) + \ln(\eta_c)\,, \tag{2}$$

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 housing price.[32]

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, and 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. 2, the inverse labor supply equation for city c is,

[32]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).

$$w_{ct} = P_t^{\gamma} \, L_{ct}^{(1-\gamma)\lambda} \, \eta_c^{1-\gamma} \, A_{ct}^{-1} \, v_t^* \, . \tag{3}$$

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,

$$A_{ct} = \delta_c \, CP_{ct}^{-\psi} \, , \tag{4}$$

where CP_{ct} is a measure of coal pollution in the city, δ_c represents a fixed city amenity and the ψ parameter determines the impact of local coal use on the 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} and the impact of local coal use on firm productivity, $CP_{ct}^{-\nu}$, where the parameter $\nu \geq 0$ determines the impact of local coal use on firm productivity.[33] Thus, I have $a_{ict} = a_{it} CP_{ct}^{-\nu}$.

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 in the level of city amenities due to industrial pollu-

[33] We can include an additional idiosyncratic shock to city-industry productivity here. This makes the exposition of the theory more complicated but does not meaningfully alter the results.

tion. The starting point for this derivation is the industry labor demand expression given in Eq. 1 and the city labor supply expression in Eq. 3. Differencing these expressions over time and substituting out the wage terms, I obtain,

$$\begin{aligned}\frac{L_{ict}}{L_{ict-1}} &= \left(\frac{a_{it}p_{it}}{a_{it-1}p_{it-1}}\right)^{\frac{1}{1-\alpha_i-\beta_i}} \left(\frac{\phi_t}{\phi_{t-1}}\right)^{\frac{-\beta_i}{1-\alpha_i-\beta_i}} \left(\frac{P_t}{P_{t-1}}\right)^{\frac{-(1-\beta_i)\gamma}{1-\alpha_i-\beta_i}} \left(\frac{v_t^*}{v_{t-1}^*}\right)^{\frac{-(1-\beta_i)}{1-\alpha_i-\beta_i}} \\ &\quad \left(\frac{L_{ct}}{L_{ct-1}}\right)^{\frac{-(1-\gamma)(1-\beta_i)\lambda}{1-\alpha_i-\beta_i}} \left(\frac{CP_{ct}}{CP_{ct-1}}\right)^{\frac{-\psi(1-\beta_i)-\nu}{1-\alpha_i-\beta_i}} .\end{aligned} \tag{5}$$

Eq. 5 will form the basis for one of the main empirical specifications used in this paper. In this equation, we can see that all of the terms on the top row of right-hand side of the equation will be absorbed in a regression that includes industry-time fixed effects, while the right-hand-side variables on the bottom row will capture the impact of changes in city size or local coal use on city-industry employment growth.

However, 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. 5. To do so, I sum employment in an industry across all cities and then take time differences, to obtain,

$$\frac{L_{it}}{L_{it-1}} = \left(\frac{a_{it}p_{it}}{a_{it-1}p_{it-1}}\right)^{\frac{1}{1-\alpha_i-\beta_i}} \left(\frac{\phi_t}{\phi_{t-1}}\right)^{\frac{-\beta_i}{1-\alpha_i-\beta_i}} \Omega_{it} , \tag{6}$$

where $\Omega_{it} = \left(\sum_c w_{ct}^{\frac{\beta_i-1}{1-\alpha_i-\beta_i}} CP_{ct}^{\frac{-\nu}{1-\alpha_i-\beta_i}} \bar{R}_{ic}\right)\left(\sum_c w_{ct-1}^{\frac{\beta_i-1}{1-\alpha_i-\beta_i}} CP_{ct-1}^{\frac{-\nu}{1-\alpha_i-\beta_i}} \bar{R}_{ic}\right)^{-1}$. In this expression, 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. I will refer to this as the adjustment factor.[34]

[34] The 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. Importantly, when summing across all cities, the adjustment factor will not vary at the city level.

Then, substituting Eq. 6 into Eq. 5, I obtain:

$$\begin{aligned}\frac{L_{ict}}{L_{ict-1}} \quad = \quad & \left(\frac{L_{it}}{L_{it-1}}\right)\left(\frac{L_{ct}}{L_{ct-1}}\right)^{\frac{-(1-\gamma)(1-\beta_i)\lambda}{1-\alpha_i-\beta_i}}\left(\frac{CP_{ct}}{CP_{ct-1}}\right)^{\frac{-\psi(1-\beta_i)-\nu}{1-\alpha_i-\beta_i}} \\ & \left(\frac{P_t}{P_{t-1}}\right)^{\frac{-(1-\beta_i)\gamma}{1-\alpha_i-\beta_i}}\left(\frac{v_t^*}{v_{t-1}^*}\right)^{\frac{-(1-\beta_i)}{1-\alpha_i-\beta_i}}\Omega_{it}^{-1}\,.\end{aligned} \tag{7}$$

In this equation, the change in city-industry employment is expressed as a function of (1) the national growth rate of that industry, which reflects industry-level technology and demand shocks up to the adjustment factor Ω_{it}, (2), a city-level congestion force related to city size, (3), the impact of changing local pollution levels, and (4), national changes in the level of prices and the outside option of workers. With the exception of the adjustment factor, all of these elements are either observed in the data or can be controlled for using time effects. This expression, which motivates my main empirical approach, makes it clear that changes in city pollution levels or city congestion forces will cause systematic deviations between city-industry employment growth and the national employment growth in that industry.

Finally, if we are willing to ignore industry heterogeneity (i.e., $\alpha_i = \alpha$ and $\beta_i = \beta$), then Eq. 7 can be summed to the city level. In this case, we have,

$$\frac{L_{ct}}{L_{ct-1}} = \left(\frac{CP_{ct}}{CP_{ct-1}}\right)^{\frac{-\psi(1-\beta)-\nu}{\sigma}}\left(\frac{P_t}{P_{t-1}}\right)^{\frac{-(1-\beta)\gamma}{\sigma}}\left(\frac{v_t^*}{v_{t-1}^*}\right)^{\frac{-(1-\beta)}{\sigma}}\left[\sum_i\left(\frac{L_{ict-1}}{L_{ct-1}}\frac{L_{it}}{L_{it-1}}\frac{1}{\Omega_{it}}\right)\right]^{\frac{1-\alpha-\beta}{\sigma}}. \tag{8}$$

where $\sigma = 1-\alpha-\beta-(1-\gamma)(\beta-1)\lambda > 0$. Note that, in Eq. 8, the last term on the right-hand side includes each industry's initial share of city employment interacted with the national industry growth rate: the building blocks of the Bartik instrument. This expression shows that endogenous city disamenities will cause city-level employment to systematically diverge from what we would expect based on the initial mix of industries in a city and national industry growth rates.

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. I then study the impact of coal use on mortality in cities and consider the share of the impact of coal use on city size that can be explained by increased mortality rates. Finally, I provide some tentative evidence on the channels through which coal use may have affected city growth.

5.1 Coal use and city-industry employment growth

The starting point for this analysis is the following regression specification,

$$\Delta \ln(L_{ict}) = b_0 + b_1 \Delta \ln(PrEMP_{i-ct}) + b_2 \Delta \ln(PrCityEMP_{ct}) + b_3 \Delta \ln(PredCoal_{ct}) + \xi_t + e_{ict} \,, \quad (9)$$

where the dependent variable is the change in log city-industry employment over some period (usually two decades) and ξ_t is a full set of time effects. The first variable on the right-hand side of this equation, $PrEMP_{ict}$, is the growth in employment that we would predicted given employment growth in industry i in all cities other than c,

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

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. The theory suggests that $PrEMP_{i-ct}$ should be a good predictor of actual city-industry employment growth on average, but that changes in either city congestion or city amenities will cause a systematic deviation between this predicted city-industry growth and actual city-industry growth. This is why I call this a *deviations from Bartik* regression approach.

The next variable on the right-hand side, $PrCityEMP_{ct}$ will capture changes in city congestion forces related to city size. This variable is defined as,

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

which 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 I use predicted changes in city-industry employment growth rather than actual changes here, since actual changes will be affected by city-level shocks, which may also influence employment growth in industry i. Also, when studying industry i, that industry is dropped when constructing $PrCityEMP_{ict}$.[35] This helps us avoid endogeneity concerns, but ultimately it does not have a large impact on the results.

The third term of the right-hand side of Eq. 9, $PredCoal_{ct}$, will capture changes in city amenities or productivity due to changes in local coal use. This variable is defined as,

[35]In practice this will cause $PrCityEMP_{ct}$ 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.

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

where θ_j is coal use per worker in industry j. It is important to note that the only difference between $PredCoal_{ct}$ and $PrCityEMP_{ct}$ is due to variation in the coal intensity of industries, represented by θ_j. As in the previous variables, $PredCoal_{ct}$ is based on predicted values generated using initial industry composition and industry growth rates in all other cities. Industry i is dropped when calculating $PredCoal_{ct}$ in order to avoid endogeneity concerns.[36] Note that this variable differs from $PrCityEMP_{ct}$ only due to differences in the intensity of coal use across industries, reflected by θ_j.

While Eq. 9 provides a good starting point for the analysis, I will also consider regressions in which the $\ln(PrEMP_{i-ct})$ term is replaced with a full set of industry-time effects. These regressions will more closely match the specification suggested by the theory (Eq. 7), except that they abstract from heterogeneous industry responses to changing levels of city pollution or city congestion forces. Later, I will also present results that explore these heterogeneous responses.

The theory makes predictions about the coefficients we should estimate using Eq. 9. In particular, the theory suggests that b_1 should be close to one, and that b_2 and b_3 should be negative.[37] Thus, the theory predicts that actual city-industry employment growth will systematically under-perform relative to national employment growth in that industry in cities that experience a more rapid rise in city coal use. In Appendix A.4.2, I present scatterplots showing that, indeed, there is a systematic negative relationship between the deviation between actual and predicted city-industry employment growth and the predicted change in local coal use, consistent with the

[36] In practice this will cause $PrCityEMP_{ct}$ 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.

