Endogenous City Disamenities: Lessons from Industrial Pollution in 19th Century Britain*

W. Walker Hanlon
UCLA and NBER

June 10, 2015

Abstract

Growing industries create jobs and attract workers to cities, but they may also generate pollution, an endogenous disamenity that drives workers and firms away. Separating these positive and negative effects is a central challenge in understanding the role of industries and pollution in shaping city growth. To make progress on this issue, this paper introduces a model incorporating endogenous disamenities related to industrial pollution into a standard urban framework. I then propose a new theoretically-grounded estimation strategy that can be used to separate the positive direct effect of employment growth in polluting industries on overall city employment from the negative indirect effect that occurs through the pollution these industries generate. This strategy is applied to study the influence of industrial pollution on the long-run growth of British cities from 1851-1911, the first widespread episode of modern industrial pollution. My results suggest that coal-based pollution substantially impeded the growth of British cities during this period.

*I thank Leah Boustan, Karen Clay, Dora Costa, Neil Cummins, Roger Foquet, Erik Hurst, Matt Kahn, Ed Kung, Miren Lafourcade, Adriana Lleras-Muney, Werner Troesken and seminar participants at UC Davis, UCLA, the Urban Economic Association Annual Meeting, the Economic History Society Annual Meeting, and the Population Association of America Annual Meeting for their helpful comments. I thank Reed Douglas for his excellent research assistance. 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

From the mill towns of 19th century England to the cities of modern China, urbanization often goes hand-in-hand with pollution. Much of this pollution comes from industry, an unfortunate by-product of the engines that drive city employment growth. Pollution represents an endogenous disamenity that acts as a drag on urban growth, pushing residents away from productive areas and forcing firms to pay higher wages. As a result, policy makers face a trade-off between encouraging the growth of industry and addressing local pollution.

One of the central challenges in understanding industrial pollution is separating the positive direct effect of employment growth in polluting industries on cities from the negative effect generated by the pollution these industries produce. The main contribution of this study is to introduce a strategy that can be used to separate these positive and negative effects, and then apply it to study the impact of pollution on British cities in the 19th and early 20th century.

I begin by introducing a model that incorporates endogenous city disamenities generated by industrial pollution into a standard spatial equilibrium framework. In the model, city economies are composed of many industries which differ in their pollution intensity. Industrial pollution acts as an endogenous disamenity that reduces the utility of workers. In response to high levels of pollution, workers are less willing to move to a city, which forces wages upwards and increases the costs faced by firms. As a result, when productivity or demand shifts cause employment in a polluting industry in a city to grow, the impact on overall city employment growth can be partially or fully offset by the endogenous pollution disamenity.

Using this model, I derive a new estimation strategy that can be used to separate the positive direct impact of employment growth in polluting industries on overall city growth from the negative indirect impact of the pollution generated by these industries. To implement this strategy, I begin by constructing two variables, building on the approach from Bartik (1991). First, I construct a *predicted city employment growth* variable based on lagged city-industry employment multiplied by industry employment growth rates in all other cities. Second, I construct a *predicted city pollution* variable based on lagged city-industry employment multiplied by industry employment growth rates in all other cities and then multiplied by industry pollution intensity per worker. Both variables are then summed across industries to obtain city-level values. The difference between these variables is driven entirely by variation in the pollution intensity of industries and variation across cities in initial industrial composition.

I then run regressions in which the growth of actual city employment is explained by the growth in predicted city employment, the growth in city industrial pollution, and control variables. When both predicted employment and city industrial pollution are included in the regression, the predicted employment variable will capture the direct effect of industry employment growth while the predicted industrial pollution variable will capture the indirect effect of industry growth through the pollution disamenity. Thus, this approach allows me to separate the positive and negative impacts of polluting industries on city employment.

This estimation strategy is related to the approach suggested by Bartik (1991). In Bartik's study, and many other studies using a similar approach, the initial industrial composition of a location is interacted with national industry growth rates in order to generate an instrument for local industry employment in future periods. The approach I offer is different. It begins with the idea that the predicted city employment growth variable obtained using the Bartik approach should, on average, correctly predict actual city employment growth once other relevant factors – such as industrial pollution – are accounted for. Failing to account for these factors will cause a systematic deviation between actual city employment and the predicted employment obtained using Bartik's approach. My estimation strategy is based on the idea that the deviation from the Bartik prediction is meaningful and can be analyzed. In particular, this study focuses on the extent to which the deviation can be explained by endogenous city disamenities caused by industrial pollution. To my knowledge this is the first paper to propose this empirical approach.

My estimation strategy has several attractive features. One important characteristic is that it can be implemented in relatively data-sparse environments in which direct pollution measures are unavailable. Many of the high-pollution settings that we would like to study – such as historical cities, or cities in developing economies – are lacking in direct pollution measures. The approach offered here requires only panel data on city-industry employment and measures of the pollution intensity of industries. These data can often be obtained from standard administrative sources. Moreover, even if direct pollution measures are available, one may still want to implement an instrumental variables strategy based on the methodology introduced here in order to obtain exogenous variation in local pollution levels.

Another advantage of this estimation strategy is that it comes with a natural test of the regression specification. If the specification is fully accounting for other factors affecting city employment growth, then the predicted employment growth variable should correctly predict actual employment growth. This implies that, once other factors affecting city growth are accounted for, the coefficient on the predicted employment growth term should be close to one. This is a helpful feature because it allows me to evaluate the quality of alternative regression specifications.

One other useful feature of this estimation strategy is that it is not subject to bias from time-invariant unobserved local features, a common concern when studying city amenities. For example, when studying cities polluted by coal-using industries, we may be worried that easier access to coal reserves also affects the residential price of coal, which in turn impacts the cost of living. By using panel data and focusing on city growth, time-invariant features such as access to nearby coal deposits will be differenced out. This eliminates many of the potential confounding factors that can affect studies of city amenities.

I apply this estimation strategy to study the impact of pollution on city growth during the first widespread case of modern urban industrial pollution: British cities in the late 19th and early 20th century. During this period, industrialization in the form of coal-powered factories drove economic growth while clouding city skies with thick dark smoke. Contemporary sources suggest that pollution reached extreme levels in cities containing the most coal-intensive industries, such as iron and steel production. In a companion paper to this study, Hanlon (2015), I examine the impact of industrial pollution on health during this period. The results of that study show that industrial pollution substantially increased mortality in more polluted areas and that this impact grew as coal use expanded in the second half of the 19th century. This study builds upon those results to consider the impact of pollution on city growth.

While this setting is of particular historical interest, it also offers several features that make it a clean environment for tracing out the impact of industrial pollution on city growth, an issue of broader concern. For example, the British economy during this period was largely free of regulatory constraints due to the strong laissez faire ideology that dominated British policy. There were no zoning laws, few local land use or environmental regulations, essentially no place-based policies, and federal taxes and transfers were limited. When present, these factors complicate the analysis of the relationship between pollution and city growth. The lack of government interference makes the setting I consider a good proving ground for validating the new estimation strategy that I introduce. A second advantage of using a historical setting is that I am able to examine impacts over many decades. This matters because it may take many years for city size to respond to changes in local amenity levels. A third useful feature is that the industrial pollution I consider was clearly visible to city residents. Also, this setting was characterized by fairly high levels of mobility, particularly mobility of rural residents into cities. These features mean that residents were able to observe pollution levels and were mobile enough to respond by adjusting their migration decisions.

My results suggest that coal use had a substantial impact on city employment and population growth during the study period. Based on my preferred regression results, if industrial pollution growth in a city was one standard deviation (s.d.) above the national average, the annual growth rate of city employment was reduced by roughly 2 percentage points. For comparison, the average growth rate of employment observed across all cities and years in my data was 1.5 percent. Thus, high rates of growth in industrial pollution substantially reduced overall city growth during this period. However, because high rates of pollution growth were accompanied by high rates of employment growth, it was difficult for contemporaries to discern the substantial negative effects of industrial pollution. This highlights the importance of having an estimation strategy that can separate out the positive and negative effects of growth in polluting industries.

Drawing on estimates of the impact of pollution on mortality from Hanlon (2015), I also provide a back-of-the-envelope calculation of the impact of pollution on city size that occurs through increased mortality in more polluted cities. This analysis suggests that the increased mortality associated with industrial pollution explains roughly 15% of the impact of industrial pollution on overall city population. This implies that the majority of the impact of industrial pollution on city size can be attributed to the decisions made by migrants in response to the disamenity generated by industrial pollution and the commensurate negative health effects.

The next section briefly reviews the related literature. The model is introduced in Section 3, followed by the empirical setting in Section 4. Section 5 describes the data, presents the strategy used to measure industrial pollution, and verifies this measure using information on the quality-of-life in cities. The main analysis is in Section 6, while Section 7 concludes.

2 Related literature

The main focus of this study, and a central challenge in the pollution literature, is separating the economic and amenity (including health) effects of polluting industries. One common approach to addressing this issue is to argue that the relevant spatial scale for pollution impacts differs from the relevant scale for economic impacts (see, e.g., Banzhaf & Walsh (2008), Bayer *et al.* (2009), and Currie *et al.* (2014)).[1] Differences in spatial scale can then be exploited using a differences-in-differences approach. An second possible approach is to consider changes in local environmental conditions that are driven by historical rather than current industrial activity, such as the cleanup of Superfund sites (Cameron & McConnaha (2006), Greenstone & Gallagher (2008)). One of the main contributions of this study is to add a third strategy that can be used by empirical researchers interested in separating the impact of industrial pollution from the effect of job creation in polluting industries. This new strategy can complement existing approaches by allowing us to separate the economic and amenity/health costs of pollution in a wider variety of settings than was previously possible.

The typical approach to assessing the costs of pollution (or the value of pollution regulations) is to look at the impact on housing prices using hedonic regressions, following Roback (1982).[2] A seminal example is Chay & Greenstone (2005), which uses the impact of plausibly exogenous shifts in local air pollution caused by the Clean

[1] Banzhaf & Walsh (2008) argue that, "The relevant spatial scale for considering economic conditions and labor market opportunities in a locational choice is likely much wider than the scale of environmental amenities." Currie *et al.* (2014) take a similar approach; they provide evidence that the geographic scale for toxic emission effects are less than one mile, and then look at how the opening and closing of toxic plants affects housing values within that range compared to to housing values just outside of it. Mastromonaco (Forthcoming) takes a similar approach, also with the EPA's Toxic Release Inventory. In contrast, Bayer *et al.* (2009) look at the effects of pollution generated by distant sources (at least 80km away) and argue that at this large geographic scale only the pollution effect, and not the economic impacts, are present.

[2] Smith & Huang (1995) provides a review of some of the older hedonic literature. In addition to this literature, there is a much larger set of studies looking at the health and mortality effects of pollution. This literature is too large to review here.

Air Act on housing prices to estimate the value of reducing local air pollution levels.[3] Unlike the hedonic literature, this study will focus on the impact of pollution on employment. This approach is less common, but has been used in studies by Banzhaf & Walsh (2008) and Bayer *et al.* (2009). In theory, housing prices and population are linked, so studying employment or population as an outcome is complementary to studying housing prices. The advantage of studying city population is that these data are available on a consistent basis in a wider set of circumstances, including in panels over long periods of time, such as the one I consider. The disadvantage, however, is that studying employment or population effects does not allow me to estimate the monetary cost of pollution.

This study also contributes to a growing literature on endogenous amenities in cities. Existing work in this area, including Rauch (1993), Moretti (2004), Shapiro (2006), and Diamond (2012), has largely focused on the impact of endogenous amenities related to the composition of city residents. This study applies a similar logic to endogenous disamenities related to the composition of city industries.

Methodologically, this study follows existing work by Kahn (1999) and Chay & Greenstone (2003) which use economic fluctuations that affect polluting industries to assess the health impacts those industries produce. The main difference between these papers and my study is that I look at city growth and attempt to separate the economic and amenity effects of polluting industries.

Finally, this study builds on existing economic history studies looking at the impact of air pollution in historical settings (e.g., Troesken & Clay (2011), Barreca *et al.* (2014), Clay *et al.* (2014), Hanlon (2015)). The main difference between this study and previous work in this vein is that I focus on the impact of pollution on the urban system, while most existing work considers the health effects of pollution. This line of research builds on an older set of work debating the impact of pollution in 19th

[3] An alternative strategy to addressing these questions is to use a structural approach, as in Sieg *et al.* (2004) and Smith *et al.* (2004).

century British cities (Williamson (1981b), Pollard (1981), Williamson (1981a)), as well as work by historians documenting the impacts of pollution during this period (e.g., Mosley (2001), Thorsheim (2006), Brimblecombe (1987)).

