THE ECONOMIC JOURNAL

The Economic Journal, Doi: 10.1111/ecoj.12522 © 2017 Royal Economic Society. Published by John Wiley & Sons, 9600 Garsington Road, Oxford OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

COAL SMOKE AND MORTALITY IN AN EARLY INDUSTRIAL ECONOMY*

Brian Beach and W. Walker Hanlon

Air pollution was severe in the nineteenth century, yet its health consequences are often overlooked due to a lack of pollution data. We offer a new approach for inferring local coal use levels based on local industrial structure and industry-specific coal use intensity. This allows us to provide the first estimates of the mortality effects of British industrial coal use in 1851–60. Exploiting wind patterns for identification, we find that a one standard deviation increase in coal use raised infant mortality by 6–8% and that industrial coal use explains roughly one-third of the urban mortality penalty observed during this period.

Cities were incredibly unhealthy places to live during the nineteenth century. Figure 1 illustrates this point by plotting infant mortality rates (top panel) and age-standardised mortality rates (bottom panel) against population density for registration districts in nineteenth century England and Wales. The strong positive relationship observed in both panels indicates that increasing population density was associated with large increases in mortality. This relationship, which is often referred to as the urban mortality penalty, is a common feature of industrial nations during this period (Cain and Hong, 2009; Kesztenbaum and Rosenthal, 2011).

The literature offers two main explanations for the causes of the urban mortality penalty. The first emphasises the role of infectious diseases, particularly those associated with unclean water and improper sewage disposal (Troesken, 2002; Cutler and Miller, 2005; Ferrie and Troesken, 2008; Kesztenbaum and Rosenthal, 2011, 2013; Alsan and Goldin, 2014; Antman, 2016). Others have suggested that poor nutrition may have also been important (McKeown, 1976; Fogel and Costa, 1997; Fogel, 2004). The role of air pollution, however, is curiously absent from this literature despite contemporary reports making it clear that air pollution, particularly that which resulted from the burning of coal, was a severe problem.

Why has the literature largely ignored the impact of air pollution on health in the nineteenth and early twentieth centuries? The main explanation is that direct pollution measures are almost entirely unavailable before the mid-twentieth

* Corresponding author: W. Walker Hanlon, UCLA, 8283 Bunch Hall, 405 Hilgard Ave., Los Angeles, CA 90095, USA. Email: whanlon@econ.ucla.edu.

For helpful suggestions we thank Ran Abramitzky, Marcella Alsan, David Atkin, Leah Boustan, Karen Clay, Dora Costa, Dave Donaldson, Roger Foquet, Claudia Goldin, Philip Hoffman, Matt Kahn, Larry Katz, Adriana Lleras-Muney, Steven Nafziger, Jean-Laurent Rosenthal, Werner Troesken and seminar participants at Caltech, UCLA, the University of Southern Denmark, the Grantham Institute at LSE, the SITE Conference, the NBER Summer Institute, and the EHA Annual Conference. Reed Douglas provided excellent research assistance. We thank the UCLA Ziman Center for Real Estate and the California Center for Population Research for generous funding. This article was previously circulated under the title 'Pollution and Mortality in the 19th Century'.

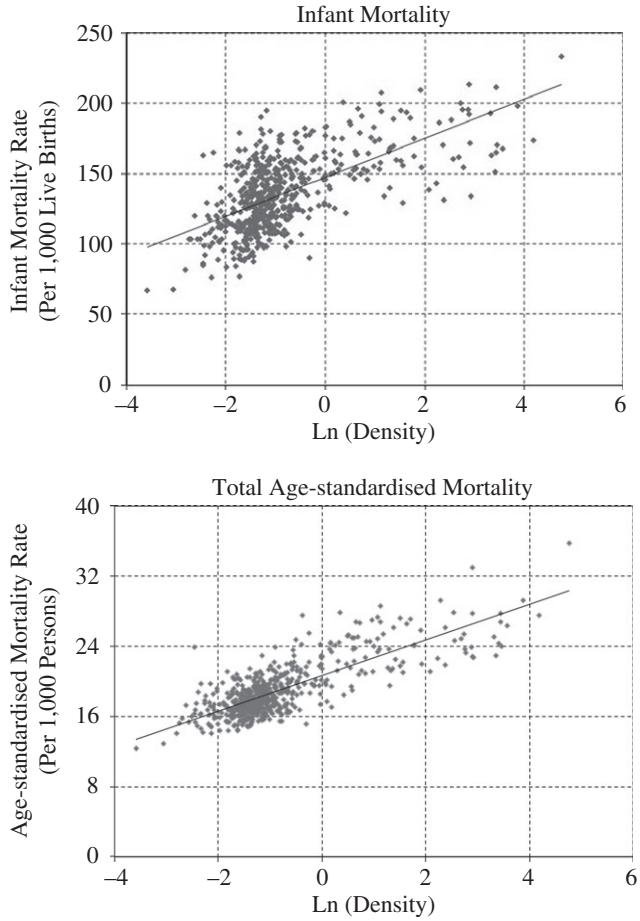

Fig. 1. *Urban Mortality Penalty in England and Wales in 1851–60*
Notes. Mortality data for 581 registration district covering all of England and Wales from 1851–60 were digitised by Woods (1997) from reports produced by the Registrar General's office to district population density in 1851 based on data from the 1851 Census of Population.