[37] I have not included a city-size agglomeration force in the model. If this is included, then the sign of b_2 is not certain to be negative.

predictions of the theory.

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

This specification is estimated using pooled cross-sections, 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).[38]

I begin the analysis, in Table 1, by exploring results with differences taken over time periods ranging from one to three decades. In the top panel, I follow the specification described in Eq. 9, which includes predicted city-industry employment growth based on the employment growth observed in that industry in all other cities on the right-hand side. 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

[38]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.

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. In the bottom panel of Table 1 I present results from regressions in which I include a full set of industry-time effects in place of the $\Delta Ln(PrEMP_{ict})$ variable.

Table 1 reveals several important patterns. First, the top row of the table shows that, on average, industry growth in all other locations does a reasonable job of predicting actual city-industry growth, i.e., the coefficient on this term is always close to one. Second, in most specifications there is evidence that employment growth in other city-industries has a negative relationship to observed city-industry employment growth, consistent with a city-size congestion force. Interestingly, this impact appears to fall over longer differences, suggesting that long-run congestion effects may be weaker than short run effects. 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. Third, the coal use variable always has a negative impact on city-industry employment growth. This impact grows stronger 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 and firms to react to the impact of rising industrial coal use. It may also reflect that pollution impacts may depend in part on cumulative exposure.[39]

[39]It is also useful to note that the R-squared values show that the available variables are able to explain more of the variation in city-industry growth as we look over longer time periods. 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.

Table 1: Baseline city-industry regression results

	Panel A: With predicted city-industry employment growth					
	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)
$\Delta Ln(PrEMP_{ict})$	0.931***	1.029***	1.068***	0.919***	1.029***	1.048***
	(0.0377)	(0.0302)	(0.0299)	(0.0597)	(0.0472)	(0.0452)
$\Delta Ln(PrCityEMP)$	-0.787*	-0.339	0.300	-0.933**	-0.444	-0.218
	(0.445)	(0.614)	(0.684)	(0.397)	(0.428)	(0.454)
$\Delta Ln(PredCoal)$	-0.389	-1.374**	-2.005***	-0.207	-1.244***	-1.593***
	(0.505)	(0.610)	(0.661)	(0.522)	(0.473)	(0.504)
Constant	0.276**	1.064***	1.559***	0.242	1.019***	1.461***
	(0.138)	(0.266)	(0.311)	(0.158)	(0.243)	(0.293)
Time effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,819	4,017	3,213	2,773	2,312	1,849
R-squared	0.211	0.313	0.381	0.198	0.286	0.343
	Panel B: With a full set of industry-time effects					
	DV: Δ Log of city-industry employment					
	All industries			Manufacturing industries		
Difference:	One decade	Two decades	Three decades	One decade	Two decades	Three decades
	(7)	(8)	(9)	(10)	(11)	(12)
$\Delta Ln(PrCityEMP)$	-0.556	0.152	1.016	-0.700	0.268	0.987
	(0.521)	(0.705)	(0.787)	(0.539)	(0.570)	(0.687)
$\Delta Ln(PredCoal)$	-0.584	-1.822**	-2.676***	-0.485	-2.095***	-3.021***
	(0.614)	(0.721)	(0.788)	(0.733)	(0.666)	(0.806)
Constant	0.557***	1.585***	2.393***	0.341	1.736***	2.621***
	(0.163)	(0.312)	(0.372)	(0.227)	(0.343)	(0.470)
Ind.-time effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,819	4,017	3,213	2,773	2,312	1,849
R-squared	0.288	0.391	0.463	0.265	0.362	0.427

*** 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.

In Appendix A.4.3 I present a variety of additional results exploring the robustness of these estimates. One set of results, in Appendix Table 13, includes additional city-level control variables, for factors such as weather (rainfall, air frost days) and innovation (using patent data), that might have influenced city growth. These variables have the expected effect on city-industry employment growth and including them does not alter my basic findings. I also calculate results while controlling for changes in city border. I omit this variable from most regressions because border changes are endogenous, but it is comforting to see that controlling for these does not alter my results. I also calculate regressions that include city employment and city coal use at the beginning of each period. These will help control for factors such as a city's initial access to coal reserves or other natural advantages that may have dynamic effects. Including these variables does not substantially change the results. Finally, I show that dropping London, the main outlier in the data, does not substantially affect the results.

We may worry that some cities are simply characterized by rapid growth in coal use but relatively slow growth in overall employment across all periods. To address this concern, I calculate results that include the predicted change in city-industry coal use in the previous two-decade period as an explanatory variable. These results, in Appendix Table 14, show that there is no clear relationship between the lagged predicted change in city coal use and current city-industry employment growth. Moreover, including these lagged variables does not substantially affect the estimated relationship between the predicted change in coal use in the current period and employment growth.

Another potential concern in this identification approach is that there may be other features that vary across industries in a way that is correlated with the feature of interest (coal use, in this case) and also affect the outcome variable. For example, recent work suggests that the skill intensity of industries may influence city wages, productivity and amenities (Rauch (1993), Moretti (2004) and Diamond (Forthcoming)). To help deal with this potential concern, I take advantage of information from the Census of Manufactures describing several industry features: average firm size, salaried worker share, share of output exported, and labor cost share. These data are used to generate city-level control variables using the same approach that I used to construct the predicted change in city coal use. Table 15 in the Appendix shows that including these controls does not meaningfully affect the estimated impact of coal use on city-industry employment growth.

We may also be worried that changes in city coal use are correlated with changes in local agglomeration forces. In particular, building on ideas put forward by Marshall (1890) and Jacobs (1969), recent studies such as Rosenthal & Strange (2001), Ellison *et al.* (2010), and Faggio *et al.* (2013) provide evidence suggesting that industries benefit from being near their buyers or suppliers, industries using similar labor forces, or industries sharing technological links. To control for these forces, I use data on input-output connections and labor force similarity between each pair of industries in order to construct control variables representing the local change in employment in buyer industries, supplier industries, or industries employing workforces that are demographically or occupationally similar.[40]. Including these additional controls has only small effect on the estimated impact of city coal use on employment growth, which remains statistically significant.

It is also possible to estimate instrumental variables 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. Instrumental variables regressions of this type are provided in Appendix A.4.5. These results also provide evidence that rising city coal use had a statistically significant effect on employment growth, though the estimated coefficients, which range from -0.670 to -0.833, are smaller than those obtained in Table 1.

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 -0.67 to

[40]A reliable measure of the technological similarity of industries is not available for this period, so I am not able to include a control for this agglomeration force.

-2, with most estimates falling between -0.8 and -1.2. To interpret these estimates, it is useful to know that the average increase in log city coal use across all periods was 0.367 with a standard deviation of 0.178. 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 12-25 percentage points over two decades. Average city-industry employment growth across all cities and periods was 43.7 percent. Thus, a one s.d. greater increase in city coal use would be expected to reduce city-industry employment growth by between one-third and one-half.

This implies 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 effect of local pollution.

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 (though 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 exactly what I find.[41] In particular, in the regression results shown in Appendix A.4.4, I observe negative and statistically significant coefficients on both the coal use variable and the interaction between the coal use and industry labor cost share variables. This shows that coal

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

use had a stronger impact on industries that were more dependent on labor inputs. This result continues to hold when I also include the interaction between city coal use and each industry's coal use, defined as coal costs relative to revenue. These results match our expectations about how pollution is likely to affect employment growth, as well as the predictions of the theory, providing some additional confidence in the results.

Next, I shift my attention to estimating the impacts on overall city employment or population, results which will be useful in thinking about the overall impact of rising coal use on urban growth.[42] The city-level regression specification is,

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

where $\Delta \ln(PrWorkpop_{ct})$ 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.[43] 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

[42]Note that these results do not map directly to the theory because they abstract from heterogeneous industry responses to rising coal use. Also, the theory is focused on growth in working population and has less to tell us about the factors that influence the growth of the non-working population of cities.

[43]Specifically,

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

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

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

growth on overall city employment as well as any negative congestion effects generated by increasing population. Thus, we should not expect this value to be negative. Rather, abstracting from industry heterogeneity and the impact of the adjustment factor, Eq. 8 suggests that we should observe $a_1 \in (0,1)$.

City-level results are presented in Table 2. 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.[44] 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.

Table 2: City-level regression results

DV: Δ Log of city employment in analysis industries						
	City employment in analysis ind.		Total city working pop.		Total city population	
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \ln(PrWorkpop_{ct})$	1.072	0.554	0.829	0.163	0.354	-0.222
	(0.694)	(0.743)	(0.695)	(0.747)	(0.645)	(0.754)
$\Delta \ln(PrCoal_{ct})$	-1.668**	-1.880**	-1.501**	-1.551**	-0.992	-1.017
	(0.716)	(0.736)	(0.714)	(0.732)	(0.700)	(0.756)
Time effects	Yes	Yes	Yes	Yes	Yes	Yes
Other controls		Yes		Yes		Yes
Observations	155	155	155	155	155	155
R-squared	0.072	0.206	0.086	0.207	0.104	0.208

*** 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.

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

5.2 Implications for urbanization levels

How large, in aggregate, was the impact of coal-based pollution on urban growth in England from 1851-1911? To shed some light on this question, I consider as an example a simple counterfactual in which the *growth* of coal use across each decade was reduced by ten percent without imposing large economic costs. While this represents a meaningful reduction in the growth of coal use, later, in Section 6, I present evidence suggesting that this was likely to have been achievable given the available technologies at relatively low cost. While these simple counterfactuals include some important assumptions, they can be useful in shedding light on the potential magnitude of the impacts of coal use on overall urbanization rates.[45]

Table 3 describes how this impacts the change in private-sector employment in the 31 analysis cities over the study period. I use four alternative approaches to estimating the impact of coal use. The first counterfactual uses the baseline estimates from Column 3 of Table 1. The next counterfactual is based on estimates done at the city-industry level with industry-specific coal-use coefficients. The third counterfactual includes industry-specific coal-use coefficients and adds a full set of industry-decade controls to the regressions. The last is based on the city-level estimates shown in Column 1 of Table 2.