3 Model

This section introduces a spatial equilibrium model building on Rosen (1979) and Roback (1982).[4] However, the standard model is modified in a few important ways. In order to incorporate the impact of industrial pollution, I model a disamenity related to the level of coal use by industry in a city. Also, to reflect the Bartik-style approach used in the empirical analysis, which relies on variation in the pollution intensity of industries, the standard framework is modified to incorporate many industries which vary in their intensity of coal use.

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 this literature, workers and firms can move freely across these cities. Goods are also freely traded across cities. I begin by modeling the demand for labor in cities.

3.1 Labor demand

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, and a fixed local industry-specific resource. In keeping with the empirical setting, I call the polluting input coal, though it may correspond to other polluting inputs in a different setting. The production function is,

[4] Within this tradition, I draw specifically on the recent models of Moretti (2011) and Diamond (2012).

$$y_{fic} = 2\, a_i \left[\left(\frac{1}{2}\right) L_{fic}^{\rho} + \left(\frac{1}{2}\right) \left(\frac{C_{fic}}{\theta_i}\right)^{\rho} \right]^{\alpha/\rho} R_{fic}^{1-\alpha} \,,$$

where L_{fic} is labor, C_{fic} is coal, R_{fic} is the local resource, and a_i is the industry technology level. The parameter $\theta_i \in (0, +\infty)$ determines the importance of coal as an input into each industry. The elasticity of substitution between labor and coal in production is governed by the parameter ρ, while $\alpha \in (0,1)$ determines the importance of local resources in production.

Within each city, there is a fixed supply of resources $\bar{R}_{ic}$ available for each industry. These resources can be thought of as natural features or local endowments of entrepreneurial ability in a particular sector. These resources play an important role in this model, by allowing multiple cities to be active in an industry even when productivity varies across cities and the market is perfectly competitive.[5]

Given this production function, firms maximize profit subject to output prices p_i, the coal price ϕ, a city wage w_c, and the price of local resources χ_{ic}.[6] The firm's maximization problem is,

$$\max_{L_{fic}, C_{fic}, R_{fic}} p_i a_i\, 2 \left[\left(\frac{1}{2}\right) L_{fic}^{\rho} + \left(\frac{1}{2}\right) \left(\frac{C_{fic}}{\theta_i}\right)^{\rho} \right]^{\alpha/\rho} R_{fic}^{1-\alpha} - w_c L_{fic} - \phi C_{fic} - \chi_{ic} R_{fic} \,.$$

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

$$\frac{C_{fic}}{L_{fic}} = \theta_i^{\frac{\rho}{\rho-1}} \left(\frac{\phi}{w_c}\right)^{\frac{1}{\rho-1}} \,.$$

[5]This approach follows Jones (1975) and has recently been used in papers by Kovak (2013) and Hanlon & Miscio (2014).

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

Next, I make the simplifying assumption that firms within an industry have very little ability to substitute between workers and coal. This assumption reflects the empirical approach and simplifies the remaining theoretical analysis, though I will also check the robustness of the empirical results to this assumption. Specifically, I consider the limit as $\rho \to -\infty$. Under these conditions, the ratio of coal use per worker for any firm in industry i is $C_{fit}/L_{fit} \to \theta_i$.

Plugging this into the first order conditions from the firm's optimization problem and summing over firms in an industry, the industry labor demand equation is,

$$w_c = \alpha\, p_i a_i\, L_{ic}^{\alpha-1} \bar{R}_{ic}^{1-\alpha}\,. \tag{1}$$

Note that, in equilibrium, the sum of firm resource use must equal total city-industry resources, which are fixed at $\bar{R}_{ic}$. Summing across industries and taking logs, the city labor demand equation is,

$$\ln(w_c) = (\alpha - 1)\ln(L_c^{dem}) + (1-\alpha)\ln\left[\sum_i (\alpha p_i a_i)^{\frac{1}{1-\alpha}}\, \bar{R}_{ic}\right]. \tag{2}$$

This expression tells us that the local labor demand curve is downward sloping and that the location of the curve depends on the sum across industries of a term that includes factors affecting industry demand (p_i) and industry productivity (a_i), weighted by city-industry resources $\bar{R}_{ic}$.

3.2 Housing supply

The supply of labor in a city will be influenced by the availability of housing. The housing market itself is not a central focus of this paper, so I model housing in a very reduced-form way,

$$\ln(r_c) = \lambda \ln(L_c) + \ln(\eta_c)\,, \tag{3}$$

where η_c represents city-specific factors that influence construction costs and $\lambda > 0$ is a parameter that determines the impact of increasing population on housing price.[7]

3.3 Labor supply

The model is populated by a continuum of homogeneous workers, each of which supply one unit of labor to the market. The utility of any worker j in city c is given by,

$$U_{jc} = \gamma \ln(X_{jc}) + (1-\gamma)\ln(h_{jc}) + \ln(A_c)\,,$$

where X_{jc} is an index over consumption goods, h_{jc} is consumption of housing, A_c is the amenity value in city c, and the $\gamma \in (0,1)$ parameter determines the relative expenditure shares of housing and goods. The budget constraint is, $w_c \geq PX_{jc}+r_c h_{jc}$ where P is the national price index over goods and r_c is the local price of housing.[8] For simplicity, the budget constraint does not reflect income from housing rents or the rent on local resources.[9] The utility of a worker in city c can be expressed using the indirect utility function,

$$V_c = \gamma \ln\left(\frac{w_c}{P}\right) + (1-\gamma)\ln\left(\frac{w_c}{r_c}\right) + \ln(A_c)\,.$$

Workers are freely mobile across cities and have an outside option utility v^* which we can think of as either the utility of emigrating or the utility of living in the rural areas of the country. Given this, and using Eq. 3, the labor supply equation for city c is,

[7] 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 the modern setting, particularly in the U.S., 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).

[8] At this point I could normalize the goods price index to $P = 1$, but to ease the mapping of the theory to the empirical exercise I continue to include P explicitly in the model.

[9] This simplifies the model, but it is not central to the results.

$$\ln(w_c) = \gamma \ln(P) + (1-\gamma)\lambda \ln(L_c^{sup}) + (1-\gamma)\ln(\eta_c) - \ln(A_c) + v^* \,. \tag{4}$$

This expression tells us that the supply of labor in a city is positively related to the city wage and the amenity offered by the city and negatively related to the cost of constructing housing.

3.4 Amenities

The amenity value in a city may be related to many features, but this study emphasizes the role of coal-based pollution in affecting city amenities. Thus, I model the amenity value as,

$$\ln(A_c) = \ln(\delta_c) - \beta \ln(CP_c)\,, \tag{5}$$

where CP_c is a measure of coal pollution in the city and δ_c represents a fixed city amenity. Let $CP_c = \sum_i C_{ic} = \sum_i \theta_i L_{ic}$ where θ_i is coal use per worker in industry i.

3.5 Equilibrium and theoretical results

Given the outside option utility $v*$, the national coal price ϕ, a set of national industry output prices $\{p_i\}$, technology levels $\{a_i\}$, and city industry resources $\{\bar{R}_{ic}\}$, equilibrium in a city is defined as the set of local wages w_c, resource prices χ_{ic}, rent r_c, and population L_c, together with a set of industry employment and coal use levels $\{L_{ic}, C_{ic}\}$ 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.

Given these equilibrium conditions, I derive two results that can be taken to the data. The first result relates the quality-of-life in a city to the level of industrial pollution. This will be useful for verifying that the pollution measure I construct is capturing a meaningful city disamenity. The second result, which motivates the main empirical specification, relates growth in city employment to employment growth in the industries present in the city as well as local industrial pollution.

To relate the quality-of-life in a city to the amenity level, I begin with Equation 4. Substituting in Equation 5 and reorganizing, I obtain,

$$[\gamma \ln(P) + (1-\gamma)\ln(r_c)] - \ln(w_c) = -\beta \ln(CP_c) + \ln(\delta_c) - v^* \,. \tag{6}$$

The left-hand side of this expression is the difference between the local price level, which depends on the prices of goods and housing weighted by their importance in the utility function, and the local wage. This is a standard quality-of-life measure in the urban economics literature. The right-hand side of the equation is composed of factors that can affect urban quality-of-life. The coefficient β reflects the extent to which industrial coal use is creating a city disamenity that affects city quality-of-life. In the empirical analysis, I use this expression to verify that my measure of industrial coal use is capturing a factor that meaningfully affects the quality-of-life of city residents.

For the main analysis, I need an expression that relates city employment to the endogenous disamenity related to industrial pollution. I begin by substituting Equation 2 in for the wage term in Eq. 4, reorganizing, and then adding the amenities expression from Eq. 5 to obtain,

$$\begin{aligned} \ln(L_c) \quad = \quad & \left(\frac{1-\alpha}{\sigma}\right) \ln\left[\sum_i (\alpha p_i a_i)^{\frac{1}{1-\alpha}} \bar{R}_{ic}\right] - \left(\frac{\beta}{\sigma}\right) \ln(CP_c) + \left(\frac{1}{\sigma}\right) \ln(\delta_c) \\ - \quad & \left(\frac{\gamma}{\sigma}\right) \ln(P) - \left(\frac{1-\gamma}{\sigma}\right) \ln(\eta_c) - \left(\frac{1}{\sigma}\right) v* \,, \end{aligned} \tag{7}$$

where $\sigma = (1 - \alpha + \lambda(1-\gamma)) > 0$. The first term on the right-hand side represents the positive impact of increased productivity or increased demand in a city industry on city population, summed over all industries. This is the *direct* impact of industry growth on city size. The second term on the right-hand side represents the negative impact of industrial coal use on city population. Note that this term will also be affected when productivity or demand grows for a polluting industry. Thus, this represents the *indirect* impact of growth in a polluting industry on city population.

Next, I need to express the unobserved industry productivity and demand shifts, represented by the $(\alpha p_i a_i)^{\frac{1}{1-\alpha}} \bar{R}_{ic}$ term in Eq. 7, using observable employment data. To make progress here, it is necessary to add time subscripts to the model and exploit the dynamic nature of the data. Using Eq. 1, summing across locations, and differencing across periods, the national growth rate of an industry i is,

$$\frac{L_{it}}{L_{it-1}} = \Omega_i \left(\frac{\alpha p_{it} a_{it}}{\alpha p_{it-1} a_{it-1}} \right)^{\frac{1}{1-\alpha}} \quad \text{where} \quad \Omega_i = \left(\frac{\sum_c w_{ct}^{-\frac{1}{1-\alpha}} \bar{R}_{ic}}{\sum_c w_{ct-1}^{-\frac{1}{1-\alpha}} \bar{R}_{ic}} \right) . \tag{8}$$

In this expression, Ω_i reflects how changes in city wage levels interact with local industry-specific resources to affect national industry growth rates.

Eq. 8 will be used to capture industry growth driven by national demand and productivity shifts. These shifts will interact with city-industry resource endowments to determine the effect on industry employment growth in a particular city. Using Eq. 1, local city-industry endowments can be expressed as,

$$\bar{R}_{ic} = L_{ict-1} w_{ct-1}^{\frac{1}{1-\alpha}} (\alpha p_{it-1} a_{it-1})^{\frac{1}{\alpha-1}} . \tag{9}$$

Putting Eqs. 8 and 9 together and summing across industries, I obtain,

$$\left[\sum_i (\alpha p_{it} a_{it})^{\frac{1}{1-\alpha}} \bar{R}_{ic} \right] = w_{ct-1}^{\frac{1}{1-\alpha}} \left[\sum_i \left(\frac{L_{it}}{L_{it-1}} \right) L_{ict-1} \Omega_i^{-1} \right] . \tag{10}$$

Finally, using Eqs. 4 and 5 to substitute out the w_{ct-1} term in Eq. 10, and then plugging this into Eq. 7, I obtain,[10]

$$
\begin{aligned}
\ln(L_{ct}) - \ln(L_{ct-1}) &= \left(\frac{1-\alpha}{\sigma}\right) \ln\left[\sum_i \left(\frac{L_{it}}{L_{it-1}}\right)\left(\frac{L_{ict-1}}{L_{ct-1}}\right)\Omega_i^{-1}\right] \\
&- \left(\frac{\beta}{\sigma}\right)\left[\ln(CP_{ct}) - \ln(CP_{ct-1})\right] - \left(\frac{1}{\sigma}\right)\left[v_t^* - v_{t-1}^*\right] .
\end{aligned} \tag{11}
$$

This equation tells us that the change in log city employment can be expressed as a function of (1) a term that depends on initial city employment interacted with the national industry growth rate (subject to the adjustment factors Ω_i), (2) the change in city coal use, and (3) time-varying national-level factors (the outside option utility). This expression forms the basis for the main empirical analysis.