century, which makes it difficult to assess the impact of pollution on health quantitatively.[1] To overcome this challenge, we introduce an approach that allows us to infer local industrial pollution levels. Specifically, we combine data describing the industrial structure of 581 districts covering all of England and Wales with information on industry coal use per worker. Together, these two

[1] While there is a larger qualitative literature on this topic (Brimblecombe, 1987; Mosley, 2001; Thorsheim, 2006), estimates of the relationship between pollution and mortality largely begin in the middle of the twentieth century (Barreca *et al.*, 2014; Clay *et al.*, 2015, 2016). A number of studies investigate the health impacts of particular pollution events in the twentieth century (Townsend 1950; Logan, 1953; Greenburg *et al.*, 1962; Ball, 2015). One nineteenth century exception to this literature is Troesken and Clay (2011), who study the evolution of mortality following large fog events. Relative to previous work, our study offers a new identification approach and provides evidence for an earlier period and a broader set of locations.

pieces of information allow us to estimate the spatial distribution of industrial coal use in 1851–60, the earliest decade for which all of the necessary data are available. These estimates are then paired with mortality data in order to analyse the impact of coal use on health.

Our baseline estimates suggest that a one standard deviation (SD) increase in log coal use raised infant mortality by 10.7–13.7 deaths per 1,000 live births, or about 8–10%. Moreover, we show industrial coal use explains roughly one-third of the urban mortality penalty for infants. Of course, we may worry that these simple cross-sectional results are subject to bias from omitted variables or population sorting. To address these concerns, we exploit variation in wind patterns to identify how coal use in one area affects mortality in neighbouring downwind and upwind districts.

These results indicate that increasing industrial coal use in upwind districts by 1 SD raised infant mortality rates in downwind districts by 2.3–3.0 deaths per 1,000 live births, or 1.7–2.2%. We also provide evidence that the mortality effects of coal use were stronger in hilly areas where air pollution was likely to be trapped in more densely inhabited valleys, which provides an alternative check on our results. In another check, we run a series of placebo tests to show that district-level characteristics and deaths from non-pollution-related causes of death are not affected by upwind pollution exposure.

As a final check, we show how results generated by comparing upwind and downwind districts can be used to evaluate the magnitude of bias present when comparing coal use to mortality within a district. We start by generating predicted pollution diffusion plumes using modern pollution modelling data.[2] These predicted diffusion patterns allow us to scale up our estimates based on upwind and downwind exposure in order to predict the impact of the same amount of coal use on mortality within the source district. This exercise predicts that raising local industrial coal by 1 SD increases infant mortality by 8.12–10.82 deaths per 1,000 live births, or 6.0–8.0%. These values are just below the estimates obtained in naive regressions comparing local industrial coal use directly to infant mortality, which suggests that any bias in our baseline estimates is likely to be small.

Finally, we apply the same approach to study the relationship between local industrial coal use and mortality across all age categories. These results suggest that most of the effects of coal use were concentrated among children under five, though there is also evidence that coal use led to smaller mortality increases among adults. In terms of life expectancy, the estimated impact of coal use on infant mortality alone implies that a 1 SD increase in local industrial coal use lowered life expectancy at birth by 0.33–0.56 years or 0.86–1.44%. If instead we look at the impact on all children under five, the same increase in coal use is associated with a decrease in life expectancy at birth of 0.84–1.58 years or 2.19–4.11%.

These results complement the existing pollution literature by quantifying the health effects of pollution in an environment characterised by high levels of pollution without

[2] Specifically, we follow Heblich *et al.* (2016) in using the ADMS-5 software provided by Cambridge Environmental Research Consultants. More information about this software is available at http://www.cerc.co.uk/environmental-software/ADMS-model.html.

access to modern medical services.[3] The results also shed new light on the magnitude of the impact of pollution in early industrialising economies and open the door to further studies on the long-run impacts of local pollution. A growing literature has shown that historical environmental conditions can play an important role in shaping economic development.[4] By documenting the importance of historical air pollution, our article provides a foundation for further work assessing the economic and developmental consequences of coal-based pollution. In this vein, Heblich *et al.* (2016) document the sorting response to pollution within 19th century British cities and show that these sorting patterns persist today. In another recent study, Hanlon (2016) provides evidence that local industrial coal use had a substantial negative effect on long-run city growth in Britain in the nineteenth and early twentieth centuries. By documenting the mortality impact of coal-based pollution, this article provides support for both of these studies, while their results show that the pollution effects that we document continue to influence modern conditions.

1. Empirical Setting

In England and Wales, air pollution was a problem reaching back at least to the seventeenth century, when Evelyn published his Fumifugium (1661) decrying the smoke of London. Most of this pollution was generated by the burning of coal, the main source of power in Britain during the nineteenth century. Two features of this period exacerbated air pollution. First, coal was often burned inefficiently, resulting in incomplete combustion, which released additional particulates into the air.[5] Second, prior to electrification, coal had to be burned on-site at factories, which tended to be in urban areas where they could be reached on foot by large numbers of workers.

In 1854, an estimated 58 million tons of coal were consumed in England and Wales. Industry accounted for 65% of this coal consumption (Mitchell, 1984, 1988).[6] The intensity of coal use, however, varied enormously across sectors. Together, variation in coal use intensity across industries along with industry agglomeration patterns provide a significant amount of spatial variation in coal use (and thus air pollution), which we exploit for our analysis.

[3] Seminal examples highlighting substantial health impacts of pollution, even at the relatively low concentrations observed in modern developed countries, include Chay and Greenstone (2003) and Currie and Neidell (2005). For recent reviews see Currie (2013) and Graff Zivin and Neidell (2013).