The counterfactual estimates in Table 3 suggest that slowing the growth of coal use by just 10% would have led to substantial increases in employment in the analysis cities. My preferred estimates, in the fourth column, suggest that these cities could have held half a million additional private-sector workers after three decades, and

[45]Note that this counterfactual includes assumptions about the elasticity of supply of workers. In particular, my estimating strategy allows for each city to face an upward-sloping city labor supply curve, and it allows for all of these curves to 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 assumption implicit in this counterfactual is that the the supply curve for workers to any particular city is not affected by the growth of the other analysis cities. For moderate changes in the growth of coal use, this assumption may not be too far off 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).

Table 3: Actual and counterfactual working population of the 31 analysis cities

	Actual population of 31 analysis cities	**Counterfactuals**			
		Baseline city-industry estimates	City-industry estimates with heterogeneous effect of coal use by industry	City-industry estimates with heterogeneous effect of coal use by industry and industry-year controls	City-level estimates
1851	2,117,154	2,117,154	2,117,154	2,117,154	2,117,154
1881	3,278,360	3,541,119	3,673,214	3,760,396	3,775,766
1911	4,972,994	5,536,270	5,817,268	6,106,211	6,056,537
Growth: (1851-1911)	134.9%	161.5%	174.8%	188.4%	186.1%

around a million by 1911. An important 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 very similar to the results based on the city-level estimates. This is a useful finding, because it suggests that studies using a Bartik-style instrumentation approach at the city-level and abstracting from industry heterogeneity can generate results that are quite similar to the theoretically-consistent estimates done at the city-industry level.

Table 4 considers the same counterfactual for total city population. These results show that the total population of the 31 analysis cities in 1911 would have been larger by about 1.5 million under the counterfactual. As a result, these cities would have included 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.

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

Year	Actual		Counterfactual	
	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	9185357	0.37
1911	11,626,649	0.34	13,072,470	0.38

5.3 The role of mortality

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, and to provide additional evidence on the impact of industrial coal use in British cities, I consider the relationship between coal use and local mortality rates. Industrial pollution can raise city mortality rates by directly impacting the health of residents, or because less healthy populations may sort into more polluted cities. In this section I examine the overall relationship between pollution and mortality, which will reflect both of these channels, since both of these are important for the relationship between pollution and city growth. In addition, I will also provide more suggestive evidence giving a lower bound on the direct causal effect of pollution on mortality for adults.

This analysis is conducted using panel data describing mortality across each decade from 1851-1900 in the districts covering each of the 31 cities in the main city-industry database.[46] Regressions are run separately for different age categories. The specification is,

$$MR_{ct}^{age} = \alpha_0 + \alpha_1 \ln(COAL_{ct}) + \alpha_2 \ln(COAL_{ct-10}) + \alpha_3 \ln(DEN_{ct}) + \kappa_c + \xi_t + \epsilon_{ct} \quad (11)$$

[46]Summary statistics for the mortality data are available in Appendix A.4.1, Table 12.

where MR_{ct}^{age} is the mortality rate for age group age in city c averaged across the decade starting in year t, $COAL_{ct}$ is coal use in year t, DEN_{ct} is population density in year t, κ_c is a full set of city fixed effects and ξ_t is a full set of year effects. I include both current and lagged coal use in this specification to allow past coal use levels to influence future mortality rates. This will be particularly important for older adults, where mortality can be influenced by long-run pollution exposure.[47] Density is included because crowded living conditions were an important factor in influencing mortality levels during this period.

The estimated coefficients on both contemporaneous and lagged coal use are described in Figure 1, while full regression results are available in Appendix A.4.6. These results show evidence that coal use was associated with a large increase in mortality for children under five as well as some evidence of increased mortality for older adults associated with past pollution exposure. We can see that lagged coal use is associated with increased total mortality for ages over 35, but these results are only statistically significant for those 65-74. While the estimates shown in Figure 1 are not precise, the coefficients imply large effects. Using national mortality rates in 1881 as a baseline, a one log-point increase in industrial coal use is associated with over two additional deaths per thousand, equal to a 10% increase in the mortality rate.[48] Overall industrial coal use increased by 0.97 log points from 1855-1903.[49] Thus, rising industrial coal use was associated with a substantial increase in the mortality rate in the analysis cities. Note that these results could partially reflect the sorting of less healthy populations into more polluted areas, so they should not be interpreted as the direct causal effect of pollution on health.

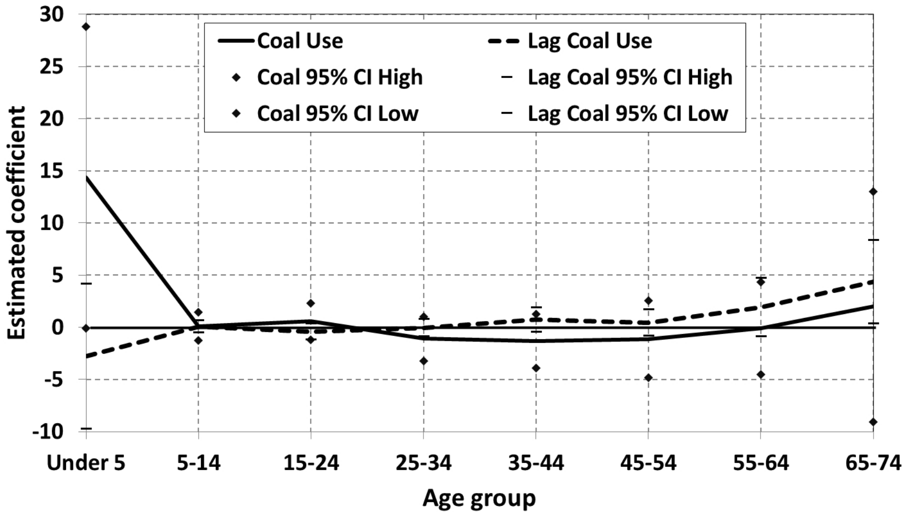
[47] One caveat to this analysis is that people could have been exposed to pollution in one location but died in another. This would bias the estimated effects towards zero. This is unlikely to be a major concern for children but it may be an important concern for the effects estimated for adults. Thus, the adult mortality effects that I estimate should be thought of as a lower bound on the true impact of pollution.

[48] These results are based on the age distribution of the population and the overall mortality rate in 1881, the middle of the study period.

[49] Based on data from Mitchell (1984).

Figure 1: Estimated effect of contemporaneous and lagged coal use on mortality

DV: Total mortality rate



These figures show coefficients and 95% confidence intervals for the coal use and lagged coal use terms. The results are generated using the regression specification shown in Eq. 11. Standard errors are clustered by city to allow for serial correlation.

It is also possible to look at the impact of coal use on specific mortality categories. This analysis, presented in Appendix A.4.6, suggests that local industrial coal use was associated with an increase in the share of deaths due to diseases of the respiratory system, the category that is most closely associated with the effects of air pollution. In particular, I find a statistically significant relationships between lagged industrial coal use and mortality from respiratory diseases starting at age 35, as well as a statistically significant increase in the share of mortality due to respiratory diseases.

One way to obtain a lower-bound estimate of the causal effect of pollution on mortality is to consider *excess respiratory mortality*, defined as the excess percentage increase in respiratory mortality associated with coal use above the percentage increase in total mortality. Excess respiratory mortality will be free of selection effects or omitted variables concerns under reasonable assumptions.[50] Focusing on excess

[50]In particular, excess respiratory mortality will be free of selection concerns if the populations that select into more polluted locations are less healthy across a broad set of disease categories,

respiratory mortality only for those 35 and over, I find that a one log-point increase in industrial coal use is associated with an additional 0.25 deaths per thousand, or a 1% increase in the overall mortality rate.[51] Thus, even these results, which should be thought of as a conservative lower bound, imply an important direct causal impact of industrial coal use on mortality.

How much of the impact of industrial pollution on city population growth documented in the previous section can be explained by these increased mortality rates? To gain insight into this question, I conduct a back-of-the-envelope comparison of the expected increase in mortality due to rising industrial coal use to the impact that I expect rising coal use to have on overall city population, based on the estimates in Column 6 of Table 2. These calculations, described in Appendix A.4.7, show that increased mortality can explain just 4.4% of the estimated impact of rising coal use on city population growth. This implies that most of the impact of coal use on city population occurred through changing migration decisions.

5.4 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.[52] 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 results we have observed.

rather than being specifically unhealthy in terms of respiratory diseases. Similarly, this will be free of concerns about omitted variables that are correlated with coal use and affect local mortality rates as long as the omitted variables do not specifically affect just respiratory diseases.

[51]These results are based on the age distribution of the population and mortality rates in 1881, the middle of the study period.

[52]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.

To begin, I use the model to derive an standard expression relating the quality-of-life in cities to local amenities. Starting with the indirect utility function and substituting in Eq. 4, I obtain,

$$[\gamma \ln(P_t) + (1-\gamma)\ln(r_{ct})] - \ln(w_{ct}) = \ln(\delta_c) - \psi \ln(CP_{ct}) - v_t^* \,. \tag{12}$$

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.[53] Estimating this equation allows me to obtain the parameter ψ, which determines how local coal use affects city employment growth through the amenity channel.[54]

These regressions are run using wage data for skilled builders and skilled engineers, occupations that are found in most or all of the cities.[55] 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.[56] In all specifi-

[53]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. Of these, non-wage income and income taxes are not a concern in my empirical setting. I 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.

[54]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.

[55]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). Further details on the Board of Trade data are presented in Appendix A.2.3.

[56]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 use the implementation from Hsiang (2010). 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.

cations, city coal use is negatively related to the amenity value of the city, and this relationship is statistically significant in most of the results.

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.4.8 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, 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.