Note that, by exploiting time variation in the data, I have eliminated any time-invariant city-specific factors, including the housing costs shifters (η_c) and other city amenities (δ_c), from Eq. 11. Also, note that the first term on the right-hand side of this equation is based on initial industry composition interacted with national industry growth rates, as in the standard Bartik instrument. Thus, Eq. 11 provides a theoretical motivation for this instrument, up to the adjustment factors Ω_i. These adjustment factors, which will be present in any study using the Bartik approach to investigate endogenous city amenities, will be discussed in more detail later.

4 Empirical setting

Coal was a major and visible source of pollution in the cities of 19th century Britain. While pollution was prevalent in 1851, when this study begins, it grew substantially over the period I consider. British coal consumption averaged 65 million tons annually in 1852-1862 and rose to 181 million tons in the 1903-1912 period.[11] This amounted to

[10]Here I have normalized the price index to one in each period.
[11]These figures are from the U.K. Department of Energy and Climate Change.

4.3 tons per person in 1911.[12] The foul smell, reduced visibility, and negative health effects of coal-based pollution were widely recognized and discussed. An example is provided by *The Times* (Feb. 7, 1882, p. 10)[13]:

> *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.*

Many other contemporary sources also commented on the negative health effects of coal-based pollution. For example, in 1880 the British Medical Journal estimated that a particularly bad London fog in the winter of 1880 led to over 1000 excess deaths in one week.[14] These health impacts are corroborated in more recent studies, such as Troesken & Clay (2011) and Hanlon (2015)

Coal-based pollution came from both industrial and residential sources. Residential coal use was more polluting, per ton burnt, than industrial pollution because combustion was generally less efficient and smoke was emitted at lower altitudes.[15] However, overall coal use by industry substantially exceeded residential use. Data from Mitchell (1988) show that industrial coal use in manufacturing accounted for 60-65% of total British domestic coal use over the period I consider, while residential use accounted for 17-25% of domestic coal consumption.[16] This suggests that industrial pollution was likely to be an important part of overall city pollution. Moreover, because some industries were particularly intensive users of coal, and these industries tended to agglomerate, industrial coal use was more geographically concentrated than residential use.

[12] 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.

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

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

[15] This feature, together with the important role played by residential smoke in London, has motivated much of the previous literature on this topic to focus on residential coal use.

[16] The remainder is composed of use by transportation and utilities.

While contemporaries were at least partially aware of the negative effects of pollution, this did not translate into a strong regulatory response, due largely to the laissez faire ideology that dominated British policy-making during this period. New legislation was passed, but the goal of these regulations was to reduce excess pollution from factories to the level practical for each manufacturing trade, not to eliminate those trades that necessarily produced high levels of air pollution.[17] In addition, these acts allowed for substantial interpretation, contained important loopholes, and imposed relatively small fines.[18] Historians suggest that, as a result of these features, pollution regulation was largely ineffective during the 19th century (Thorsheim (2006), Fouquet (2012)), though there is evidence that regulation began to have some impact towards the end of the 19th century (Troesken & Clay (2011)).

A similar lack or regulatory interference applied to city growth during this period. For example, there were no zoning laws in England until 1909. This meant that cities were able to grow in response to increases in local labor demand. Another feature that allowed cities to adjust their size during this period was the high levels of internal and international migration. Summarizing our knowledge in this area, Long & Ferrie (2003) characterizes the British population as "highly mobile."[19] Much of this migration was made up of rural residents moving to the cities. Baines (1985) suggests that, controlling for the age structure of migrants, internal migration accounted for about 40% of the population growth in British cities. These features mean that, in the empirical setting I consider, city employment and population should be responsive to changing employment opportunities and amenity levels.

[17] The new regulations included The Sanitary Act of 1866, The Public Health Act of 1875 , and The Public Health (London) Act of 1891. For further details, see Appendix A.1.1.

[18] 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.

[19] The sentiment is echoed by Baines (1994), who argues that "both the housing and labor markets were more open than today and migrants were less likely to be deterred by the problems of educating children or looking after relatives."

5 Data and measurement

5.1 Data

The first necessary piece of information for this study is a measure of the coal intensity of each industry. This information is drawn from the first British Census of Production, which was completed in 1907. While these data come from near the end of the study period, this is the earliest available consistent source for this information. Because of the central role coal played in the British economy, this Census collected detailed information on the amount of coal used in each industry, as well as industry employment.[20] This allows me to construct a measure of coal use per worker in each industry (corresponding to θ_i in the model).

Table 1 describes the resulting measures of industry coal use per worker, as well as industry employment in 1851. The most intensive industrial users of coal were Metal & Machinery, Earthenware & Bricks, and Chemical & Drug Manufacturing. These were industries that used coal to heat material. Moderate coal-using industries, such as Textiles, generally used coal to power steam engines. Other industries, such as Apparel or Tobacco products, used very little coal. As we will see, one consequence of the substantial variation in coal use intensity at the industry level was substantial variation in the amount of industrial coal use at the city level.

To translate coal use by each industry into the amount of industrial coal burned in each city, I use information on the size of each industry in each city over time. These data come from the Census of Population, which collected the occupation of each person at each ten-year census interval. These occupational categories generally correspond to industries, such as "Cotton spinner" or "Steel manufacturer".[21] Consistent city-industry series have been constructed for the period 1851-1911 by

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

[21] 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.

combining occupational categories from the various censuses. The resulting British city-industry database covers 31 cities and 26 industries, spanning nearly the entire private-sector economy.[22]

Table 1: Industry coal use per worker and employment in 1851

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

Coal per worker 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.

Out of this set of 26 available industries, the main analysis focuses on 16 manufacturing industries, corresponding to those described in Table 1. There are two reasons to focus primarily on manufacturing industries.[23] First, manufacturing industries are more likely to produce products that can be traded outside of the city, so this group of industries is a better fit for the theory than industries producing non-traded goods.

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

[23]The focus on manufacturing industries leads me to exclude several service industries, road, rail and water transportation, local utilities, and construction.

Second, some important control variables can only be constructed for the manufacturing industries for which data are available from the Census of Production. However, I will also conduct additional robustness exercises where I analyze impacts across all private sector industries.

The set of cities in the database includes the largest cities in England based on 1851 population. Figure 2 in the Appendix shows the location of these cities. I have also collected data for additional cities in 1901 that will be used when analyzing city quality-of-life.

To measure the quality-of-life in cities, I use data on rent, goods prices, and wage levels. These data are available for a cross-section of cities in 1905 based on a report prepared by the Board of Trade. This report covers 77 major English towns in 1905. I am able to construct pollution measures for 51 of these. Further details on these data are available in Appendix A.2.2.

5.2 Measuring industrial coal use in cities

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, ψ_t:

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

The 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.2.5, shows two results. First, the 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 ψ_t term.

Another assumption implicit in these coal use measures is that industry coal use does not vary substantially across locations in response to variation in the relative level of wages or coal prices. In the model, this is reflected in the assumption that $\rho \rightarrow -\infty$. Without data on the coal intensity of industries by location, it is not possible to directly assess the extent to which this occurred. However, using data on wages and coal prices by city, available for 1905 from the Board of Trade, I also assess the sensitivity of my results under various values of ρ.

Estimates of θ_i for manufacturing industries are provided by the 1907 Census of Production, while Census of Population data provide city-industry employment. To calculate ψ_t, I use the fact that,

$$\ln(\psi_t) = \ln(COAL_t) - \ln\left(\sum_c \sum_i L_{ict} * \theta_i\right) .$$

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 ψ_t term will not affect the estimated coefficients because regressions are run on the log of coal use across locations within a period. In these regressions, the ψ_t term will be absorbed into the year effects included in the regressions. However, this term will affect the overall impact of coal use on city growth.

The coal use measure reflects only pollution from manufacturing industries. In general, other industries, such as services, were not likely to be major coal users, so this measure should capture most industrial coal use. An exception is local utilities, particularly gas, which was a major user of coal.[24] 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.

Estimates of industrial coal use per worker at the city level are described in Table 8 in the Appendix. These estimates suggest that there was substantial variation across cities in the expected level of coal use per worker. Cities specializing in heavy industry, such as Birmingham and Sheffield, show levels of coal use per worker that are nearly double the national average. Textile manufacturing towns, such as Manchester and Leeds, show moderate levels, near the national average. Commercial and trading cities, such as Liverpool and Bristol, are among the least intensive users of industrial coal. Even among similarly sized cities, coal use intensity varied substantially.

To verify that this approach generates a measure that accurately reflects variation in local coal use intensity going back in time, I have calculated estimates of industrial coal use at the county level for 1871 and compared the results to county-level coal use measures based on an 1871 parliamentary report which surveyed coal use in industries in some English counties. This analysis, available in Appendix A.2.6, suggests that my approach does a good job of replicating industrial coal use at the county level in 1871, particularly in more industrial and urbanized locations.

Finally, it is necessary to settle on a functional form for the impact of city coal use on city residents. The mortality results from Hanlon (2015) can be helpful here. The patterns documented in that study suggest that city mortality is very close to linear in log coal use. To the extent that mortality patterns reflect amenity value, this suggests that log coal use should be used as the main explanatory variable. Similar

[24]Coal was used to make gas, which was then pumped to users in the city, where it was burned for light or heat. Another potential exception is transportation, particularly rail transportation, which used a substantial amount of coal. However, most of this coal would have been burnt 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.

concave patterns are also documented by Clay *et al.* (2014), and appear in the the medical literature (e.g., Pope III *et al.* (2011)). Based on these results, the log of coal use is used as my measure of coal pollution in the primary analysis.

5.3 Verifying the measure: coal use and city quality-of-life

Before turning to the main analysis, it is useful to establish that the coal use measure is capturing a meaningful city disamenity. One way to do so is to estimate the relationship between industrial coal and city quality-of-life. The estimating equation is suggested by Equation 6, where city quality-of-life is measured as the difference between the local cost of living and local wages, as in a standard Rosen-Roback approach.[25] Equation 6 suggests that quality-of-life should be negatively related to my measure of industrial pollution, if it is capturing a meaningful city disamenity.

The estimation uses data on wages and costs for 51 cities in 1905 from a study conducted by the Board of Trade.[26] The wage series are available for specific occupations which were present in many cities and relatively similar across locations. I use wage series for two skilled occupations: skilled builders and skilled engineers (engine operators).[27] 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. Both series are presented as index values relative to London.

Using the wage, price and rent data, the QOL measures is[28]:

[25] 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.

[26] The Board of Trade data cover 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.

[27] 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. Moreover, 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.

[28] Plots of the quality-of-life measures are available in Appendix A.2.7.

$$QOL_c = COST_c - Wage_c\,.$$

Table 2 presents the results of regressions comparing these quality-of-life measures to city coal use. Columns 1-3 describe results using the QOL measure based on the wages of skilled builders while Columns 4-6 use the QOL measure 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.[29] In all specifications, city coal use is negatively related to the amenity value of the city, and this relationship is statistically significant in most of the results. None of the other control variables appear to have a consistent and statistically significant relationship with the city amenity value. These results suggest that that city coal use was negatively related to the amenity value of a city. However, given the nature of these cross-sectional regressions and the sample size, these results should be interpreted with some caution.

A second piece of evidence suggesting that the coal use measure is capturing a meaningful disamenity is provided by Hanlon (2015), which compares an identical coal use measure (constructed at the district level) to mortality data for over 500 English districts. That paper shows that coal-based pollution increased British mortality by 5-7% in 1851-1860. Moreover, increased coal use from 1851-1900 raised age-standardized mortality by 1.35-1.46 per thousand, compared to an average overall level of 19.6 per thousand in 1851-1860. These mortality effects were concentrated in cause-of-death categories associated with pollution, particularly respiratory deaths, which suggests that they were not driven by the selection of less healthy people into

[29]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 2 reports the larger robust standard errors.

more polluted areas or by on-the-job accidents. These mortality results provide further evidence that the coal use measure is a reasonable reflection of the city disamenity related to industrial pollution.

Table 2: Comparing QOL 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)
Log salaried shr.			-0.0188			0.0101
			(0.0837)			(0.0701)
Latitude			-0.00320			-0.0108
			(0.0103)			(0.0101)
Air frost days			-0.00200*			5.41e-05
			(0.00116)			(0.000893)
Rainfall			0.0276			-0.0635
			(0.0727)			(0.0691)
Constant	0.207	0.148	0.372	0.384**	0.375***	0.971*
	(0.125)	(0.110)	(0.541)	(0.145)	(0.139)	(0.532)
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. The salaried worker share variable is constructed using city-industry data from 1901. Note that wage data for skilled engineers is available for fewer cities than wage data for skilled builders.