[4] One example is Bleakley (2007), which showed that hookworm eradication in the American south in the early twentieth century had important effects on human capital development. Beach *et al.* (2016) show that the use of water purification technologies to reduce typhoid had similar benefits for human capital. Another example is provided by Almond (2006), which documents the long-term effects of the 1918 influenza pandemic. Similarly Isen *et al.* (2017) use pollution reduction caused by the 1970 Clean Air Act to show that early-life pollution exposure can affect adult income.

[5] For example, the 1871 Coal Commission report (p. 135) describes the inefficient stoking of coal-powered boilers, stating that, 'The careless and wasteful manner of stoking in most of the coal-producing districts is not only a source of vast waste, but of extreme annoyance to all the surrounding neighbourhood. Coal is piled upon the fire without any discretion, producing dense volumes of the blackest smoke, which is so much fuel actually thrown away; nor is the waste the worst part of it; vegetation is destroyed, or seriously injured, for miles, and that which acts so seriously on the plant cannot fail to be injurious to man'.

[6] Figure A2 in online Appendix A displays the allocation of coal consumption in nineteenth century England and Wales.

Though nineteenth century contemporaries did not fully understand the consequences of exposure to air pollution, they did associate pollution with negative health effects, particularly respiratory diseases. One health official described the situation in Manchester as follows:

> Coal smoke forms a continual dark and dense canopy over the town, and causes a murkiness in the streets from which they are never free ... The constant inhalation of these black particles ... must be highly irritating to the lungs.
>
> (*The Times*, 17 May 1866, p. 11)

This report is typical of descriptions of air pollution during in this period.[7] Beyond the health effects, coal smoke in industrial cities was so intense that light-coloured clothes went out of fashion and the lack of visibility could cause traffic accidents.

Today, we know that burning coal releases soot and other particulate matter that can increase mortality. As noted by the US Environmental Protection Agency, the release of small particulate matter (less than 10 micrometres in diameter) is especially dangerous as the matter can get deep inside of a person's lungs or even their bloodstream.[8] Consistent with this, several studies have documented a link between exposure to particulate matter and increased risk for respiratory and cardiovascular disease.[9] The coal combustion process also releases a variety of other chemicals, such as sulphur dioxide, carbon dioxide and nitrogen oxides, and metallic pollution such as mercury, lead and cadmium. These pollutants are associated with a wide variety of negative health effects, particularly on the respiratory and cardiovascular systems.[10]

Despite high levels of pollution, regulation was in its infancy in the middle of the nineteenth century. While the idea of regulating pollution emerged in the 1860s, with provisions for the reduction of excess smoke included in the Sanitary Act of 1866 and the Public Health Act of 1875, historical sources suggest that these measures had limited effectiveness until the beginning of the twentieth century (Thorsheim, 2006; Fouquet, 2012). Pollution regulation faced an uphill battle against the influence of local industrialists and the strong *laissez faire* ideology that dominated British policy-making during this period.

In the absence of regulation, those who could afford to protected themselves from pollution by sorting into neighbourhoods that were upwind of pollution sources, as documented by Heblich *et al.* (2016). However, while sorting took place across neighbourhoods within cities, the ability of most people to avoid pollution by sorting

[7] See Thorsheim (2006) for many more examples.
[8] See https://www.epa.gov/pm-pollution.
[9] See Pope *et al.* (2009) on the relationship between air pollution and life expectancy. For infant health, see Chay and Greenstone (2003), Currie and Neidell (2005), Currie *et al.* (2009), Arceo *et al.* (2016) and Knittel *et al.* (2016). Currie (2013) provides a review of studies examining the impact of pollution on mortality within the economics literature.
[10] Rückerl *et al.* (2011) provides an excellent review of the epidemiological literature on the effects of short and long-term exposure to air pollution. As for the underlying mechanism of these effects, recent research indicates that the channel is due to the pro-inflammatory effects of particulates as well as their activation of stress-signalling pathways (see discussion in Brunekreef and Holgate, 2002).

across cities was constrained by the high cost of commuting during this period and the need for many different types of workers to be present in cities.[11] Given this, and the fact that our analysis is conducted at the district level (which is generally larger than cities), we expect sorting to have only a modest effect on the relationship between coal use and mortality in our study.

2. Data and Measurement

2.1. *Mortality Data*

Our mortality data come from the reports of the Registrar General. The data were originally collected as part of an extensive system aimed at registering every birth, marriage, and death in England and Wales.[12] Although annual reports are available, we draw primarily on the age and cause tabulated data from the 1851–61 decennial supplement, which were digitised by Woods (1997) and are available through the UK Data Archive. Decadal averages are desirable in this setting because the presence of epidemic diseases can introduce substantial noise into annual mortality rates. The data are available at the district level for over 600 districts covering all of England and Wales. For consistency, we combine a small number of districts that experienced border changes between 1851 and 1861. We also treat London as a single 'super-district' by collapsing the many small districts within the traditional borders of London.[13] As a result of these adjustments, our main analysis covers 581 districts.