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**	-0.0458*	-0.0294***	-0.0452**	-0.0219
	(0.00946)	(0.0203)	(0.0240)	(0.0108)	(0.0174)	(0.0266)
$Ln(POP_c)$		0.0421**	0.0334		0.0185	-0.00326
		(0.0208)	(0.0245)		(0.0187)	(0.0273)
Controls			Yes			Yes
Observations	51	51	51	47	47	47
R-squared	0.053	0.133	0.206	0.139	0.153	0.206

*** 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: city latitude, air frost days, rainfall, and the share of salaried workers in the city as predicted by industry salaried worker shares and the local composition of industries.

6 Discussion

The results above suggest that moderate reductions in the growth of coal use could have led to substantially increased urban growth, while also reducing harmful health effects. But how difficult would it have been for the British to reduce industrial coal use during this period? To answer this question, I draw on the findings of the 1871 Coal Commission report. This extremely detailed report was commissioned by Parliament and produced by leading experts.[57] The task was divided among subcommittees, with Committee B assigned to, "inquire whether there is reason to believe that coal is wasted by carelessness or neglect of proper appliances for its economical consumption."

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. The committee highlighted three major areas where improvements could be made. The first was the procedures used for adding coal to boilers.[58] 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,

[57]The report stretches to over 1300 pages. Among the luminaries interviewed by the committee focused on the topic of savings in coal use were Henry Bessemer, the inventor of (among many other things) the Bessemer process, and Charles William Siemens, the inventor of the regenerative furnace.

[58]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."

> *...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; most especially such as are used in collieries and iron workers, 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)

Finally, while efficiency improvements had been achieved through the adoption of new technologies, particularly in the iron and steel industries, the report suggests that there was scope for further improvements, though these required more costly investments. For example, they interviewed Sir Charles William Siemens, a civil engineer and inventor of the regenerative furnace, who stated that,

> *...there appears to be a saving of fuel [from using regenerative furnaces], which amounts in many cases to nearly 50 per cent. The cost however, of erecting the regenerative gas furnace is rather more than double that of the ordinary furnaces...of the 6,243 puddling furnaces at work in 1869 a very small number of them had adopted any arrangement for utilizing the waste heat, consequently even at the lowest estimate the loss of heat must be enormous.* (p. 101)

Thus, it appears that by the middle of the study period there existed several alternative approaches for reducing the use of coal by industry. At least two of these approaches – improved procedures for feeding coal into furnaces, and improved insulation to reduce heat loss – could potentially be implemented at low cost.

Why, then, were these efficiency gains not achieved? Three factors appeared to play a key role. First, coal was abundant and relatively inexpensive. The committee writes that, "in places where coal is cheap 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. Under the circumstances which prevail in most of our iron producing districts, coal being cheap, there has not been, until very recently, much attention given to economy of use" (p. 129). Second, while several pollution regulations were passed during this period, they were generally weak and largely ineffective ((Thorsheim (2006), Fouquet (2012)). This was due to a combination of the strong *laissez faire* ideology that dominated British policy-making during this period and the influence of local industrialists. Third, coal pollution imposed city-level externalities, which meant that producers had little incentive to unilaterally reduce their coal consumption.

7 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 substantial costs for local economic growth, influenced mortality patterns, and slowed overall urbanization in England.

These results agree with qualitative historical work, as well as contemporary reports, emphasizing the severity of the industrial pollution problems faced by English cities in the 19th century. However, they differ substantially from previous quantitative work on this topic, particularly a series of papers by Jeffrey Williamson, who used evidence similar to the quality-of-life regressions described in Section 5.4 to conclude that the negative externalities of industrial pollution were "trivial".[59] One key difference between our studies is that the quality-of-life regressions will reflect only resident's responses to observable disamenities, while my analysis will also capture other impacts of industrial pollution, such as it's impact on productivity, regardless

[59](Williamson, 1981b, 1982).

of whether contemporaries were aware of these effects. In addition, this study draws on a much richer set of data than previous work.

The problems of industrialization and pollution experienced by 19th century English 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. The results of this study provide a warning against ignoring the economic consequences of local pollution, as well as tools that can help measure these consequences in relatively data-sparse settings. Thus, one promising avenue for future work is to apply the tools developed here to study the impact of pollution in high-pollution modern settings.

References

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

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.

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.

Chay, Kenneth Y., & Greenstone, Michael. 2003. The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession. *The Quarterly Journal of Economics*, **118**(3), pp. 1121–1167.

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

Clay, Karen, Lewis, Joshua, & Severnini, Edson. 2015b (April). *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, T.G. 1999. GMM Estimation with Cross Sectional Dependence. *Journal of Econometrics*, **92**(1), 1 – 45.

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

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.

Faggio, Guilia, Silva, Olmo, & Strange, William. 2013 (October). *Heterogeneous Agglomeration.* Mimeo.

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

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

Graff Zivin, Joshua, & Neidell, Matthew. 2013. Environment, Health, and Human Capital. *Journal of Economic Literature*, **51**(3), 689–730.

Haines, Michael R. 2001 (July). *The Urban Mortality Transition in the United States, 1800-1940.* NBER Working Paper, Historical Paper 134.

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.

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.

Jacobs, Jane. 1969. *The Economy of Cities.* New York: Vintage Books.

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.

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

Marshall, Alfred. 1890. *Principles of Economics.* New York: Macmillan and Co.

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

Mitchell, B.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.

Rosenthal, Stuart S, & Strange, William C. 2001. The Determinants of Agglomeration. *Journal of Urban Economics* , **50**(2), 191–229.

Shapiro, Jesse M. 2006. Smart Cities: Quality of Life, Productivity, and the Growth Effects of Human Capital. *The Review of Economics and Statistics*, **88**(2), pp. 324–335.

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.

Woods, R. 1997 (March). *Causes of Death in England and Wales, 1851-60 to 1891-1900 : The Decennial Supplements*. [computer file].

Woods, Robert. 2000. *The Demography of Victorian England and Wales*. Cambridge, UK: Cambridge University Press.

A Appendix

A.1 Empirical setting appendix

A.1.1 Further details on the empirical setting

Figure 2: English cities included in the study

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 more detail in the online appendix to Hanlon & Miscio (2014). At the end, I discuss the construction of additional control variables using the Census of Production data.

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 3 shows an example of what these data look like for the iron and steel industries.

Figure 3: 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 Mortality data

The mortality data used in this study come from the reports prepared by the Registrar General's office and digitized by Woods (1997).[60] The data were collected by an extensive system aimed at registering every birth, marriage, and death in England and Wales. Of the data collected by the Registrar's office, those on mortality are considered to be the most accurate and comprehensive, the "shining star of the Victorian civil registration" (Woods (2000)). For every death, registration with the local official (the "Registrar") was required within five days before the body could be legally disposed of. The Registrar was required to document the gender, age, and occupation of the deceased, together with the cause of death. The Registrar General's office put a substantial amount of effort into improving the registration of causes of death in the 1840s. This included sending circulars to all registrars and medical professionals, constructing a standardized set of disease nosologies, and providing registrars and medical professionals with standardized blank cause-of-death certificates. These efforts paid off with more accurate data in the 1850s and beyond. Thus, while there is surely some measurement error in the cause-of-death reporting, the error rates are not likely to be too large, particularly in the broad cause-of-death categories used in

[60]These data were obtained through the UK Data Archive.

this analysis.

Of the causes of death listed in the data, the most important for the purposes of this study are respiratory diseases. The disease category, which is the most closely associated with the impacts of airborne pollution, includes bronchitis, asthma, pneumonia, influenza, and other related diseases.

The mortality data are reported at the level of districts, which do not correspond perfectly to the borders of the cities used in the main analysis, though in general they are reasonably close. In the smallest cities, the city borders may be entirely within the district borders, while for the larger cities, such as Manchester and Birmingham, the borders of the city often exceed the borders of the district. London is an exception, where I can aggregate districts within the city to match the city borders. However, given that industrial pollution does not respect city or district borders, the pollution levels for the city are likely to be a good approximation of the pollution levels experienced by the residents of the district corresponding to the city center. Thus, issues in matching districts to cities should not be a major source of error in this analysis.

A.2.3 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.2.4 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 impact 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 (Forthcoming) 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. These are admittedly rough estimates.

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.

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.

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 (since British census data do not simultaneously measure occupation and industry until much later). See Hanlon & Miscio (2014) and the online appendix to that paper for more details.

A.2.5 Industry coal use intensity data

Table 7: Industry coal use per worker and employment in 1851

Industry	Coal/worker	Workers in 1851
Earthenware, bricks, etc.	48.9	135,214
Metal and engine manufacturing	43.7	894,159
Chemical and drug manufacturing	40.1	61,442
Mining related	28.9	653,359
Oil, soap, etc. production	20.7	54,751
Brewing and beverage production	19.4	100,821
Leather, hair goods production	12.1	27,146
Food processing	12.0	220,860
Textile production	10.1	1,066,735
Paper and publishing	9.7	226,894
Shipbuilding	6.1	169,770
Wood furniture, etc., production	5.4	114,014
Vehicle production	2.6	53,902
Instruments, jewelry, etc.	2.0	43,296
Apparel	1.6	243,968
Tobacco products	1.1	35,258
Employment-weighted average:	21.6	

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 city-industry database constructed from the Census of Population Occupation reports.