6 Analysis

The main focus of the paper is analyzing the impact of industrial pollution on city employment. The analysis, which is based on Eq. 11 from the theory, begins with three ingredients. The first ingredient,

$$\Delta \ln(WORKPOP_{ct}) = \ln\left(\sum_i L_{ict}\right) - \ln\left(\sum_i L_{ict-\tau}\right),$$

is the change in the log working population of a city, where the working population is just the sum of employment in the analysis industries. This corresponds to the term on the left-hand side of Eq. 11.

The second ingredient,

$$\Delta \ln(PredWP_{ct}) = \ln\left(\sum_i (L_{ic,t-\tau})\, GREMP_{it,t-\tau}^{-c}\right) - \ln\left(\sum_i (L_{ic,t-\tau})\right),$$

is the predicted growth of the working population in a city. This is constructed by taking the employment share of each industry in a city in year $t-\tau$, where τ will generally be two decades, multiplying by the industry growth rate in all cities other than c in the intervening period, and then summing across all industries in a city. This variable captures the impact of changes in industry productivity or demand that affect employment in a city. It corresponds to the first term on the right-hand side of Eq. 11, up to the adjustment factors Ω_i, which are discussed below.

The third ingredient,

$$\Delta \ln(PredCOAL_{ct}) = \ln\left(\psi_t \sum_i \left(L_{ic,t-\tau} * GREMP_{it,t-\tau}^{-c} * \theta_i\right)\right) - \ln\left(\psi_\tau \sum_i (L_{ic,t-\tau} * \theta_i)\right),$$

is the predicted change in the level of log coal use in a city. The construction of this variable is similar to that of $\Delta \ln(PredWP_{ct})$, except for the additional θ_i and ψ_t terms. The ψ_t terms will be absorbed by year effects and therefore will not affect estimation results. Thus, the only variation between $\Delta \ln(PredWP_{ct})$ and $\Delta \ln(PredCOAL_{ct})$ that is relevant for the regression results is generated by variation in the pollution intensity of industries, through the θ_i terms. The $\Delta \ln(PredCOAL_{ct})$ variable corresponds to the pollution terms in Eq. 11.

Given these ingredients, estimates are obtained using pooled cross-sectional regressions of the form,

$$\Delta \ln(WORKPOP_{ct}) = b_0 + b_1 \Delta \ln(PredWP_{ct}) + b_2 \Delta \ln(PredCOAL_{ct}) + X_{ct}\Gamma + \mu_t + e_{ct}\,, \quad (12)$$

where X_{ct} is a set of control variables and μ_t is a full set of time effects. Summary statistics for the data used to estimate Eq. 12 are available in Appendix A.3.2.

Note that the specification in Eq. 12 mirrors Eq. 11 up to the adjustment factors Ω_i, with μ_t capturing the effect of changes in the outside option of workers. The adjustment factors are industry specific and represent the extent to which employment growth in industry i differs from the level that would be predicted if we could observe the actual demand and productivity values (p_i and a_i). This deviation is caused by the influence of growth in an industry on overall city wage levels in cities other than c, which are used to calculate the predicted national industry growth rates.

Given the data available in this study, it is not possible to directly account for the industry adjustment factors. However, there are two reasons to think that ignoring the adjustment factors (i.e., setting $\Omega_i = 1$) is reasonable. First, the effect that the adjustment factors represent is second-order. This is because they depend on the influence of growth in industry i on city-level wages affecting employment growth in industry i, but industry i is just one of many industries that influence city wage levels. Thus, we should expect the adjustment factors to be very close to one. Second, it is possible to sign the bias generated by the adjustment factors. The adjustment factors will bias the estimated effect of city coal use because when growth occurs in industries that are more polluting, the employment growth rate in that industry in cities other than c will understate the employment growth that we would have predicted if p_i and a_i were observed. This will act to bring the predicted employment growth closer to actual employment growth, leaving a smaller employment gap for the coal use variable to explain. As a result, the estimated impact of city coal use on city employment growth will be biased towards zero. Finally, note that in ignoring these adjustment factors I am following existing work using the Bartik approach to study endogenous city disamenities.

There is a simple intuition underlying the estimation strategy described by Eq. 12: on average, initial city-industry employment shares interacted with national industry growth rates should correctly predict actual city employment growth once other factors affecting city growth – such as industrial pollution – are accounted for. This implies that the coefficient on $\ln(PredWP_{ct})$ should be close to one once other important factors affecting city employment growth are included in the regression. Specifications that deliver an estimated b_1 coefficient that differs substantially from one are likely to be missing some important factors influencing city employment growth.

Before running regressions, it is helpful to build intuition graphically. Let,

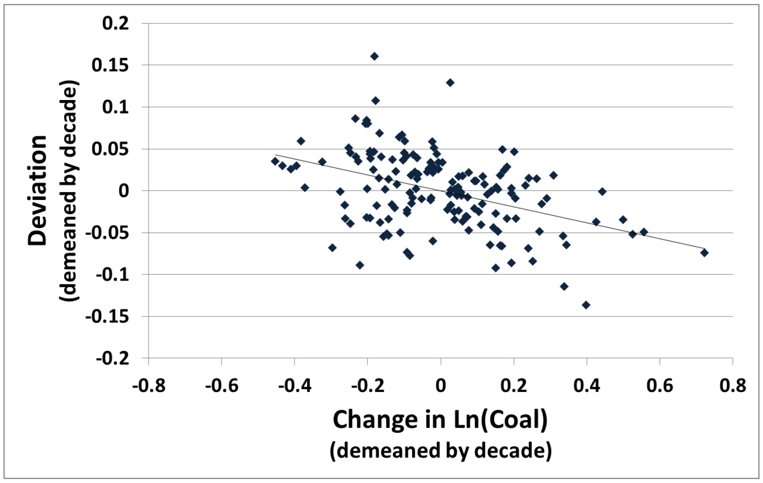
$$DEVIATION = \Delta \ln(WORKPOP_{ct}) - \Delta \ln(PredWP_{ct}) \, ,$$

represent the deviation between the actual change in log population observed in the data over any 20 year period and the change predicted using initial city-industry employment and national industry growth rates. Figure 1 plots the deviation against the change in log coal use, pooling all cities and two-decade periods in the data. To allow me to compare all decades, I demean the variables by decade.[30] This figure reveals a clear negative relationship between the increase in coal use in a city and the deviation between actual and predicted city employment growth. Put another way, this figure suggests that industrial pollution can help explain why city employment

[30]Similar results hold when looking separately at each year. Figures are available in the Appendix.

levels deviate from the level we would expect given the initial industrial composition of cities and national industry growth rates.

Figure 1: Deviation of actual from predicted growth vs. change in city coal use



Deviation is the difference between actual and predicted growth in the working population of the city. Both series are demeaned by their average across all cities in each year, so that all years can be plotted in the same graph. Graphs for each year separately are available in the Appendix.

Turning to the estimation, note that in a regression of $\Delta \ln(WORKPOP_{ct})$ on $\Delta \ln(PredCOAL_{ct})$ that excludes the predicted employment variable, the coefficient on $\Delta \ln(PredCOAL_{ct})$ will capture both the negative effects of pollution on city growth as well as the positive direct effect of industry growth. Similarly, if I run a regression of $\Delta \ln(WORKPOP_{ct})$ on $\Delta PredWP_{ct}$ but omit the coal use variable, then the estimated coefficient on $\Delta PredWP_{ct}$ will pick up a mix of the positive and negative effects. However, when both the predicted employment and coal use are included in the regression, $\Delta PredWP_{ct}$ will capture the direct effect of predicted industry growth on city growth, while the coal use variable will capture the additional impact of industrial pollution on city growth.

While Eq. 12 is estimated using pooled cross-sectional variation, it is important to keep in mind that many of the factors that we might worry about in standard cross-sectional regressions will be differenced out or accounted for by the $PredWP_{ct}$ variable. For example, the $PredWP_{ct}$ variable will incorporate any factors that are present in the base year, including the initial level of coal use in a city, other factors that affect city amenities, fixed city housing cost shifters, etc. As a result, omitted variables that affect the level of city population but do not vary over time are not a concern in this estimation approach.

One threat to identification is that there may be features that vary across industries, and therefore across cities, in a way that is correlated with industry coal use, and affect city employment growth. For example, coal using industries may be more innovative than other industries. To help control for variation in the level of innovation across cities, I include patents per capita in a city from 1852-1858. Another example is that coal using industries may use more (or fewer) high skilled workers, which previous work has shown to be important in modern economies.[31] To help deal with this I use variation in the salaried workers share across industries to generate a variable representing the predicted change in the log share of salaried workers in a city.[32] I have also collected a set of environmental variables – latitude, air frost days, and rainfall – which may affect city amenities, the impact of coal burning on the city environment, and the level of residential coal use.[33] Further details on the construction of these control variables are available in Appendix A.2.3.

Serial correlation is a concern when pooling observations for the same city from multiple decades. Spatial correlation is also a potential concern. To deal with these issues, I have calculated results allowing for serial correlation within the same city over two decades, following Newey & West (1987), and spatial correlation between cities within 50km of each other following Conley (1999).[34] In almost all specifica-

[31] E.g., Rauch (1993), Moretti (2004), Shapiro (2006), and Diamond (2012).

[32] This follows the approach in Diamond (2012).

[33] An air frost day is a day in which the temperature of the air one meter above the ground drops below zero degrees Celsius. Rainfall is measured in meters per year. Both of these series come from the Met weather service and are based on averages for 1981-2010 from the weather station nearest each city.

[34] I use the implementation from Hsiang (2010).

tions, allowing serial and spatial correlation reduces the confidence intervals, leading to stronger results. This is consistent with negative spatial correlation.[35] To be conservative, in the regression tables I report the larger robust standard errors.

As a starting point for the analysis, I consider pooled regression results covering 1871-1911. For each year, I use values from two decades before as the base year for constructing the $\Delta \ln(PredWP_{ct})$ and $\Delta \ln(PredCOAL_{ct})$ variables. Table 3 presents the results. In Column 1, the $\Delta \ln(PredWP_{ct})$ variable is included, but $\Delta \ln(PredCOAL_{ct})$ is not. We can see that this delivers a negative coefficient on $\Delta \ln(PredWP_{ct})$, suggesting that some important factors affecting city employment growth are missing from this specification. Column 2 includes only the $\Delta \ln(PredCOAL_{ct})$ while omitting $\Delta \ln(PredWP_{ct})$. We can think of this specification as combining both the positive direct effect of employment growth in polluting industries together with the negative indirect effects created by the pollution they emit. The results suggest that the net effect of city coal use on city employment growth, including both the positive direct and negative indirect effects, is negative. It may be surprising to find a negative net effect, but this is consistent with existing work such as Chay & Greenstone (2005).

In Column 3, both $\Delta \ln(PredWP_{ct})$ and $\Delta \ln(PredCOAL_{ct})$ are included. In this specification, $\Delta \ln(PredWP_{ct})$ will pick up the positive direct effect of employment growth and, as expected, this results in a positive coefficient. Moreover, we can see that the coefficient on $\Delta \ln(PredWP_{ct})$ is now closer to, and statistically indistinguishable from, one. This suggests that including city coal use partially explains the deviation between predicted and actual city population. The $\Delta \ln(PredCOAL_{ct})$ term picks up the additional impact of industrial coal use. As expected, the coefficient on this term is negative.

[35] Hanlon (2014) provides evidence indicating that employment growth in English cities was characterized by negative spatial correlation during this period. Specifically, that study shows that when cities experienced negative local shocks, other nearby cities tended to benefit, resulting in negative spatial correlation.

The impact of increased coal use on city growth, suggested by the estimates from Column 3 of Table 3, are substantial; relative to the average city, a city with coal use growth that is one standard deviation (s.d.) above the average is expected to have annual employment growth that is lower by 2.05 percentage points.[36] This suggests that cities with modest employment growth but substantial growth in coal use could have experienced negative growth as a result. However, because both employment growth and industrial pollution are driven by industry growth, it is unlikely that a city would experience growth in coal use that was far above average without also having more rapid employment growth. Consistent with this, it is rare to observe cities experiencing negative growth in the data.