Our primary outcome variable is infant mortality, which we measure as deaths per 1,000 births. When we study mortality across all age categories, the population denominators are based on average population figures from the 1851 and 1861 censuses.[14] To eliminate the potential concern that polluting industries may have had more on-the-job accidents, which could influence adult mortality rates, we exclude deaths from accidents or violence from our analysis.[15]

[11] Consistent with this, Hanlon (2016) finds that the urban wage premium associated with increased local industrial coal use, using data from 1905, was quite small. Further, in their analysis of mobility in England and Wales between 1851 and 1881, Long and Ferrie (2003) and Long and Ferrie (2013) find that fewer than 27% of migrants crossed county borders. The mean migration distance was only 35 miles, while 25% of all moves were less than 5 miles. These results, combined with the fact that the average registration district radius was just over 5.5 miles suggest that most sorting likely occurred within rather than across districts.

[12] Of the data collected by the Registrar's office, those on mortality are considered to be the most accurate and comprehensive (Woods, 2000). This is because, for every death, registration was required within five days before the body could be legally disposed of.

[13] Because many of the London districts were small, workers often lived and worked in different districts, which undermines our ability to infer local industrial coal use from occupation data for these small districts. Also, because these districts were so small, much of the coal use in one area likely impacted residents in other nearby areas. Working with a single London district avoids these issues and results in a district that is more similar to other districts in the data in terms of area. In the online Appendices we show that our results are robust to excluding London from the analysis.

[14] We may be concerned that these values may not do a good job of capturing district population if there are flows into and out of districts in between the census years. To check this, we digitise the 1851 to 1853 annual reports. In the online Appendices, we present additional results using only deaths between 1851 and 1853 with population levels from the 1851 census used to calculate mortality rates. Analysis on this narrower window, where inaccurate denominators should pose less of an issue, produces similar results.

[15] Note, however, that our results will still capture the impact of on-the-job pollution exposure, which will show up in causes of death other than accidents and violence. This reflects a potentially important channel through which pollution increases mortality that we want to capture.

While the age-tabulated data provide the foundation of our analysis, we also draw on the cause-of-death data to generate a control variable that helps address concerns about the sorting of poorer or less healthy populations into more polluted areas. To construct this variable, we look for diseases satisfying three criteria. First, they should not be among the causes-of-death typically associated with air pollution exposure.[16] Second, the diseases should be related to patterns of poverty and poor living conditions. Third, these causes of death should have clearly identifiable symptoms that would have been well known in the nineteenth century.[17] Based on these criteria, we construct a control variable, which we label 'Child NPR mortality' (as in 'not pollution related'), that includes infant deaths from: cholera, diarrhoea, dysentery and other digestive disorders, diphtheria, smallpox, scarlet fever and typhus. Additional information on these diseases as well as a discussion of the accuracy of the cause-of-death information is available in online Appendices.

2.2. *Measuring Local Industrial Coal Use*

To measure local industrial coal use, we begin by calculating district-level employment by industry. Specifically, we follow Hanlon and Miscio (2017) and collapse occupation data from the 1851 census into 26 distinct categories, covering nearly the entire private sector economy. We focus on 1851 as it is the earliest census for which reliable occupation data are reported.[18] These industry employment tabulations are then paired with data describing the coal use intensity of each industry. We use coal use per worker estimates from the first Census of Manufacturers as our measure of coal use intensity.[19] Because this census was taken in 1907, we conduct several checks, discussed below, to assess whether they allow us to accurately estimate local industrial coal use levels in the 1850s.

These two pieces of data allow us to model coal use at the district level in a particular year t in the following way:

$$\text{COAL}_{dt} = \rho_t \sum_i \theta_i L_{idt}, \tag{1}$$

where L_{idt} denotes local employment in industry i in district d and year t and θ_i denotes the coal use intensity in that industry. We also include a time-varying term representing efficiency gains in coal use, which we denote ρ_t. This

[16] We also exclude infectious diseases that affected the respiratory system, such as measles and whooping cough, where the effects could have been exacerbated by air pollution exposure.

[17] We exclude tuberculosis, an important infectious disease, because it was primarily a disease of the respiratory system and also because it was not well diagnosed during this period.

[18] Occupation data was gathered in the 1841 census as well, but concerns have been raised about the accuracy of the 1841 data since it was the General Register Office's first experience in gathering occupation data.

[19] Table A2 in online Appendix A provides a full breakdown of coal use intensity by industry. We exclude public utilities from this calculation. Some public utilities, particularly gas, were important coal users and did create local pollution. However, by converting coal into gas, which was then pumped into cities, this industry may have actually decreased pollution in city centres. Thus, these industries are excluded because of their ambiguous effects on local pollution and local health.

procedure follows Hanlon (2016); however, because we use a log specification and only analyse a single time period, the ρ_t component will not affect our results.

This approach relies on two important assumptions. First, that relative coal intensity per worker across industries did not change substantially over time. By comparing industry coal use intensity in 1907 and 1924, Hanlon (2016) shows that the relative coal use intensity of industries tends to be quite stable over time.[20] A recent Chinese report provides further support for this assumption, as their list of heavily polluting industries closely matches our estimates of industry coal use intensity.[21]

Our second assumption is that variation in coal use across locations is driven primarily by local industrial structure and the technological features of industries. To assess the validity of this assumption, Hanlon (2016) compares inferred county-level coal use for 1871 (based on the method described above) to county-level coal use data from the 1871 Coal Commission report and concludes that the above methodology does a good job of reproducing actual industrial coal use. In online Appendix A.5, we implement a similar test by comparing results from regressing county mortality rates in 1871 on predicted coal use to results from regressing mortality on actual coal use. Results from both sets of regressions are very similar, suggesting that our approach can accurately reproduce the coal use–mortality relationship.