A.2.6 City coal use intensity data

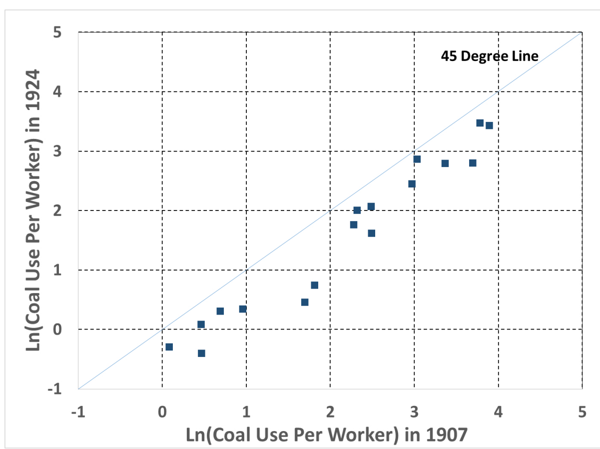
Table 8: Industrial coal use per private-sector worker for analysis cities (tons/year)

City	1851	1861	1871	1881	1891	1901	1911	Avg.	Growth
BATH	1.9	2.3	2.7	2.9	2.9	2.7	2.9	2.6	0.40
BRIGHTON	2.2	2.5	3.2	3.4	3.2	3.1	3.0	2.9	0.34
NORTHAMPTON	2.5	3.0	3.7	3.1	2.9	2.9	3.0	3.0	0.21
PORTSMOUTH	2.8	3.7	4.2	5.2	4.8	4.9	4.7	4.3	0.54
LIVERPOOL	2.8	3.5	4.3	4.6	4.2	4.1	4.1	3.9	0.40
LONDON	2.8	3.4	3.9	4.0	4.0	3.8	3.5	3.6	0.29
LEICESTER	2.9	3.9	4.7	4.1	3.6	4.3	5.1	4.1	0.43
SOUTHAMPTON	3.1	3.4	4.4	4.9	4.0	3.2	3.2	3.7	0.21
BRISTOL	3.3	4.2	4.9	4.8	4.6	4.8	4.9	4.5	0.36
HULL	3.3	4.6	6.4	6.0	5.9	5.8	6.3	5.5	0.63
NORWICH	3.5	4.1	4.9	5.2	4.7	4.4	4.2	4.4	0.28
NOTTINGHAM	3.9	5.0	6.2	7.6	7.4	7.1	7.0	6.3	0.63
IPSWICH	4.0	4.9	6.0	6.3	5.9	6.0	6.9	5.7	0.42
HUDDERSFIELD	4.7	5.7	7.1	7.7	7.8	7.4	7.3	6.8	0.45
BLACKBURN	5.1	6.8	7.8	8.5	8.3	7.8	7.9	7.5	0.45
MANCHESTER	5.2	6.5	7.1	7.4	7.5	7.1	6.9	6.8	0.32
SUNDERLAND	5.2	6.5	9.3	9.1	8.3	8.3	8.2	7.8	0.50
PRESTON	5.3	6.8	7.9	8.4	7.8	7.2	7.1	7.2	0.37
HALIFAX	5.5	6.7	8.6	9.5	9.7	9.3	9.6	8.4	0.54
STOCKPORT	5.5	6.6	7.3	7.3	6.4	5.8	6.6	6.5	0.19
SOUTH_SHIELDS	5.7	5.7	8.2	8.0	7.6	7.9	8.5	7.4	0.29
DERBY	5.8	7.2	9.2	10.4	9.7	8.7	8.6	8.5	0.46
NEWCASTLE	5.9	7.8	9.3	8.4	9.2	8.8	8.6	8.3	0.41
BRADFORD	6.0	7.3	8.1	8.4	7.8	7.3	7.1	7.4	0.23
LEEDS	6.4	8.7	10.2	10.1	9.2	8.6	8.7	8.8	0.39
BOLTON	6.5	8.9	9.8	10.5	10.4	9.7	9.6	9.4	0.43
OLDHAM	6.8	9.4	10.0	11.3	11.9	11.2	10.8	10.2	0.51
BIRMINGHAM	9.7	11.7	13.0	14.1	13.7	12.4	11.6	12.3	0.27
GATESHEAD	10.4	13.1	15.3	14.4	13.4	11.8	11.2	12.8	0.24
WOLVERHAMPTON	11.5	13.4	16.3	15.9	14.3	12.2	11.7	13.6	0.18
SHEFFIELD	12.5	14.9	17.8	17.3	16.6	15.3	15.5	15.7	0.26
Average	5.2	6.5	7.8	8.0	7.7	7.2	7.2		
Std. Dev.	2.6	3.2	3.7	3.7	3.5	3.1	3.0		

A.2.7 Analyzing the change in relative industry coal intensity over time

Figure 4 provides a scatterplot of industry coal use per worker for each industry in 1907 and 1924 as well as corresponding regression results. 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. This 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 for most of the 1851-1911 period burning coal was the dominant energy source for industries and there were few alternatives.

Figure 4: 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

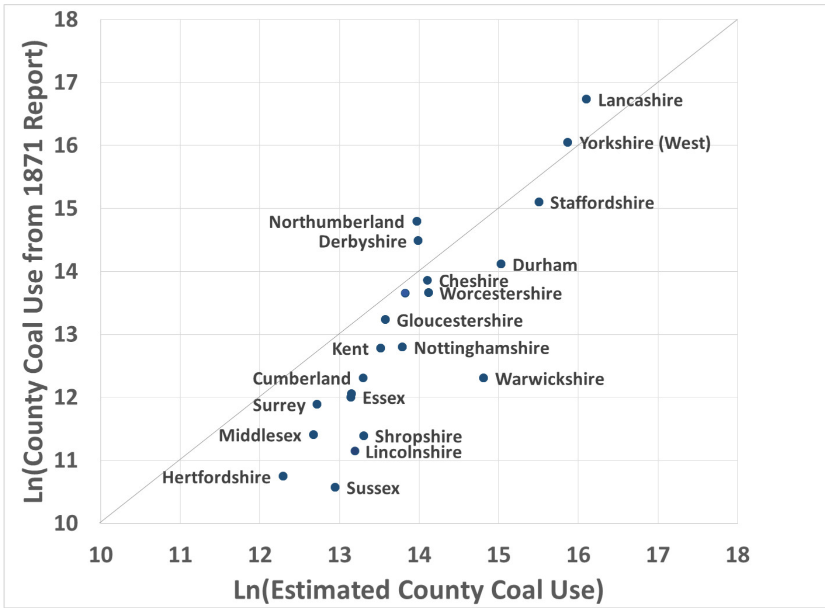
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

A.2.8 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 House of Commons, *Report of the Commissioners Appointed to Inquire into the Several Matters Relating to Coal in the United Kingdom*. 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 5 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 5: Comparing county industrial coal use in 1871



A.3 Theory appendix

In this appendix I consider some simple extensions 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. 7 in the main text:

$$\frac{L_{ict}}{L_{ict-1}} = \left(\frac{L_{it}}{L_{it-1}}\right)\left(\frac{L_{ct}}{L_{ct-1}}\right)^{\frac{-(1-\gamma)(1-\beta_i-\iota_i)\lambda}{1-\alpha_i-\beta_i-\iota_i}}\left(\frac{CP_{ct}}{CP_{ct-1}}\right)^{\frac{-\psi(1-\beta_i-\iota_i)-\nu}{1-\alpha_i-\beta_i-\iota_i}} \left(\frac{P_t}{P_{t-1}}\right)^{\frac{-(1-\beta_i-\iota_i)\gamma}{1-\alpha_i-\beta_i-\iota_i}}\left(\frac{v_t^*}{v_{t-1}^*}\right)^{\frac{-(1-\beta_i-\iota_i)}{1-\alpha_i-\beta_i-\iota_i}}\Omega_{it}^{-1}. \tag{13}$$

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 does not change as a result of the inclusion of capital into the model, so that the denominator of the exponent 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 grown 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:

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

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:

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

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

Variable	Mean	Std. Dev.	Min.	Max.
$\Delta Ln(L_{ict})$	0.437	0.492	-5.032	3.689
$\Delta Ln(PrEMP_{ict})$	0.365	0.253	-0.151	1.251
$\Delta Ln(PrCityEMP)$	0.27	0.056	0.084	0.479
$\Delta Ln(PredCoal)$	0.367	0.178	0.144	0.848
Ln(City Patenting)	4.311	1.508	0	8.875
City Air-frost Days	39.625	9.94	22.7	56
City Rainfall	0.805	0.19	0.557	1.294
N =		4017		

Table 10: Summary statistics for analysis of manufacturing industries only

Variable	Mean	Std. Dev.	Min.	Max.
$\Delta Ln(L_{ict})$	0.387	0.5	-1.762	3.689
$\Delta Ln(PrEMP_{ict})$	0.322	0.24	-0.151	1.251
$\Delta Ln(PrCityEMP)$	0.25	0.077	0.035	0.506
$\Delta Ln(PredCoal)$	0.366	0.185	0.108	0.852
Ln(City Patenting)	4.311	1.508	0	8.875
City Air-frost Days	39.625	9.94	22.7	56
City Rainfall	0.805	0.19	0.557	1.294
N =		4017		

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	0.918
Δ Emp., All Workers	0.321	0.18	-0.128	0.915
Δ Total Population	0.334	0.179	-0.039	0.915
$\Delta Ln(PrCityEMP)$	0.269	0.055	0.105	0.413
$\Delta Ln(PredCoal)$	0.367	0.178	0.167	0.835
N		155		

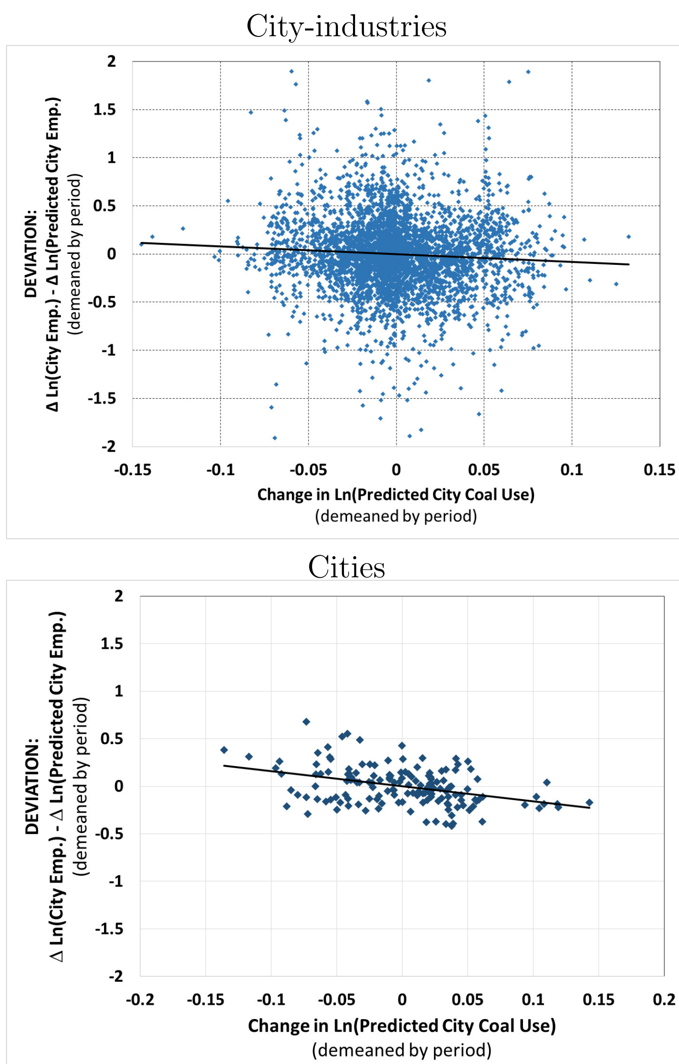
Table 12: Summary statistics for the mortality analysis