Table 3: Pooled regression results

DV: Δ Log of city working population in analysis industries			
	(1)	(2)	(3)
$\Delta \ln(PredWP_{ct})$	-0.413		1.361**
	(0.255)		(0.611)
$\Delta \ln(PredCOAL_{ct})$		-0.813**	-2.290***
		(0.350)	(0.833)
Constant	0.390***	0.840***	1.508***
	(0.0798)	(0.250)	(0.447)
Year effects	Yes	Yes	Yes
Observations	155	155	155
R-squared	0.049	0.072	0.095

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Robust standard errors are displayed because they are more conservative (deliver larger confidence intervals) than standard errors allowing spatial or serial correlation. The data cover 1871-1911, with values values from two decades before each year used as the basis for constructing the $\Delta \ln(PredWP_{ct})$ and $\Delta \ln(PredCOAL_{ct})$.

In interpreting these estimates, it is important to keep in mind that they do not incorporate the general equilibrium effects on the outside option of workers caused by the reduction in migration to cities due to pollution. Thus, it is better to interpret

[36]This figure is obtained by multiplying a one s.d. increase in coal use (0.178) by the coefficient from Table 3 and then converting the impact over two decades to an annual rate.

the impacts over shorter periods, such as one year, or smaller changes in the level of coal use.

Table 4 presents results that include the additional control variables I have collected. Column 1 includes region effects for five regions.[37] Column 2 adds in a control variable based on the number of patents per capita in a city from 1852-1858. This is a strong predictor of city growth, suggesting that more innovative cities were growing more rapidly during this period. However, including this control does not substantially change the estimated impact of city coal use. Columns 3 and 4 add in an additional control for the share of salaried workers in a city. I do not find strong evidence that more salaried workers increased city growth.[38] Column 5 adds in additional controls for initial coal use in the city as well as geographic features. Including these additional variables has little effect on the estimated outcomes. The geographic controls do not have explanatory power, which suggests that these effects are being differenced out or captured by the $\Delta \ln(PredWP_{ct})$ variable. The level of city coal use also has limited impact on the results, which suggests that initial city features related to coal use are adequately captured by the $\Delta \ln(PredWP_{ct})$ variable. This result also makes it clear that it is the change in coal use, rather than the level, that is impacting city growth. Finally, note that the estimated coefficients on $\Delta \ln(PredWP_{ct})$ in all of these regressions are statistically indistinguishable from zero, suggesting that these are reasonable specifications.

In these regressions, each city is given equal weight. Additional results, available in Appendix A.3.3, are calculated while weighting by either the log of city population or overall city population. These weighted regressions deliver similar results.

I have explored the sensitivity of the results described in Table 4 to dropping individual cities from the database. Focusing on the specification in Column (5),

[37]These regions are London, the South, the Midlands, the Northwest (including Yorkshire), and the North.

[38]This is surprising given the modern literature on this topic. This may reflect a bias in technology growth towards unskilled workers during the period I study.

when individual cities are dropped, the coefficients on $\Delta \ln(PredCOAL_{ct})$ range from -1.71 to -3.11. These coefficients are all statistically significant at the 90% level or above, with most statistically significant at the 95% level. I have also explored the sensitivity of the results to using lag lengths that are longer or shorter than two decades when calculating the predicted employment and coal use variables. These results are available in Appendix A.3.3.

Table 4: Regression results with additional control variables

	DV: Δ Log of city working population in analysis industries				
	(1)	(2)	(3)	(4)	(5)
$\Delta \ln(PredWP_{ct})$	0.805	0.448	0.819	0.477	0.532
	(0.635)	(0.615)	(0.629)	(0.609)	(0.669)
$\Delta \ln(PredCOAL_{ct})$	-1.858**	-1.758**	-1.952**	-2.085**	-2.049**
	(0.889)	(0.835)	(0.946)	(0.901)	(0.946)
Patents per capita		14.54**		15.39**	18.01***
		(6.135)		(6.038)	(5.942)
Initial coal use					-0.0501*
					(0.0259)
Δ Log salaried share			-0.317	-1.122	-0.615
			(1.594)	(1.502)	(1.539)
Latitude					-0.0125
					(0.0410)
Air frost days					-0.00462
					(0.00314)
Rainfall					-0.0488
					(0.165)
Constant	1.325***	1.134**	1.395**	1.369**	2.812
	(0.502)	(0.464)	(0.571)	(0.547)	(2.363)
Year effects	Yes	Yes	Yes	Yes	Yes
Region effects	Yes	Yes	Yes	Yes	Yes
Observations	155	155	155	155	155
R-squared	0.194	0.229	0.194	0.232	0.285

*** $p<0.01$, ** $p<0.05$, * $p<0.1$. Robust standard errors in parentheses. The data cover 1871-1911, with values from two decades before each year used as the basis for constructing the $\Delta \ln(PredWP_{ct})$, $\Delta \ln(PredCOAL_{ct})$, as well as the city salaried worker share.

How did the impact of coal use on city growth evolve over the 1851-1911 period? Cross-sectional regression results, available in Appendix A.3.3, suggest that the impact of city coal use peaked in the 1871-1891 period and then declined. However, with only 31 observations, the estimated coefficients in these cross-sectional regressions are not statistically significant. Nevertheless, this pattern fits the results of Troesken & Clay (2011), who find that coal-based pollution in London peaked around 1890 and then declined.

The results presented thus far focus on the working population employed in manufacturing industries. We may also be interested in how coal use affects overall city employment, which includes employment in services, transportation, local utilities, etc., or overall city population, which also includes non-workers. Columns 1-2 of Table 5 provide results looking at the relationship between city coal use and city employment in all private-sector industries. Columns 3-4 of Table 5 provide results with total city population as the dependent variable.[39] Whether looking at all private-sector industries or overall city population, I continue to observe evidence of the negative impact of industrial pollution on city growth. The magnitude of the effects is smaller when including non-traded industries or non-workers, suggesting that these sectors were less able to reallocate spatially in the face of variation in city amenity levels, as we would expect.

[39] In all of these regressions, the $\Delta \ln(PredWP_{ct})$ variable is calculated using exactly the same set of occupations used to construct the dependent variable. I.e., in Columns 1-2 the $\Delta \ln(PredWP_{ct})$ variable is calculated while including non-manufacturing workers, while in Columns 3-4 this variable is calculated while also including all workers and non-workers.

Table 5: Results for all private-sector industries and total city population

	DV: Δ Log of city private-sector employment		DV: Δ Log of total city population	
	(1)	(2)	(3)	(4)
$\Delta \ln(PredWP_{ct})$	1.060	0.728	0.137	0.304
	(0.702)	(0.651)	(0.599)	(0.614)
$\Delta \ln(PredCOAL_{ct})$	-1.649**	-1.887***	-0.906**	-1.315**
	(0.721)	(0.705)	(0.383)	(0.639)
Year effects	Yes	Yes	Yes	Yes
Region effects		Yes		Yes
Additional controls		Yes		Yes
Observations	155	155	155	155
R-squared	0.072	0.282	0.367	0.478

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. The data cover 1871-1911. The results in Columns 1-2 include 26 private-sector industries, with values from two decades before each year used as the basis for constructing $\Delta \ln(PredWP_{ct})$ and $\Delta \ln(PredCOAL_{ct})$. The data in Columns 3-4 include 40 occupation categories, including all of the private-sector industries, plus government workers and non-workers. The additional controls include patents per capita, the initial level of coal use, latitude, air frost days, and rainfall. The city salaried worker share is not included in these regressions because it cannot be constructed for all private-sector industries.

An implicit assumption in the analysis presented thus far is that industry coal use intensity does not vary too much across locations in response to variation in coal prices and wage levels. Next, I conduct an exercise to assess the sensitivity of the analysis to this assumption. In the model, industry coal use per worker is given by,

$$\frac{C_{ic}}{L_{ic}} = \theta_i^{\frac{\rho}{\rho-1}} \left(\frac{\phi}{w_c}\right)^{\frac{1}{\rho-1}} ,$$

where ϕ/w_c is the ratio of the price of coal to the local wage. Thus far I have assumed $\rho \to -\infty$ so that firms are not able to adjust their coal use in response to local price levels and $C_{ic}/L_{ic} = \theta_i$. I now relax this assumption and consider how my results are affected under alternative values of ρ while allowing the coal price ϕ_c to vary across cities. Data on coal prices by location and the wage of skilled builders in each city from the 1907 Board of Trade report are used to construct ϕ_c/w_c. Given a value of ρ, these data are used to calculate estimates of C_{ic}/L_{ic} that vary by city and industry.

Table 6 presents results calculated with city coal use measures in which the intensity of industry coal use is allowed to vary across space. Columns 1-2 present results assuming $\rho = -1$. This corresponds to a case in which the exponent on θ_i is 1/2 and the exponent on w_c/ϕ_c is 1/2. Columns 3-4 present results with $\rho = -2$, and Columns 5-6 present results with $\rho = -3$. The results in this table suggest that the estimated impact of coal use is not too sensitive to allowing for spatial variation in coal use intensity. Moreover, allowing more spatial variation in coal use intensity (ρ closer to zero) increases the estimated impact of coal use on city employment. This suggests that, to the extent that the coal use variable used thus far did not reflect variation in industry coal use intensity across cities driven by variation in wages and coal prices, the main results may be understating the true impact of industrial pollution.[40]

Table 6: Results allowing for spatial variation in industry coal use intensity

DV: Δ Log of city working population in analysis industries						
Substitution parameter:	$\rho = -1$		$\rho = -2$		$\rho = -3$	
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \ln(PredWP_{ct})$	2.103*	1.941	1.741*	1.691*	1.618**	1.597*
	(1.170)	(1.209)	(0.883)	(0.932)	(0.787)	(0.843)
$\Delta \ln(PredCOAL_{ct})$	-2.486*	-2.469*	-2.216**	-2.288*	-2.147**	-2.236**
	(1.308)	(1.453)	(1.048)	(1.193)	(0.966)	(1.109)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Region effects		Yes		Yes		Yes
Additional controls		Yes		Yes		Yes
Observations	145	145	145	145	145	145
R-squared	0.071	0.269	0.076	0.273	0.079	0.275

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. The data cover 1871-1911, with values from two decades before each year used as the basis for constructing $\Delta \ln(PredWP_{ct})$, $\Delta \ln(PredCOAL_{ct})$, and the city salaried worker share. The set of observations is smaller than that used in the main analysis because wage and coal price data are not available for Bath and Brighton. The additional controls include the log change in city salaried worker share, patents per capita, the initial level of coal use, latitude, air frost days, and rainfall.

As an additional test of the estimation procedure, I conduct a permutation exercise

[40] A caveat to this sensitivity check is that it only applies to variation in industry coal use intensity across locations that is driven by variation in the price of coal relative to wages. This analysis cannot address variation that is driven by other factors, such as stronger regulation in some cities, though this is not likely to be a major factor in the empirical setting I consider.

in which I randomly re-assign the industry coal use per worker values available in the data across industries, construct placebo measures of city coal use, and then apply my estimation procedure to generate placebo estimates of the impact of coal use on city growth. After repeating this procedure 1000 times, I then compare the results estimated using the true industry coal use measures to the set of placebo estimates in order to obtain an alternative assessment of the statistical significance of my results. When no controls are included, as in Column 3 of Table 3, I find that only 1.8% of the placebo coefficients exceed the coefficient estimated with the true coal use intensity measure. This suggests a statistical significance level of 98.2% for the city coal use coefficient in Column 3 of Table 3. When controls are included, as in Column 5 of Table 4, I find that 6.1% of the placebo coefficients exceed the coefficient estimated with the true coal use intensity measure. This suggests a statistical significance level of 93.9%. I also find that the estimated coefficient on $\Delta \ln(PredWP_{ct})$ obtained when using the true industry coal use measures is closer to one, in absolute value, than roughly 90% of the coefficients on this variable obtained in the placebo regressions. This suggests that the specification using the true coal use measure offers a better fit for the data than most of the randomly generated alternatives. Further details on the permutation exercise and additional results are described in Appendix A.3.4.

City coal use may affect city growth either through the response of migrants or through the direct impact of pollution on population through mortality. As a final step in the analysis, I use results from Hanlon (2015) in order to obtain a back-of-the-envelope estimates of the share of the pollution impact that can be explained by increased mortality. Specifically, I consider the impact of a city experiencing growth in coal use that is one s.d. above the average on overall city population, using results from Table 5. This can then be compared the the expected impact of the same growth in coal use on overall mortality using estimates from Hanlon (2015).

Consider a city of 100,000 that experiences growth in employment and coal use at the average rate observed across all cities and periods in the data. After one year, we would expect the city population to be 101,187. However, if instead the city had experienced a growth in coal use that was one s.d. above the average, then the population after one year would be only 100,010 based on the estimated impact of industrial pollution on overall city population from Table 5.[41] Thus, the additional coal use growth reduces population growth by 1,117 people over the course of one year. The estimates from Hanlon (2015) suggest that this additional growth in coal use would have increased mortality by 184 people per year.[42] Thus, the impact of coal use on mortality can explain 15.7% of the reduction in population growth in the city that experienced more rapid growth in coal use. This suggests that, while the direct mortality impact of pollution was important, the majority of the population response was due to other channels, the most important of which must have been the decisions made by migrants in response to the local pollution disamenity.