2.3. *Additional Control Variables*

Here, we briefly describe the set of controls we collected to help capture other factors that may have influenced mortality. Online Appendix A.3 describes these variables in greater detail. Motivated by contemporary reports that air pollution was particularly severe in hilly locations where the smoke was likely to be trapped, we include each district's mean altitude and standard deviation of altitude (which we refer to as hilliness). Next, we include agricultural suitability, which may have influenced income or nutrition. We also include an indicator for whether each district contained a seaport, and if so, a measure of the tonnage passing through the seaport. These controls are meant to capture the role of international trade in spreading disease. Finally, we include female labour force participation, which was high during this period and likely affected mortality through channels such as reduced breastfeeding or supervision of children. This variable was constructed from occupation reports in the 1851 Census. We have also collected information on medical employment and illiteracy at marriage in each district. Because these may be influenced by pollution levels, we do not include these in our main regressions but we do include them in some robustness results.

[20] This is a particularly strong test given that the period from 1907 to 1924 saw the widespread adoption of electricity as a power source, which disrupted the level of industry coal use. In contrast, there was much less change in the sources of industrial power in the period from 1851 to 1907.

[21] See Hanlon and Tian (2015) for a discussion of those data.

3. Analysis

3.1. *Patterns in the Raw Data*

Figure 2 describes the raw relationship between local industrial coal use and infant mortality (top panel) or mortality across all ages (bottom panel) at the district level.[22]

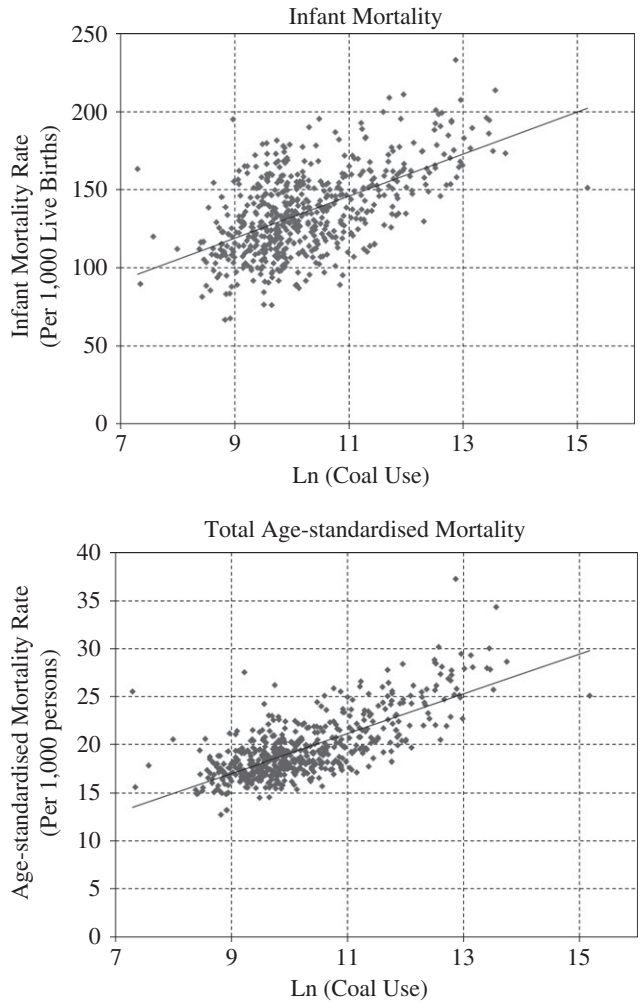

Fig. 2. *Coal Use and Mortality in England and Wales in 1851–60*
Notes. Local industrial coal use is based on the industrial composition of districts in 1851. The mortality rates are calculated using data from 1851 to 1860.

[22] The all-age mortality rates are age-standardised to account for differential mortality patterns at different ages. The formula is $\text{MORT}_d = \Sigma_g^G \text{MR}_{gd}\text{PS}_g$ where MORT_d is the age-standardised mortality rate for district d, MR_{gd} is the raw mortality rate in age-group g in district d and PS_g is the share of population in age-group g in the country as a whole. Thus, this formula adjusts a location's mortality rate to account for deviations in the age distribution of residents from the national age distribution.

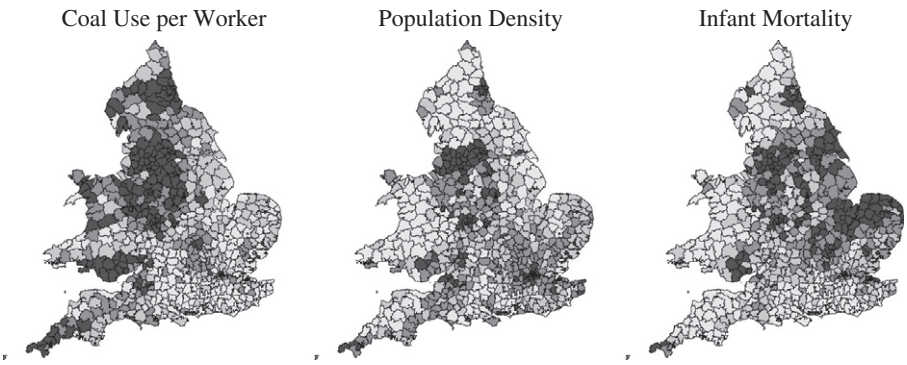

Fig. 3. *Maps of Industrial Coal Use, Population Density and Mortality*
Notes. Colours correspond to quartiles of each variable, where darker colours indicate higher values. We are grateful to the Cambridge Project on The Occupational Structure of Nineteenth Century Britain (funded by the Economic and Social Research Council) for their generosity in providing us with shapefiles for the 1851 Registration Districts. Infant mortality rates are calculated relative to births.