Variable	Mean	Std. Dev.	Min.	Max.
Total MR - Under 5	75.932	13.666	43.382	109.351
Total MR - 5-14	5.451	1.594	2.518	9.024
Total MR - 15-24	6.04	1.526	3.157	9.709
Total MR - 25-34	8.713	1.659	5.058	13.931
Total MR - 35-44	12.887	2.121	8.446	20.684
Total MR - 45-54	19.337	3.174	13.409	31.284
Total MR - 55-64	35.617	5.418	25.534	50.63
Total MR - 65-74	73.430	9.202	53.718	94.699
Respiratory MR - 5-14	0.489	0.145	0.172	0.950
Respiratory MR - 15-24	0.404	0.126	0.16	0.725
Respiratory MR - 25-34	0.851	0.28	0.332	1.631
Respiratory MR - 35-44	1.839	0.678	0.733	3.738
Respiratory MR - 45-54	3.947	1.569	1.427	8.557
Respiratory MR - 55-64	9.291	3.277	4.158	18.356
Respiratory MR - 65-74	19.507	5.666	8.83	32.117
Ln(Coal)	12.542	1.108	10.31	15.697
Ln (Density)	2.196	1.565	-0.768	5.192
N		155		

A.4.2 Scatterplots of employment growth against coal use

Figure 6 plots the difference between actual employment growth and predicted employment growth against the predicted growth in city-level coal use using all available 20-year differences, demeaned by year, for the 31 cities available in the database. The top panel shows this relationship for employment in each city-industry, which matches the approach taken in the main analysis. While there is a lot of noise in city-industry growth rates, I observe that city-industry growth consistently underperformed relative to national industry growth in cities where we predict a larger change in local industrial coal use. The bottom panel shows that a similar relationship emerges when employment is aggregated to the city level. Here the pattern is much clearer because aggregating to the city level has eliminated the noise in the city-industry level data.

Figure 6: Deviation vs. predicted change in city coal use



Top panel: The y-axis is the difference between actual city-industry employment growth over each two-decade period in city c and industry i and predicted growth in that city-industry based on employment growth in industry i in all other cities. The x-axis is the predicted change in city-level industrial coal use over the two-decade 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. **Bottom panel:** The y-axis is the difference between actual city employment growth over each two-decade period in city c, summing across all industries, and predicted growth in that city-industry based on 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.

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. Column 2 adds in a control based on the level of patenting in the city in the 1850s, which can reflect the innovative potential of the local economy.[61] In Column 3, 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 4, 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. Finally, Column 5 presents results with London excluded from the data. In general, the signs on the city controls are as expected. However, the inclusion of these variables does not change the baseline results.

[61] 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				
	(1)	(2)	(3)	(4)	(5)
$\Delta Ln(PrEMP_{ict})$	1.022***	1.020***	1.028***	1.019***	1.034***
	(0.0298)	(0.0305)	(0.0304)	(0.0299)	(0.0299)
$\Delta Ln(PrCityEMP)$	-0.459	-0.512	-0.385	-1.188**	-1.170**
	(0.562)	(0.581)	(0.583)	(0.513)	(0.520)
$\Delta Ln(PredCoal)$	-1.742***	-1.745***	-1.503**	-1.298**	-1.315**
	(0.661)	(0.662)	(0.631)	(0.560)	(0.565)
City Air-frost Days	-0.00281*	-0.00282*	-0.00299*	-0.00267*	-0.00269*
	(0.00152)	(0.00152)	(0.00155)	(0.00149)	(0.00149)
City Rainfall	-0.0950	-0.102	-0.0999	-0.181*	-0.185*
	(0.101)	(0.105)	(0.0971)	(0.104)	(0.109)
Ln(City Patenting)		0.00312	0.00161	0.0221*	0.0224*
		(0.00895)	(0.00834)	(0.0118)	(0.0121)
Ln(Initial city pop.)				-0.154***	-0.151***
				(0.0282)	(0.0322)
Ln(Initial coal use)				0.115***	0.113***
				(0.0285)	(0.0288)
Border Chg. Flag			0.118***		
			(0.0281)		
Constant	1.539***	1.549***	1.345***	1.658***	1.642***
	(0.378)	(0.385)	(0.374)	(0.348)	(0.359)
Time effects	Yes	Yes	Yes	Yes	Yes
Dropping London					Yes
Observations	4,017	4,017	4,017	4,017	3,887
R-squared	0.318	0.318	0.329	0.333	0.329

*** 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.

Table 14 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 14: City-industry regression results with lagged changes in coal use

	DV: Δ Log of city-industry employment			
	All industries		**Manufacturing only**	
	(1)	(2)	(3)	(4)
$\Delta Ln(PrEMP_{ict})$	1.015***	1.019***	1.010***	1.010***
	(0.0394)	(0.0393)	(0.0550)	(0.0548)
$\Delta Ln(PrCityEMP)$	-1.586**	-0.363	-2.018***	-1.333***
	(0.645)	(0.845)	(0.368)	(0.477)
$\Delta Ln(PredCoal)$		-2.414**		-1.328
		(0.978)		(0.880)
Lagged $\Delta Ln(PredCoal)$	-0.0494	0.619	0.414	0.651
	(0.395)	(0.454)	(0.356)	(0.452)
Constant	0.491***	0.396**	0.248	0.287
	(0.173)	(0.179)	(0.189)	(0.188)
Time effects	Yes	Yes	Yes	Yes
Observations	2,405	2,405	1,382	1,382
R-squared	0.291	0.302	0.314	0.316

*** 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. Lagged predicted coal use is the predicted change in city level coal use from the two previous decades.

Table 15 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, and the ratio of labor costs to revenue. 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.2.4.

Table 15 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 15: Results including controls based on other industry characteristics

	DV: Δ Log of city-industry employment				
	(1)	(2)	(3)	(4)	(5)
$\Delta Ln(L_{i-ct})$	1.024***	1.030***	1.029***	1.022***	1.026***
	(0.0477)	(0.0470)	(0.0471)	(0.0475)	(0.0486)
$\Delta Ln(PrCityEMP)$	-0.456	-0.341	-0.432	-0.248	-0.133
	(0.424)	(0.485)	(0.449)	(0.416)	(0.588)
$\Delta Ln(PredCoal)$	-1.387***	-1.355***	-1.270**	-1.524***	-1.447***
	(0.483)	(0.501)	(0.527)	(0.466)	(0.515)
$\Delta Ln(SalariedWkr.Shr.)$	-0.624				0.254
	(0.972)				(1.463)
$\Delta Ln(Avg.FirmSize)$		0.204			0.413
		(0.403)			(0.609)
$\Delta Ln(ExportsShr.)$			0.0711		-0.811
			(0.877)		(1.212)
$\Delta Ln(LaborCostShr.)$				8.163*	9.555**
				(4.328)	(4.654)
Constant	1.127***	1.073***	1.033***	1.161***	1.089***
	(0.270)	(0.252)	(0.269)	(0.241)	(0.268)
Time effects	Yes	Yes	Yes	Yes	Yes
Observations	2,312	2,312	2,312	2,312	2,312
R-squared	0.286	0.286	0.286	0.287	0.288

*** 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.2.4.

Another potential concern in my main analysis regressions 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.2.4.

Results obtained while including these controls are shown in Table 16. 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. Of the controls, only the IO measures show evidence of positive effects, though these are not statistically significant. The DEM and OCC variables both show negative effects and appear to absorb much of the impact of the city-size variable. This suggests that much of the city-size effect may be transmitted through competition in local labor markets.

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 included controls for the dynamic agglomeration forces documented in that study. The main results are also robust to the inclusion of these controls.