7 Conclusion

Standard urban economics models predict that features that make cities less desirable places to live – such as high levels of pollution – will negatively impact city population. Pollution is a particularly important disamenity to consider because, unlike disamenities such as bad weather, it is endogenous and therefore can be influenced by the decisions of policy makers. However, the endogenous nature of this disamenity also makes identifying the causal effect on city population more difficult. The main contribution of this study is to provide a framework for isolating the negative impact of the disamenity generated by industrial pollution and then apply it to show that industrial pollution had a large effect on population growth in the industrial cities of

[41] One standard deviation is equal to a change in log coal use of 0.178. I need to work with city population here in order to compare to the mortality results, so I multiply this by the estimated effect of coal use on city population from Column 4 of Table 5.

[42] The estimated impact of a change in coal use on mortality, from Column 2 of Table 10 in Hanlon (2015), is 10.36. This will include both the direct impact of coal use on mortality and the increase in mortality due to the selection of less healthy populations into more polluted areas. Multiplied by an increase in coal use of 0.178, this implies 1.84 additional deaths per thousand.

19th century Britain.

My preferred results suggest that a city with an increase in industrial pollution that was one standard deviation greater than the national average would experience annual city employment growth that was 2 percentage points lower than the national average. Given that the average growth in city employment over the period I study was 1.5%, this is a substantial effect. In interpreting these estimates, it is important to keep in mind that they do not include general equilibrium effects of rising city pollution levels on the outside option of workers across all cities. Thus, they can reflect the impact of increasing coal use in an individual city, but they should not be extrapolated to assess the impact of rising coal use across all British cities during this period. Analyzing these general equilibrium impacts is one direction for future research. I also provide evidence that most of this effect of industrial pollution – roughly 85% – was due to factors other than the direct mortality impact of pollution, such as the decisions of migrants flowing into and out of the cities.

One advantage of the approach I propose is that it can be applied in other settings, such as emerging industrial economies, where pollution levels are often much higher than in cities in more developed countries. In these settings, pollution measures are often unavailable but policymakers must grapple with the tradeoffs between employment growth and increased pollution. Applying the approach introduced here to these settings is another promising direction for future research.

References

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

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

Baines, Dudley. 1994. Population, Migration and Regional Development, 1870-1939. *In:* Floud, R, & McCloskey, D (eds), *The Economic History of Britain Since 1700*, second edition edn. Cambridge University Press.

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

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

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.

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

Cameron, Trudy Ann, & McConnaha, Ian T. 2006. Evidence of Environmental Migration. *Land Economics*, **82**(2), 273–290.

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.

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

Clay, Karen, Lewis, Joshua, & Severnini, Edson. 2014 (November). *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.

Currie, Janet, Davis, Lucas, Greenstone, Michael, & Walker, W. Reed. 2014 (April). *Environmental Health Risks and Housing Values: Evidence from 1600 Toxic Plant Openings and Closings.* Mimeo.

Diamond, Rebecca. 2012 (November). *The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000*. Working paper.

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

Greenstone, Michael, & Gallagher, Justin. 2008. Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program. *The Quarterly Journal of Economics*, **123**(3), pp. 951–1003.

Hanlon, W. Walker. 2014 (September). *Temporary Shocks and Persistent Effects in the Urban System: Evidence from British Cities after the U.S. Civil War*. NBER Working Paper No. 20471.

Hanlon, W.W. 2015 (March). *Pollution and Mortality in the 19th Century*. Working paper.

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

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.

Jones, Ronald W. 1975. Income Distribution and Effective Protection in a Multicommodity Trade Model. *Journal of Economic Theory*, **11**(1), 1 – 15.

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

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

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

Mastromonaco, Ralph. Forthcoming. Do Environmental Right-to-Know Laws Affect Markets? Capitalization of Information in the Toxic Release Inventory. *Journal of Environmental Economics and Management*.

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

Moretti, E. 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.

Pope III, CA, Burnett, Richard T, Turner, Michelle C, Cohen, Aaron, Krewski, Daniel, Jerrett, Michael, Gapstur, Susan M, & Thun, Michael J. 2011. Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure-response relationships. *Environmental Health Perspectives*, **119**(11), 1616–1621.

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

Roback, Jennifer. 1982. Wages, Rents, and the Quality of Life. *The Journal of Political Economy*, **90**(6), pp. 1257–1278.

Rosen, Sherwin. 1979. Wage-Based Index of Urban Quality of Life. *In:* Mieszkowski, Peter, & Straszheim, Mahlon (eds), *Current Issues in Urban Economics*. Baltimore and London: Johns Hopkins University Press.

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.

Sieg, Holger, Smith, V Kerry, Banzhaf, H Spencer, & Walsh, Randy. 2004. Estimating the general equilibrium benefits of large changes in spatially delineated public goods. *International Economic Review*, **45**(4), 1047–1077.

Smith, V. Kerry, & Huang, Ju-Chin. 1995. Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models. *Journal of Political Economy*, **103**(1), pp. 209–227.

Smith, V Kerry, Sieg, Holger, Banzhaf, H Spencer, & Walsh, Randall P. 2004. General equilibrium benefits for environmental improvements: projected ozone reductions under EPA's Prospective Analysis for the Los Angeles air basin. *Journal of Environmental Economics and Management*, **47**(3), 559–584.

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.

A Appendix

A.1 Empirical setting appendix

A.1.1 Further details on the empirical setting

Figure 2: English cities included in the study

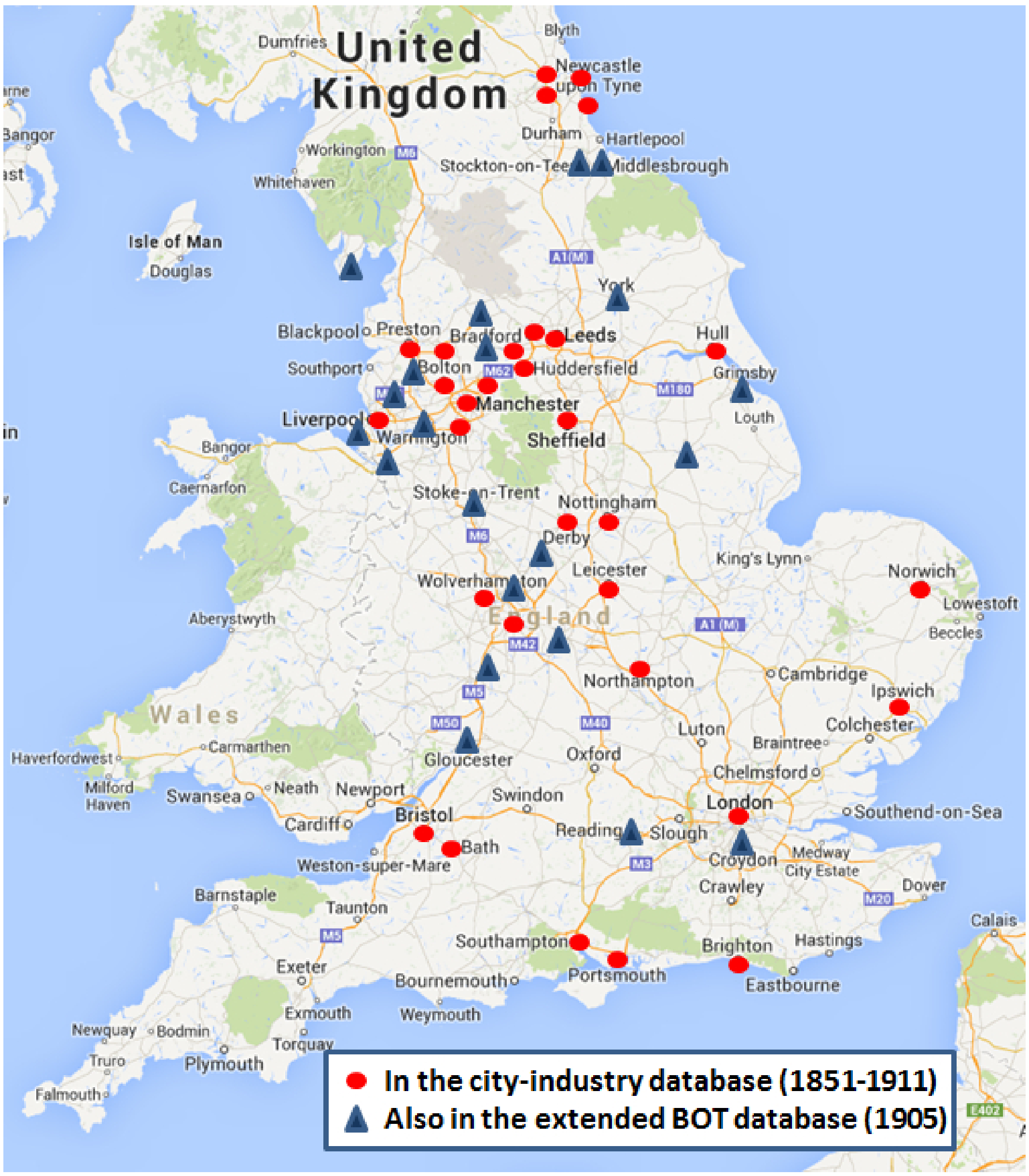


Table 7: History of British air pollution regulation, 1851-1911

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 and measurement 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 burnt 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 burnt, 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 burnt 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 burnt 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.

In addition to data on coal use, the Census of Production provides counts of the number of wage and salaried workers in each industry. These are used to construct measures of the salaried worker share by industry.

A.2.2 Board of Trade data

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

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

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

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

A.2.3 Constructing additional control variables

One threat to identification in this study is the possibility that there may be other industry features that vary across industries in a way that is correlated with industry coal use and impact overall city size.[43] In particular, recent research suggests that the share of skilled workers in a city can have a substantial effect on city growth, either through productivity or amenity spillovers. If coal using industries use either more or less skilled labor than non-polluting industries then these effects may bias the estimated impact of coal use on city size.

One way to address this concern is to draw on additional data describing other industry features. In particular, I use information from the 1907 Census of Production giving the share of salaried (higher skilled) workers in each industry. I then construct a control variable based on the share of salaried workers in city employment:

$$PredSALARIED_{ct} = \frac{\left(\sum_i L_{ict} * GREMP^{-c}_{it,t-\tau} * SALARIED_i\right)}{\sum_i L_{ict} * GREMP^{-c}_{it,t-\tau}}.$$

Here $SALARIED_i$ represents the share of salaried workers in the industry. Note that the construction of this control variable is very similar to the construction of the $PredCOAL_{ct}$ variable. When including this variable in regressions, I always use $\ln(PredSALARIED_{ct})$, following Diamond (2012).

A.2.4 City coal use intensity data

Table 8 describes the level of coal use estimated for each city in each year based on the city's industrial composition and the industry's coal use intensity per worker from the 1907 Census of Production.

[43]A similar concern applies to all studies using a Bartik instrumentation approach.