In addition to showing a strong positive relationship between coal use and mortality, this Figure also sheds light on the appropriate way to model the relationship between these variables. In particular, the relationship appears to be close to linear when coal use is in logs, suggesting that this is a reasonable way to model pollution effects. This implies a concave relationship between coal use and mortality, which is consistent with the concave relationship between pollution and mortality found in existing papers such as Pope *et al.* (2011) and Clay *et al.* (2016).

Figure 3 describes the spatial variation of the key variables used in the analysis: coal use, population density and infant mortality. While there are some similar geographic patterns (e.g. high levels in the industrial districts of Northwest England, as well as the areas around London, Birmingham, Cardiff, Bristol, Newcastle-upon-Tyne, and in Cornwall). Figure 3 also reveals a substantial amount of variation between these three variables. Further evidence on this point is provided in online Appendix Figure A3, which plots population density against district coal use. There we see that districts sharing similar levels of population density can differ substantially in their coal use levels, in some cases by as much as two to four log points. Online Appendix Table A3 builds on this point by describing population density, two pollution measures and mortality for a set of districts with similar populations.

3.2. *Baseline Results and the Urban Mortality Penalty*

We begin our analysis by studying infant mortality. This is a useful group to study as infants are unlikely to be exposed to pollution in one location and then die in another, which can be an important issue when studying adult mortality. Another advantage is that infant death rates are calculated using births as the denominator, which are available annually, rather than population, which is observed only in census years.

Our baseline specification is as follows:

$$\text{IMR}_d = \alpha_0 + \alpha_1 \ln(\text{DENSITY}_d) + \alpha_2 \ln(\text{COAL}_d) + X_d\Lambda + \epsilon_d, \quad (2)$$

where IMR_d is the number of deaths of children under one in district d divided by the number of births (measured in 1,000s) over the 1851–60 decade, DENSITY_d is the population density (measured in 1851), COAL_d is local industrial coal use, and X_d is a vector of control variables. For ease of interpretation, we standardise $\ln(\text{DENSITY}_d)$ and $\ln(\text{COAL}_d)$ to each have a mean of zero and standard deviation of one. To deal with spatial correlation, we follow Conley (1999) and allow for correlated standard errors between any pair of districts within 50 kilometres of each other.[23]

Table 1 presents our first set of results. Columns (1) and (2) present univariate regressions with coal use and population density, respectively, as explanatory variables. Both of these are strongly related to infant mortality. In particular, the results in column (2) document the substantial urban mortality penalty that existed in 19th century Britain. However, from historical evidence we know that coal use and the resulting air pollution was more severe in urban areas, so we expect these two variables to be correlated. In column (3), we add in the available set of control variables as well

Table 1

Coal Use, Infant Mortality and the Urban Mortality Penalty

	DV: Infant mortality rate				
	(1)	(2)	(3)	(4)	(5)
Ln(Coal use)	15.09***			13.68***	10.92***
	(1.358)			(2.462)	(2.328)
Ln(Population density)		16.55***	7.797***	5.255***	−1.351
		(1.440)	(1.531)	(1.432)	(1.521)
Ln(District population)			8.503***	−2.465	−1.924
			(1.421)	(2.612)	(2.393)
Child NPR mortality					1.500***
					(0.170)
Other controls			Yes	Yes	Yes
Observations	581	581	581	581	581
R^2	0.324	0.390	0.498	0.530	0.621
Implied lower bound on coal use coefficient					
Recommended approach ($\tilde{R} = 1.3 \times \hat{R}$):				12.6	8.3
Most conservative approach ($\tilde{R} = 1$):				10.5	5.6

Notes. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors, in parenthesis, allow spatial correlation between any pair of districts within 50 kilometres of each other. The dependent variable is the infant mortality rate: deaths under age one divided by births using data from 1851–60 and excluding deaths due to accidents or violence. Pollution measures are based on each district's industrial composition in 1851. The population, population density, and pollution variables are standardised. The controls in columns (3)–(5) include: mean altitude, hilliness, female labour force participation, agricultural suitability, a seaport indicator, and seaport tonnage. Child NPR mortality is based on the mortality rate per 1,000 births from a basket of childhood infectious diseases that are less likely to be affected by pollution: cholera, diarrhoea/dysentery, diphtheria, smallpox, scarlet fever and typhus. Implied lower bound on coal use coefficient is calculated using the approach from Oster (2016).

[23] Our choice of a 50 kilometre cut-off is motivated in part by, Clay *et al.* (2016), who argue that most of the air pollution exposure from coal burning power plants occurred within 30 miles (48 kilometres).

as log district population. Here we observe that both district population and population density have a strong positive association with the infant mortality rate.

In column (4) we add local coal use. Note that because log district population is included as an explanatory variable, the coal use variable can be interpreted as changes in the intensity of district industrial coal use per person.[24] There are a couple of important points to notice in column (4). First, there is a strong positive and statistically significant correlation between local industrial coal use and the infant mortality rate, which is only slightly weaker than the association suggested by column (1). Second, the estimated coefficient on population density drops from 7.8 in column (3) to 5.25 in column (4), a reduction of one-third. This suggests that coal use may help explain about one-third of the urban mortality penalty. The impact of overall district population on mortality also completely disappears.