Table 16: City-industry regression results with industry connections controls

	DV: Δ Log of city-industry employment			
	(1)	(2)	(3)	(4)
$\Delta Ln(PrEMP_{ict})$	0.968***	0.969***	1.017***	0.971***
	(0.0490)	(0.0490)	(0.0490)	(0.0491)
$\Delta Ln(PrCityEMP)$	-0.400	-0.404	0.233	0.0278
	(0.437)	(0.439)	(0.448)	(0.550)
$\Delta Ln(PredCoal)$	-1.205***	-1.182**	-0.783*	-1.144**
	(0.464)	(0.467)	(0.459)	(0.466)
$\Delta Ln(IOin)$	0.0676			
	(0.0663)			
$\Delta Ln(IOout)$		0.00277		
		(0.0557)		
$\Delta Ln(DEM)$			-0.960***	
			(0.173)	
$\Delta Ln(OCC)$				-0.478
				(0.337)
Constant	0.970***	0.972***	0.842***	0.967***
	(0.241)	(0.244)	(0.236)	(0.241)
Time effects	Yes	Yes	Yes	Yes
Observations	2,004	2,004	2,004	2,004
R-squared	0.238	0.237	0.253	0.238

*** 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.

A.4.4 Heterogeneous effects regressions

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 17 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. 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 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 17: Heterogeneous effects in more labor intensive industries

	DV: Δ Log of city-industry employment			
	(1)	(2)	(3)	(4)
$\Delta Ln(PrEMP_{ict})$	1.011***	1.001***	0.948***	0.955***
	(0.0470)	(0.0478)	(0.0469)	(0.0497)
$\Delta Ln(PrCityEMP)$	-0.550	-1.609***	-0.585	-1.559***
	(0.465)	(0.499)	(0.473)	(0.507)
$\Delta Ln(PrCityEMP)$*Labor Shr.	0.611	0.570	0.669	0.577
	(1.262)	(1.268)	(1.271)	(1.279)
$\Delta Ln(PredCoal)$	-1.079**	-0.722	-1.391***	-1.058**
	(0.491)	(0.467)	(0.531)	(0.506)
$\Delta Ln(PredCoal)$*Labor Shr.	-0.968*	-0.989*	-1.158**	-1.134**
	(0.513)	(0.516)	(0.515)	(0.520)
Industry Labor Cost Shr.	0.361	0.380	0.433	0.444
	(0.350)	(0.353)	(0.352)	(0.355)
$\Delta Ln(PrCityEMP)$*Coal Use			0.0120	0.0129*
			(0.00739)	(0.00737)
$\Delta Ln(PredCoal)$*Coal Use			0.00450*	0.00363
			(0.00265)	(0.00267)
Industry Coal Per Worker			-0.00717***	-0.00647***
			(0.00227)	(0.00227)
Time effects	Yes	Yes	Yes	Yes
Additional controls		Yes		Yes
Observations	2,312	2,312	2,312	2,312
R-squared	0.288	0.305	0.295	0.310

*** 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.4.5 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,

$$\Delta \ln(L_{ict}) = a_0 + a_1 \Delta \ln(PrEMP_{ict}) + a_2 \Delta \ln(PrCityEMP_{ct}) + a_3 \Delta \ln(CoalPW_{ct}) + \xi_t + e_{ict} \,,$$

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

$$\Delta \ln(CoalPW_{ct}) = b_0 + b_1 \ln(PrCityEMP_{ct}) + b_2 \Delta \ln(PredCoal_{ct}) + \xi_t + \epsilon_{ict} \,.$$

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 the coal use.

Table 18 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 results including the $PrEMP_{ict}$ term, as in the top panel of Table 1, while Columns 3-4 present results that instead include a full set of industry-time effects, as in the bottom panel of Table 1. These results 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, but in general the differences are not statistically significant. 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.

First stage regression results are presented in the bottom panel of Table 18. These show that the predicted change in local industrial coal use is a sufficiently strong predictor of the observed change in the coal use intensity of local industries when controlling for the overall change in local employment.

Table 18: IV regression results for manufacturing industries over two-decade differences

	DV: Δ Log of city-industry employment			
	(1)	(2)	(3)	(4)
$\Delta \ln(CoalPW_{ct})$	-0.913**	-0.653	-1.526**	-1.182*
	(0.448)	(0.397)	(0.733)	(0.680)
$\Delta Ln(PrEMP_{ict})$	1.025***	1.013***		
	(0.0568)	(0.0568)		
$\Delta Ln(PredCityEMP)$	-1.728***	-2.529***	-1.888***	-2.795***
	(0.300)	(0.493)	(0.343)	(0.562)
Time effects	Yes	Yes		
Ind-time effects			Yes	Yes
Other controls		Yes		Yes
	First-stage results: DV is $\Delta Ln(CoalPW)$			
$\Delta Ln(PredCoal)$	1.362***	1.363***	1.372***	1.278***
	(0.301)	(0.267)	(0.420)	(0.390)
$\Delta Ln(PrEMP_{ict})$	-.004	-0.010		
	(0.006)	(0.006)		
$\Delta Ln(PredCityEMP)$	-1.406***	-1.571***	-1.413***	-1.498***
	(0.290)	(0.294)	(0.377)	(0.397)
Time effects	Yes	Yes		
Ind-time effects			Yes	Yes
Other controls		Yes		Yes
F-stat	20.50	25.99	10.67	10.73
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.

Table 19 explores some additional IV results that also include the additional controls based on industry characteristics shown in Table 15. In each of these I use changes in the predicted values of each of these controls as instruments for the changes in the values based on actual city-industry growth rates, while also instrumenting for the change in coal use. For space reasons, I omit the first stage results (available upon request – they look quite reasonable) and report only F-statistics. This table shows that the basic results are essentially unchanged when these additional controls are included.

Table 19: IV regression results including controls for other industry characteristics

	DV: Δ Log of city-industry employment			
$\Delta \ln(CoalPW_{ct})$	-0.924**	-0.798*	-0.714*	-1.073**
	(0.403)	(0.442)	(0.421)	(0.424)
$\Delta Ln(PrEMP_{ict})$	1.024***	1.025***	1.026***	1.018***
	(0.0586)	(0.0569)	(0.0574)	(0.0581)
$\Delta Ln(PredCityEMP)$	-1.753***	-1.688***	-1.594***	-1.841***
	(0.366)	(0.286)	(0.292)	(0.296)
$\Delta Ln(SalariedWkr.Shr.)$	-0.120			
	(1.432)			
$\Delta Ln(Avg.FirmSize)$		-0.147		
		(0.389)		
$\Delta Ln(ExportsShr.)$			-0.560	
			(0.771)	
$\Delta Ln(LaborCostShr.)$				6.446
				(5.151)
Constant	0.921***	0.839***	0.784***	1.007***
	(0.227)	(0.247)	(0.230)	(0.238)
Observations	2,312	2,312	2,312	2,312
R-squared	0.247	0.248	0.245	0.239
F-stat Coal F.S.	24.99	25.72	31.53	21.91
F-stat control F.S.	6.35	12.78	15.74	45.35

*** 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. F-stat indicates the Angrist-Pischke multivariate F-test of excluded instruments

A.4.6 Mortality regression results

This subsection presents further details related to the discussion of mortality in Section 5.3. First, Table 20 presents the full set of regression results corresponding to the coefficient estimates described in Figure 1. It is interesting to note here that density had a strong negative effect on mortality during this period. This does not appear to be an artifact of reverse causality, since a similar pattern is observed if I use lagged density as a control rather than density at the beginning of the decade. Rather, this result reflects the fact that, while dense areas were initially less healthy places to live, they also experienced more rapid mortality reductions in the 19th century (Haines (2001)). This was due largely to sanitary improvements, particularly the provision of clean water and sewerage. Larger cities had a natural advantage in implementing these technologies because they involved large fixed costs which could be spread over more consumers.

Table 20: Relationship between coal use and mortality

Age:	(1) Under 5	(2) 5 - 14	(3) 15 - 24	(4) 25 - 34	(5) 35 - 44	(6) 45 - 54	(7) 55 - 64	(8) 65 - 74
Ln(Coal)	14.37*	0.0720	0.559	-1.096	-1.333	-1.148	-0.0969	1.987
	(7.366)	(0.694)	(0.894)	(1.084)	(1.321)	(1.890)	(2.260)	(5.631)
Lag Ln(Coal)	-2.794	0.0721	-0.432	-0.0771	0.726	0.407	1.899	4.354**
	(3.540)	(0.305)	(0.398)	(0.428)	(0.600)	(0.642)	(1.417)	(2.035)
Ln (Density)	-11.50	0.109	-1.155	1.267	1.885	2.220	1.159	-6.010
	(9.613)	(0.961)	(0.955)	(1.342)	(1.828)	(2.627)	(3.150)	(6.547)
Constant	-54.37	1.390	5.554	18.66*	14.82	22.84	9.961	7.026
	(78.44)	(7.382)	(10.90)	(10.44)	(12.64)	(18.56)	(26.50)	(57.15)
Decade effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	124	124	124	124	124	124	124	124
R-squared	0.803	0.917	0.866	0.857	0.608	0.268	0.124	0.183
Number of loc_code	31	31	31	31	31	31	31	31

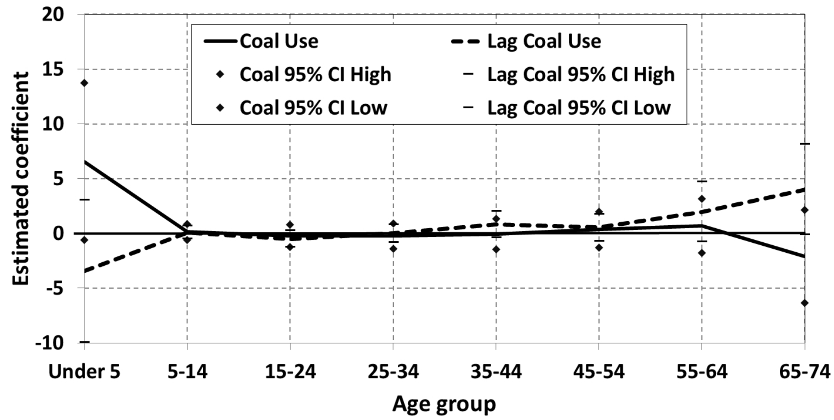
*** p<0.01, ** p<0.05, * p<0. Standard errors, in parenthesis, are clustered by location to allow serial correlation in mortality rates over time.

Next, Figure 7 presents a set of additional regression results, similar to those shown in Figure 1, but using alternative approaches to controlling for density. The top panel presents results in which I do not include a control for density. These show smaller but qualitatively similar pollution effects. It is not surprising that the pollution effects are smaller since highly polluted areas also tended to be more dense and dense areas were experiences reductions in mortality during this period. The middle panel shows results when I include only lagged density, rather that current density, as a control. The use of lagged density in place of contemporaneous density will reduce concerns that reverse causality may be affecting the results. These results are similar to the baseline estimates except that they suggest that both current and lagged coal use influenced mortality among the young. The bottom panel shows results when both contemporaneous and lagged density are included as controls. These look very similar to the baseline results. In addition to the results shown here, I have also calculated results while including the log of coal use one decade ahead as an explanatory variable. Those results show that the lead of coal use is not related to mortality in a statistically significant way for any age group.