Table 8: Industrial coal use per manufacturing worker for a selection of cities

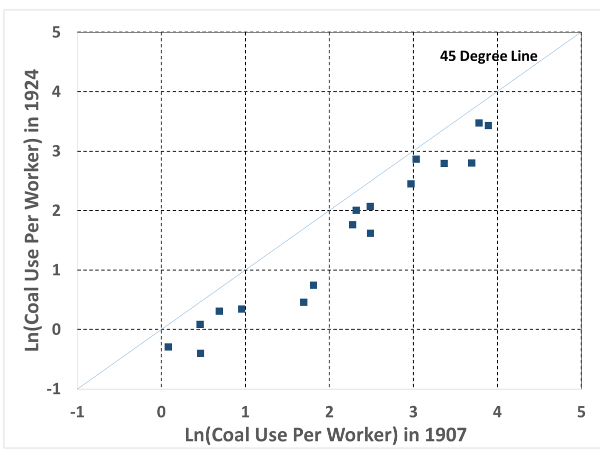
City	1861	1871	1881	1891	1901	1911	Average	Std. Dev.	Growth 1861-1911
Northampton	6.3	6.5	5.1	4.8	5.5	5.7	5.6	0.7	-0.6
Leicester	8.0	8.6	6.9	5.9	7.9	9.3	7.8	1.2	1.3
Norwich	9.5	10.5	10.5	9.9	9.5	9.5	9.9	0.5	0.0
Nottingham	10.2	11.1	12.9	12.9	13.6	13.8	12.4	1.4	3.6
Portsmouth	10.9	11.4	13.2	12.6	13.6	12.8	12.4	1.0	1.9
Brighton	10.9	11.7	12.3	11.9	11.8	11.6	11.7	0.5	0.7
London	11.1	11.6	11.7	11.7	11.6	10.9	11.4	0.3	-0.2
Stockport	11.9	11.7	11.3	10.4	10.6	12.3	11.4	0.8	0.4
Blackburn	12.0	12.0	12.4	12.3	12.7	12.9	12.4	0.4	0.9
Bristol	12.2	12.9	12.0	11.5	11.9	12.1	12.1	0.4	-0.1
Preston	12.4	12.5	12.7	12.3	12.6	12.5	12.5	0.2	0.1
Southampton	12.7	14.1	14.8	13.5	13.4	13.1	13.6	0.8	0.4
Liverpool	12.7	14.5	14.6	13.9	14.0	14.0	14.0	0.7	1.3
Huddersfield	13.0	12.9	13.9	13.3	13.7	13.5	13.4	0.4	0.5
Halifax	13.2	14.9	15.2	15.4	16.3	16.9	15.3	1.3	3.7
Ipswich	13.7	15.2	14.7	14.5	15.7	18.2	15.3	1.6	4.5
Kingston-upon-Hull	13.9	16.5	16.1	16.6	17.6	18.3	16.5	1.5	4.4
Bradford	14.1	13.9	13.6	12.9	13.2	12.9	13.4	0.5	-1.2
Manchester	14.7	15.4	15.3	15.6	15.4	15.5	15.3	0.3	0.8
Derby	15.4	18.3	20.7	20.5	19.7	18.8	18.9	1.9	3.4
Bolton	16.5	16.1	16.1	16.2	16.6	16.4	16.3	0.2	0.0
Oldham	16.5	15.4	16.3	17.7	18.1	17.6	16.9	1.1	1.1
Leeds	17.4	18.6	18.0	16.6	17.1	17.1	17.5	0.7	-0.3
Sunderland	18.8	21.1	20.5	19.0	18.9	18.7	19.5	1.0	-0.1
South Shields	19.3	22.4	22.1	21.3	21.4	21.0	21.2	1.1	1.7
Newcastle	21.5	22.0	20.9	21.9	22.4	21.3	21.7	0.5	-0.2
Birmingham	24.9	25.0	25.6	25.9	24.6	23.1	24.8	1.0	-1.8
Gateshead	29.4	30.4	29.2	28.1	27.3	25.4	28.3	1.8	-4.0
Wolverhampton	29.8	30.5	30.3	28.3	26.7	25.7	28.5	2.0	-4.0
Sheffield	30.9	31.6	31.0	30.6	31.2	32.0	31.2	0.5	1.1
Average	15.5	16.3	16.3	15.9	16.1	16.1			
Std. Dev.	6.3	6.4	6.3	6.3	5.9	5.6			

Values are in tons per worker per year.

A.2.5 Analyzing the change in relative industry coal intensity over time

Figure 4 provides a scatter plot 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.2.6 Comparing estimated coal use to 1871 Coal Commissioner's Report

In this appendix, I check the ability of my methodology to accurately reflect variation in local coal use measures going back in time. To do so, I take advantage of data collected as part of the *Report of the Commissioners Appointed to Inquire into the Several Matters Relating to Coal in the United Kingdom*, which was submitted to the House of Commons in 1871. As part of this report, surveys of industrial coal use were conducted for a selection of counties in England. These surveys were conducted by sending circulars to the major industrial establishments in each county asking for information about coal usage. I digitized the coal usage information reported for each county as part of this report. Because only some circulars were returned, I adjust the coal usage figures to account for non-reporting. I then add up coal use by county.

There are some potential shortcomings with the county-level industrial coal use values based on the 1871 report. Because circulars were sent only to large industrial establishments, they will miss industrial coal use by smaller users (e.g., local blacksmith shops). This is unlikely to be a major issue in the major industrial counties, but it may be a concern in less industrial counties, where smaller users are likely to be relatively more important. Also, to adjust for circulars that were not returned, I am forced to assume that the coal use levels of businesses that returned circulars are identical to those that chose not to respond, which introduces a second margin of error into these calculations. Nevertheless, the data from the 1871 report appear to be the most accurate and comprehensive data available at the sub-national level for the early part of my study period.

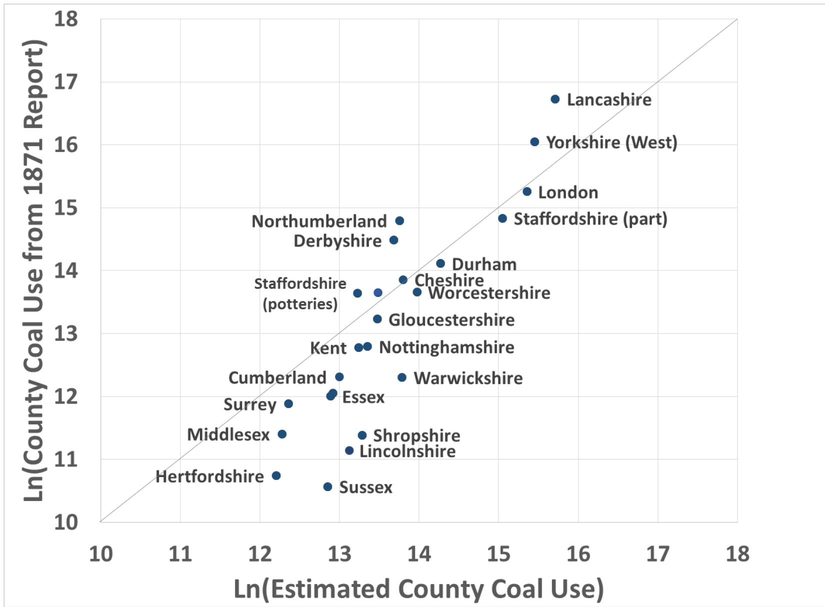
I want to compare the resulting county-level coal use values to values constructed using the methodology described in this paper, i.e., values based on data on local employment by industry from the Census of Population multiplied by industry coal use intensity from the 1907 Census of Production. To construct county-level measures using this approach, I start with the district level occupation data from Hanlon (2015) for 1851. These occupation data are aggregated to the county level. I then use national industry growth rates between 1851 and 1871 to calculated predicted county-industry employment values for 1871. These are interacted with the coal use per worker data from the 1907 Census of Production to generate industry coal use at the county level. I then sum across industries to obtain estimates of county coal use. Finally, I adjust all county level coal use values by my estimate of ψ_t for 1871, which is calculated using national coal use data, as described in the main text. The result is estimated county-level coal use measures for 1871.

Figure 5 provides a scatterplot comparing the county industrial coal use estimated using the methodology introduced in this paper (x-axis) to county industrial coal use from the 1871 Coal Commissioner's Report (y-axis). Both values are in natural logs. This figure suggests that, in general, my estimation approach does a reasonable job of replicating the county-level industrial coal use values calculated using the 1871 Coal Commissioner's Report. This is particularly true for the larger and more industrial counties of the country. It is comforting that the estimation procedure works well for more industrial areas, since the cities studied in this paper will also be more industrial than the rest of the country.

For the smaller and more rural counties, located towards the bottom left of the graph, it appears that my estimation approach is overstating industrial coal use relative to the 1871 report. However, this does not necessarily indicate that my approach is not working for these counties. This is because the values from the 1871 report are likely to be further off of true industrial coal use in less industrialized counties, because they will miss small users, which are likely to be relatively more important in the less industrial areas.

Overall, the reasonably close correspondence between the estimated coal use values at the county level for 1871 and the values calculated using the Coal Commissioner's Report suggests that the methodology I have introduced provides an accurate measure of local industrial coal use, even in the earlier decades of my study period.

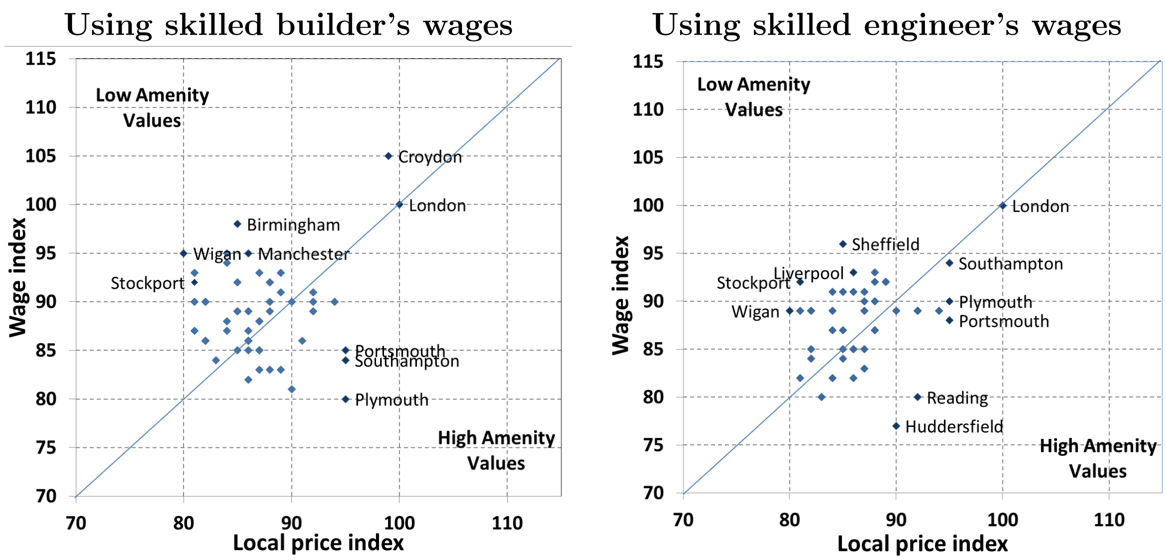
Figure 5: Comparing county industrial coal use in 1871



A.2.7 City quality-of-life estimates

Figure 6 plots the quality-of-life measures. Cities toward the top left on these graphs have high wages relative to their cost of living, indicating low quality-of-life. Cities towards the bottom left have high quality-of-life. In general, these figures are consistent with historical evidence: port cities, such as Portsmouth, Plymouth and Southampton, tended to offer residents higher quality-of-life, while industrial towns were characterized by lower quality-of-life. A couple examples can help illustrate this point. Wigan, one of the cities showing the lowest quality of life in the sample, was a notoriously poor place to live, as described by George Orwell in "The Road to Wigan Pier." In contrast, one of the towns with high quality of life measures is Southampton, which John Choules described in his 1854 book "The Cruise of the Steam Yacht North Star," saying, "I hardly know a town that can show a more beautiful Main Street than Southampton, except it be Oxford...The shops are very elegant, and the streets are kept exceedingly clean."