There are two natural concerns with the results presented in column (4). First, there is the possibility of omitted variables that are correlated with local industrial coal use and also affect infant mortality. The set of control variables partially address this issue but in the Tables below we offer several additional approaches to addressing concerns about omitted variables. Second, we may be concerned that less healthy populations sort into more polluted areas, so that the estimated relationship between coal use and infant mortality is not driven by a causal effect of coal use on health.

To help address issues of population sorting, in columns (5), we include our child NPR mortality control, which is based on infant deaths due to cholera, diarrhoea and dysentery, diphtheria, smallpox, scarlet fever and typhus. Here the logic is that if the sorting of less healthy populations into more polluted areas is driving the relationship between pollution and infant mortality, then these sorting effects should also show up in diseases that are less directly affected by air pollution. After controlling for child NPR mortality, we continue to find strong evidence of a relationship between coal use and infant mortality. Furthermore, relative to column (4), the coefficient on coal use falls only slightly despite the fact that the R^2 value increases substantially. It is worth noting that coal use is likely to have some positive influence on each of the mortality-based control variables, for example through the impact of in utero or early-life exposure on health later in life, or through the interaction of pollution with infectious disease (Clay *et al.*, 2015). Any such effects should generate a downward bias in the estimated coefficient on coal use.

At the bottom of Table 1, we conduct a bounding exercise on the coal use coefficient following Oster (2016). The intuition behind this exercise is to use the movement of the coal use coefficient in response to the inclusion of the control variables relative to the amount of variation in the dependent variable that these controls explain (the change in R^2) to make inferences about the expected impact of including other unobserved controls. The key assumption that we make in this exercise is that the unobserved controls are as related to log coal use as the observed controls.[25] We also need to pick a maximum R^2 value, *Rmax*, that the unobserved control variables could

[24] If we do not include log district population in the regression in column (3), the estimated coefficient on coal use is 11.461 with a standard error of 1.15, which is not statistically different from the coefficient obtained when log district population is included.

[25] In Oster's notation, we are assuming $\delta = 1$.

achieve. We consider two alternatives here. First, we follow Oster's recommendation of setting $Rmax = 1.3 \times \hat{R}$ where $\hat{R}$ is the R^2 value from the fully controlled regression. Second, we consider the most conservative option, $Rmax = 1$. In either case, the lower bound estimate of the impact of coal use on infant mortality continues to imply substantial mortality impacts.

In online Appendix B.1, we present a variety of robustness results. In one exercise, we show that alternative approaches to measuring local industrial coal use, such as coal use per acre, coal use per worker, or an indicator variable for whether industries tended to be heavily polluting, produce similar results. We also show that the results in Table 1 are largely unchanged if we include a control for medical employment, a measure of local illiteracy levels, or if we run regressions weighted by district population.[26]

There is one final pattern in Table 1 worth pointing out. Specifically, in columns (2) and (3) we observe that population density has a strong positive association with infant mortality, reflecting the urban mortality penalty. In column (4), we see that this penalty remains even when we control for local coal use. However, once we control for the non-pollution-related infectious disease environment, in column (5), we see that the relationship between population (or population density) and mortality essentially disappears. If, however, we drop the coal use variable from the regression in column (4) (results not reported), the population–mortality relationship reappears and is statistically significant. What this is telling us is that together infectious disease mortality and coal use can explain the infant urban mortality penalty but that neither variable can completely explain the mortality penalty alone.

3.3. *Results Using Wind Patterns and Hilliness*

In this subsection, we exploit wind patterns in order to identify causal pollution effects better. Across Britain the predominant wind direction is from the south and west towards the north and east.[27] Exploiting this fact, we construct measures of upwind and downwind coal use for each district. We begin by first identifying all districts within 25 kilometres to the south and west (upwind districts) and those within 25 kilometres to the north and east (downwind districts).[28] We then calculate the coal used in each of these two sets of nearby districts. Because we want to work in logs, this analysis focuses only on the 422 districts where there are other districts in both the upwind and downwind quadrants, so that neither upwind nor downwind coal use is missing once we take logs.[29]

[26] We exclude medical employment and illiteracy from our main regression because both of these may be endogenously affected by pollution. When running weighted regressions we drop London, which would otherwise dominate the results.

[27] See online Appendix A.4 for information on wind patterns in Britain.

[28] Distances were measured between the district centre points, which were assigned to either the main district town or, for very rural districts, the geographic centre. Using a 25 kilometre window around the centre of a district allows us to capture most neighbouring districts but is small enough that the impacts of nearby pollution should be substantial. If we use values under 25 kilometres, we lose some direct neighbour districts from the sample.

[29] An alternative is to add one to the coal use in all districts, take logs, and then conduct the analysis using all of the districts in the data. Results in online Appendix Table B5 show that this generates qualitatively similar results, but that adjusting the upwind and downwind coal use values in this way affects the estimated magnitude of the effect of coal use.