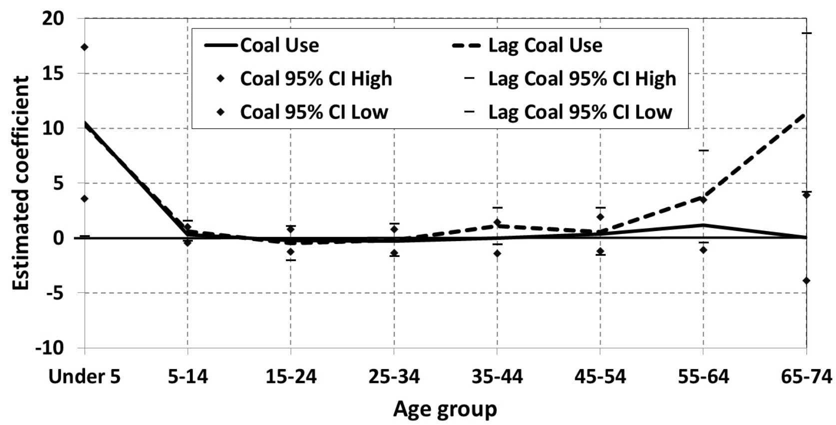
It is also possible to look at how coal use affected mortality in particular disease categories. Figure 8 presents results of this nature, focusing on diseases of the respiratory system. This disease category, which includes deaths due to bronchitis, asthma, pneumonia, and other similar diseases, is the category most closely related to the expected effects of airborne pollution. It was also the largest single cause-of-death category in 19th century England, and it was increasing during most of the study period. These results focus only on deaths among those over 5 years old because cause of death date for small children is often considered to be less reliable.

Figure 7: Results for respiratory mortality

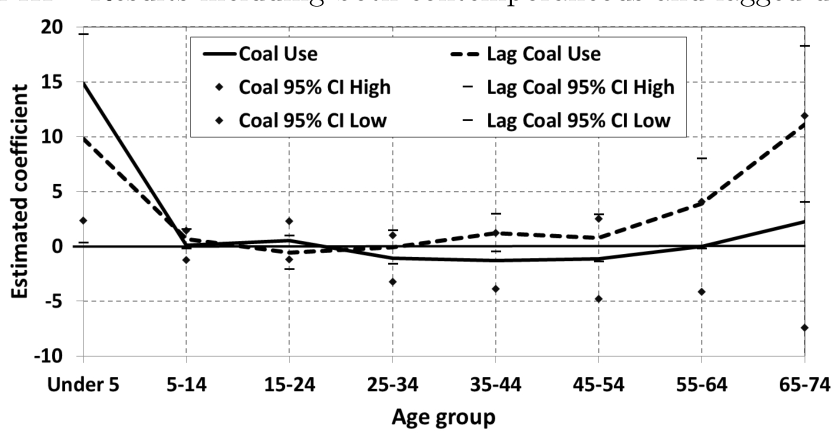
Panel I – Results without including any density controls



Panel II – Results including only lagged density as a control



Panel III – Results including both contemporaneous and lagged density

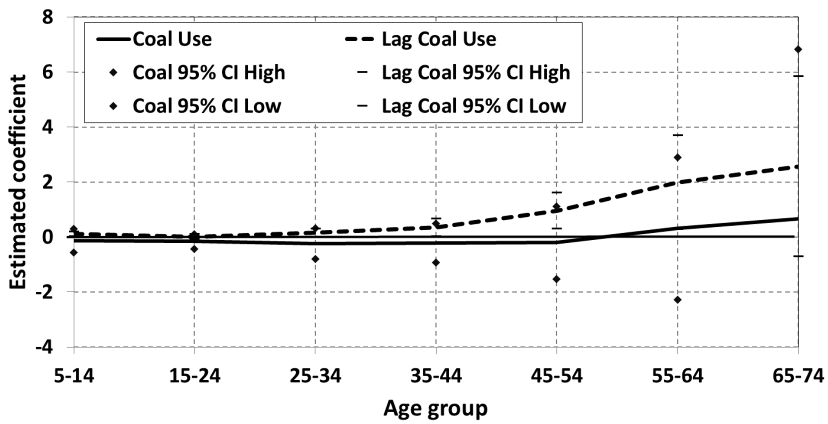


These figures show coefficients and 95% confidence intervals for the coal use and lagged coal use terms. The results are generated using the regression specification shown in Eq. 11 but with alternative approaches to controlling for population density. Standard errors are clustered by location to allow serial correlation.

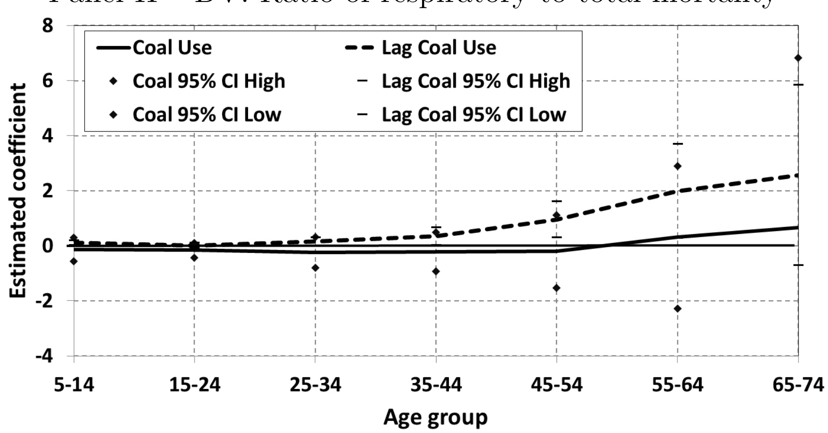
The results in the top panel of Figure 8 show that lagged coal use was associated with increased respiratory mortality for all age groups. These effects are statistically significant at the 95% level for most age categories. The bottom panel of Figure 8 shows that lagged coal use was also associated with an increase in the share of total mortality that was due to respiratory diseases. Because of the close association between respiratory diseases and airborne pollution, these findings provide us with additional confidence in the mortality results.

Figure 8: Results for respiratory mortality

Panel I – DV: Mortality rate due to respiratory diseases



Panel II – DV: Ratio of respiratory to total mortality



These figures show coefficients and 95% confidence intervals for the coal use and lagged coal use terms. The results are generated using the regression specification shown in Eq. 11 but with alternative dependent variables. Standard errors are clustered by location to allow serial correlation.

A.4.7 Comparing mortality to city growth effects

For this exercise, I consider two hypothetical cities of 100,000 people, shown in the first row of Table 21. The first hypothetical city, in Column 1, grows at the rate observed aggregating all cities in my data (0.334 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 0.178. Multiplying this by the estimated impact of rising industrial coal use on total city population in Table 2 (-1.017), the decadal growth in this city should be 0.153. 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 916 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. Using the results from Table 20 applied to the population distribution across ages from 1881, I estimate that a one log-point increase in local industrial coal use raises mortality by 2.25 deaths per thousand per year. Thus, an increase in log coal use of 0.178 is expected to generate 40 additional deaths over one year in a city of 100,000 people. Thus, rising mortality could explain about 4.6% of the reduction in overall population, which implies that the remaining 95.4% are attributable to changing migration decisions.

Table 21: 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:	0.334	0.153	
Compound annual growth:	1.0168	1.0077	
Population after one year	101,684	100,768	916
Estimated impact of coal use on mortality (per 1000):			2.25
Estimated additional deaths in coal intensive city:			40
Fraction of population growth diff. explained by mortality:			0.044

A.4.8 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 our estimates. In particular, abstracting from heterogeneity in the production function parameters, the relationship between the ψ and ν parameters is determined by,

$$\frac{-\psi(1-\beta)-\nu}{1-\alpha-\beta} = X\,, \tag{14}$$

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.

Before doing this, it is useful to gain some understanding of the relationships described by Eq. 14. In this equation, the impact of either the amenities effect (represented by ψ) or the productivity effect (represented by ν) on the overall relationship between coal use and city-industry employment growth (represented by X) will depend on the term $1-\alpha-\beta$, which appears in the denominator. This term is the cost share of industry-specific fixed city resources, and it determines how strongly changes in amenities or productivity driven by coal use impacts employment. This is because the importance of fixed local resources determines the extent to which production can be relocated to other cities. If fixed resources are important, then the denominator on the left-hand side of Eq. 14 will be large, and an equivalent change in either amenities or productivity will have a small impact on city-industry employment because it is difficult to reallocate employment across locations in response to these effects.

The term $1-\beta$ appears in the numerator of Eq. 14, where it multiplies ψ. Note that, holding fixed the cost share of industry-specific resources, variation in this term will reflect changes in the labor cost share. In particular, when the labor cost share is larger, $1-\beta$ will be larger, which implies that coal use will have a stronger impact on city-industry employment when employment is a more important part of the production function. There is no corresponding term multiplying the ν parameter, which reflects the fact that, as modeled here, the productivity effect of coal use is on total factor productivity, rather than labor productivity alone.

In Table 22, 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 -0.8 and -1.2. 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 much more important in generating the impact of coal use on city employment than the amenity channel.

Table 22: 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 -0.8	Productivity effect parameter (ν) using coal effect of -1.2
Lowest:				
0.0172	0.30	0.05	0.045	0.065
	0.30	0.10	0.085	0.125
	0.50	0.05	0.049	0.069
	0.50	0.10	0.089	0.129
	0.50	0.20	0.169	0.249
	0.70	0.10	0.092	0.132
	0.70	0.20	0.172	0.252
	0.70	0.40	0.332	0.492
Highest:				
0.0504	0.30	0.05	0.055	0.075
	0.30	0.10	0.095	0.135
	0.50	0.05	0.065	0.085
	0.50	0.10	0.105	0.145
	0.50	0.20	0.185	0.265
	0.70	0.10	0.115	0.155
	0.70	0.20	0.195	0.275
	0.70	0.40	0.355	0.515