Figure 6: City amenity graphs



Data from 51 towns

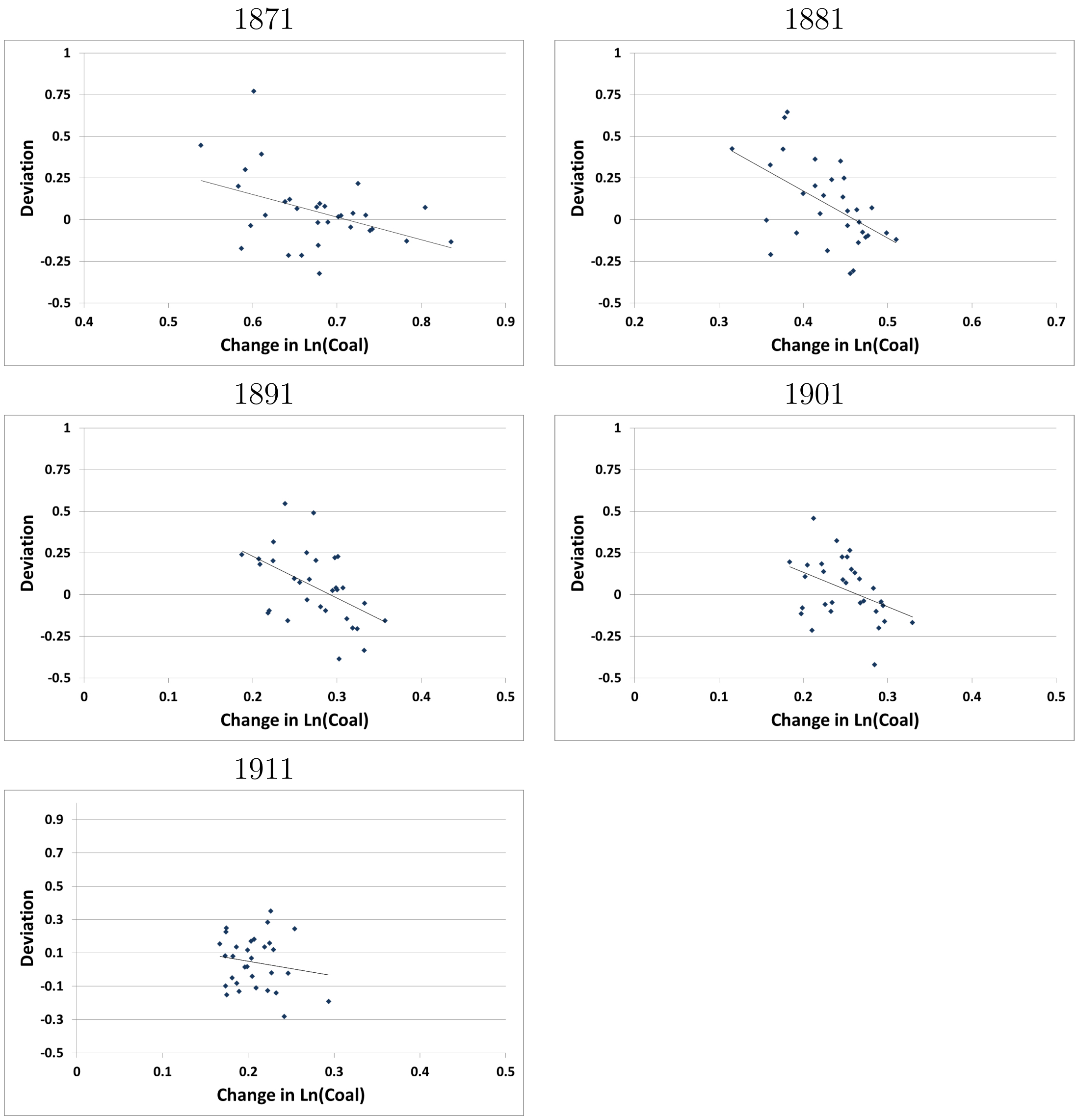
Data from 45 towns

A.3 Analysis appendix

A.3.1 Additional plots of DEVIATION against coal use

Figure 1 in the main text plots the change in log coal use against the deviation of actual city employment growth from predicted city employment growth using all cities and years. Figure 7 provides similar plots for each year separately.

Figure 7: Deviation of actual from predicted growth plotted against change in city coal use



A.3.2 Summary statistics for analysis variables

Table 9 presents summary statistics for the main analysis variables.

Table 9: Summary statistics for variables used in the main analysis

Variable	Obs.	Mean	Std. Dev.	Min	Max
$\Delta \ln(L_{ct})$	155	0.30	0.19	-0.14	0.96
$\Delta \ln(PredWP_{ct})$	155	0.37	0.18	0.17	0.84
$\Delta \ln(PredCOAL_{ct})$	155	0.27	0.07	0.09	0.49
Δ Log salaried share	155	0.02	0.02	-0.03	0.06
Patents per capita	155	0.01	0.00	0.00	0.02
Rainfall	155	0.81	0.19	0.56	1.29
Air frost days	155	39.64	9.97	22.70	56.00
Patents per manuf. worker	155	53.04	1.19	50.82	55.00

A.3.3 Additional regression results

Table 10 presents cross-sectional regression results for each decade. Because I have so few observations, these results are estimated without including the additional control variables. The estimated coefficients suggest that the impact of industrial coal use on city growth followed an inverted U-shaped pattern, with the effect peaking in the 1891-1901 period. However, with only 31 observations in each regression, the results are not estimated with enough precision to be statistically significant. I have also calculated similar results using one-decade lags to construct the predicted variables and those results exhibit a similar inverted U-shaped pattern.

Table 10: Cross-sectional results by decade

	DV: Δ Log of city working population in analysis industries				
Observation year:	1871	1881	1891	1901	1911
Base year:	1851	1861	1871	1881	1891
$\Delta \ln(PredWP_{ct})$	0.683	0.965	2.522	0.0136	-0.00465
	(1.651)	(3.114)	(2.659)	(1.941)	(0.783)
$\Delta \ln(PredCOAL_{ct})$	-0.596	-3.378	-4.856	-0.895	0.0853
	(2.027)	(3.707)	(3.264)	(2.826)	(1.935)
Constant	0.468	1.615	1.006***	0.565*	0.159
	(1.029)	(1.128)	(0.333)	(0.300)	(0.482)
Observations	31	31	31	31	31
R-squared	0.218	0.329	0.451	0.253	0.138

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. Values from two decades before each year are used as the basis for constructing $\Delta \ln(PredWP_{ct})$ and $\Delta \ln(PredCOAL_{ct})$.

While the results in the main text are generated using a two-decade lag between the base and observation year, we can also consider alternative lag lengths. This is done in Table 11, which considers pooled estimates for lag lengths of one to three decades. For comparability, all of the regressions are run using outcome variables starting from 1881, though similar patterns emerge if I use all available years for each lag length. Effects are generally larger when considered over a two decade period.

Table 11: Results with alternative lag lengths for predicted variables

	DV: Δ Log of city working population in analysis industries					
Lag length:	One decade		Two decades		Three decades	
	(1)	(2)	(3)	(4)	(5)	(6)
$\Delta \ln(PredWP_{ct})$	0.874	-0.632	1.294*	-0.166	1.469	-0.748
	(0.676)	(0.762)	(0.745)	(0.748)	(0.939)	(1.032)
$\Delta \ln(PredCOAL_{ct})$	-1.618	-1.767*	-2.656**	-3.659***	-2.406**	-1.380
	(1.051)	(1.017)	(1.103)	(1.329)	(1.186)	(1.403)
Year effects	Yes	Yes	Yes	Yes	Yes	Yes
Region effects		Yes		Yes		Yes
Additional controls		Yes		Yes		Yes
Observations	124	124	124	124	124	124

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. The data cover 1881-1911. The additional control variables included are latitude, air frost days, rainfall, and the log change in the city salaried worker share.

Tables 12 and 13 describe, respectively, results obtained when regressions are weighted by the log of city population and by the level of city population, respectively. In general, these weighted results deliver an estimated coal effect that is very similar to the results obtained from unweighted regressions.

Table 12: Regression results weighted by the log of city population

	DV: Δ Log of city working population in analysis industries				
	(1)	(2)	(3)	(4)	(5)
$\Delta \ln(PredWP_{ct})$	0.784	0.459	0.799	0.601	0.549
	(0.624)	(0.615)	(0.619)	(0.649)	(0.666)
$\Delta \ln(PredCOAL_{ct})$	-1.817**	-1.737**	-1.926**	-2.040**	-1.998**
	(0.870)	(0.829)	(0.934)	(0.968)	(0.936)
Patents per capita		13.01**			16.66***
		(6.067)			(5.961)
Initial coal use					-0.0495*
					(0.0257)
Δ Log salaried share			-0.372		-0.644
			(1.580)		(1.544)
Latitude				-0.0440	-0.0109
				(0.0443)	(0.0415)
Air frost days				-0.00351	-0.00411
				(0.00299)	(0.00312)
Rainfall				-0.0747	-0.0374
				(0.159)	(0.165)
Constant	1.302***	1.140**	1.383**	3.941	2.681
	(0.492)	(0.462)	(0.568)	(2.550)	(2.378)
Year effects	Yes	Yes	Yes	Yes	Yes
Region effects	Yes	Yes	Yes	Yes	Yes
Observations	155	155	155	155	155
R-squared	0.192	0.221	0.192	0.214	0.272

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. The data cover 1871-1911, with values from two decades before each year used as the basis for constructing the $\Delta \ln(PredWP_{ct})$, $\Delta \ln(PredCOAL_{ct})$, as well as the city salaried worker share.

Table 13: Regression results weighted by the level of city population

	DV: Δ Log of city working population in analysis industries				
	(1)	(2)	(3)	(4)	(5)
$\Delta \ln(PredWP_{ct})$	0.554	0.566	0.530	0.445	0.602
	(0.568)	(0.585)	(0.537)	(0.573)	(0.566)
$\Delta \ln(PredCOAL_{ct})$	-1.579**	-1.575**	-2.043**	-1.651**	-1.966**
	(0.769)	(0.773)	(0.842)	(0.804)	(0.846)
Patents per capita		-0.594			4.965
		(6.557)			(7.125)
Initial coal use					-0.0391
					(0.0249)
Δ Log salaried share			-2.028		-2.069
			(1.519)		(1.543)
Latitude				-0.0212	-0.00751
				(0.0538)	(0.0521)
Air frost days				-0.00202	-0.000970
				(0.00320)	(0.00311)
Rainfall				0.00410	0.0466
				(0.172)	(0.177)
Constant	1.209***	1.211***	1.586***	2.447	2.390
	(0.448)	(0.451)	(0.546)	(2.873)	(2.803)
Year effects	Yes	Yes	Yes	Yes	Yes
Region effects	Yes	Yes	Yes	Yes	Yes
Observations	155	155	155	155	155
R-squared	0.269	0.269	0.280	0.275	0.297

*** p<0.01, ** p<0.05, * p<0.1. Robust standard errors in parentheses. The data cover 1871-1911, with values from two decades before each year used as the basis for constructing the $\Delta \ln(PredWP_{ct})$, $\Delta \ln(PredCOAL_{ct})$, as well as the city salaried worker share.

A.3.4 Permutation exercise

As a final test of the empirical estimation procedure, I conduct a permutation exercise. In the first step of this exercise, I randomly re-assign the industry coal use per worker values available in the data across industries. These are then used to construct city coal use levels following the same procedure used in the main analysis. For each set of city coal use levels, I then run one regression without control variables (as in Column 3 of Table 3) and a second regression with the full set of available control variables (as in Column 5 of Table 4). There are sixteen industry coal use per worker levels observed in the data, which can be assigned across industries in $2 \times E^{13}$ potential ways. Out of these, I randomly select 1000 permutations assigning coal use per worker values to industries. Thus, I am left with 1000 sets of regression results without including controls and another 1000 results when the full set of available controls are included. For each regression, I collect the estimated coefficients and p-values for both the $\Delta \ln(PredWP_{ct})$ and $\Delta \ln(PredCOAL_{ct})$ terms. This allows me to evaluate my estimation procedure in several ways.

One outcome of interest from the permutation exercise is the share of results, out of the 1000 regressions, that deliver a statistically significant regression coefficient on the city coal use term.[44] When no controls are included, I find that 13.9% of the regressions deliver statistically significant coefficients on city coal use. With controls, 16.6% of the regression deliver statistically significant results. This suggests that the standard errors presented in the main text may lead me to slightly overstate the statistical significance of the effect of coal use on city growth.

A second outcome of interest from the permutation exercise is the share of results, out of the 1000 regressions, that deliver an estimated coefficient that is more negative than the estimates obtained when using the true industry coal use intensity values. Comparing the estimates from the results tables in the main text to the distribution of permutation coefficients provides an alternative method for determining the statistical significance of the results. When no controls are included, I find that only 1.8% of the coefficients from the permutation exercise exceed the coefficient estimated with the true coal use intensity measure in Table 3. This suggests a statistical significance level of 98.2% for the city coal use coefficient in Column 3 of Table 3. When controls are included, I find that 6.1% of the permutation coefficients exceed the coefficient estimated with the true coal use intensity measure in Table 3. This suggests a statistical significance level of 93.9% for the city coal use coefficient in Column 5 of Table 4. These alternative statistical significance measures provide some additional confidence

[44]As in the main text, statistical significance is based on robust standard errors, which tend to deliver larger confidence intervals than standard errors that allow for spatial and serial correlation.

that the estimated negative effects of city coal use on city growth are unlikely to be randomly generated.

A third outcome of interest from the permutation exercise involves the coefficient on the $\Delta \ln(PredWP_{ct})$ term. In particular, I am interested in the share of the coefficients from the placebo regressions that deliver a coefficient on this variable that is closer to one, in absolute value, than the coefficient obtained when using the true coal use intensity measure. Since the proximity of the coefficient on $\Delta \ln(PredWP_{ct})$ to one can be thought of as a test of the quality of the regression specification, this amounts to asking what fraction of the placebo regressions appear to be as good or better as the specification that uses the true coal use intensity values. When no controls are used, I find that 10.6% of the placebo regressions deliver a coefficient on $\Delta \ln(PredWP_{ct})$ that is closer to one in absolute value than that obtained using the true data. The share is 7.4% when controls are included. This suggests that true coal use measures provide regression specification that fits the data better than around 90% of the randomly generated alternatives.