Table 2

Placebo Tests – District Characteristics and the Effect of Upwind Versus Downwind Coal Use

	Ln(Population density)	Ln(Population)	Altitude (mean)	Hilliness	Seaport indicator	Seaport tonnage	Agricultural suitability	FLFP	Illiteracy share
Ln(Upwind coal)	−0.170***	−0.131***	0.202*	0.162	−0.0232	−0.0316	−0.0473***	−0.00206	0.0232**
	(0.0548)	(0.0302)	(0.108)	(0.0984)	(0.0236)	(0.0210)	(0.0155)	(0.00837)	(0.00969)
Ln(Downwind coal)	−0.163***	−0.142***	0.200**	0.138*	−0.0357*	−0.0297	−0.0308**	0.00487	0.0160*
	(0.0575)	(0.0306)	(0.0887)	(0.0765)	(0.0182)	(0.0221)	(0.0148)	(0.00721)	(0.00842)
Difference between downwind and upwind coefficients									
Coefficient difference	−0.007	0.011	0.003	0.024	0.012	−0.002	−0.016	−0.007	0.007
p-value	0.903	0.701	0.981	0.822	0.537	0.808	0.331	0.370	0.491

	Child NPR mortality	Cholera IMR	Diptheria IMR	Scar. Fev. IMR	Smallpox IMR	Measles IMR	Typhus IMR	Medical employ.
Ln(Upwind coal)	−0.995	−1.286**	−0.00770	0.127***	0.0855	0.0414	0.0590	−0.390***
	(0.632)	(0.566)	(0.0180)	(0.0439)	(0.0663)	(0.0552)	(0.0400)	(0.0810)
Ln(Downwind coal)	−0.903	−1.409***	0.000802	0.0969**	0.00611	−0.0226	−0.0142	−0.417***
	(0.564)	(0.538)	(0.0156)	(0.0427)	(0.0440)	(0.0652)	(0.0431)	(0.0614)
Difference between downwind and upwind coefficients								
Coefficient difference	−0.092	0.124	−0.009	0.030	0.079	0.064	0.073	0.027
p-value	0.879	0.827	0.708	0.542	0.203	0.344	0.185	0.761

Notes. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors, in parenthesis, allow spatial correlation between any pair of districts within 50 kilometres of each other. All regressions include controls for coal use within a district as well as population within 25 kilometres of the district. The coal use variables are all standardised. Child NPR mortality is based on the mortality rate per 1,000 births from a basket of childhood infectious diseases that are less likely to be affected by pollution: cholera, diarrhoea/dysentery, diphtheria, smallpox, scarlet fever and typhus. Medical employment is in logs. FLFP is the female labour force participation rate.

Comparing the effects of log coal use in upwind districts relative to downwind districts allows us to recover a causal effect of air pollution on infant mortality because wind patterns are unlikely to affect health aside from their role in transmitting air pollution. To implement this approach, we estimate a baseline specification that includes own district coal use, upwind coal use and downwind coal use. In this specification, the key statistic is the difference between the upwind and downwind coal use coefficients. The reason we focus on the difference is that proximity to high coal using districts may influence other outcomes (e.g. how urban the district is). Looking at the difference, however, overcomes this concern because we would not expect the impact of proximity to districts with more coal-using industries to differ depending on whether those districts were in an upwind or downwind direction except through air pollution spillovers.

Before presenting our main results, we assess the validity of this approach in Table 2. These results assess whether there is a differential relationship between coal use in upwind and downwind districts and other district characteristics. We consider variables reflecting district demographics, such as population, population density, and illiteracy, district physical features, such as altitude, hilliness and agricultural suitability, as well as mortality in diseases less likely to be associated with pollution. Out of this broad set of available variables, none exhibit a relationship to upwind coal use that is statistically

Table 3

Effect of Coal Use in Upwind and Downwind Districts

	DV: Infant mortality rate				
	(1)	(2)	(3)	(4)	(5)
Ln(Coal use)	8.727***	8.441***	10.06***	13.63***	10.74***
	(1.752)	(1.867)	(2.855)	(2.722)	(2.674)
Ln(Upwind coal)	1.694	2.849*	2.638*	1.424	1.907*
	(1.228)	(1.549)	(1.426)	(1.149)	(1.137)
Ln(Downwind coal)	−1.285	0.00599	−0.227	−1.420	−1.012
	(1.070)	(1.489)	(1.362)	(1.211)	(1.173)
Ln(Population density)	9.983***	10.68***	10.85***	4.950***	−1.509
	(1.486)	(1.492)	(1.608)	(1.615)	(1.546)
Ln(Nearby population)		Yes	Yes	Yes	Yes
Ln(District population)			Yes	Yes	Yes
Other controls				Yes	Yes
Child NPR mortality					Yes
Observations	422	422	422	422	422
R^2	0.463	0.467	0.467	0.566	0.664
Difference between downwind and upwind coefficients					
Coefficient difference	2.979	2.843	2.865	2.844	2.919
p-value	0.0602	0.0640	0.0614	0.0538	0.0426

Notes. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors, in parenthesis, allow spatial correlation between any pair of districts within 50 kilometres of each other. The additional controls include mean altitude, hilliness, female labour force participation, agricultural suitability, a seaport indicator, and seaport tonnage. Ln(Nearby population) is the population of districts within 25 kilometres. The dependent variable in all regressions is the infant mortality rate. Pollution measures are based on each district's industrial composition in 1851. The coal use variables are all standardised. Child NPR mortality is based on the mortality rate per 1,000 births from a basket of childhood infectious diseases that are less likely to be affected by pollution: cholera, diarrhoea/dysentery, diphtheria, smallpox, scarlet fever and typhus.

different than the relationship observed for downwind coal use. This shows that districts with more pollution in the upwind direction are not different in an observable way and suggests that an identification strategy that exploits wind direction is likely to generate valid results. In online Appendix Table B3, we present additional placebo results looking at patterns of adult mortality in diseases less likely to be related to pollution. For most of these diseases we also find no statistically significant difference between coal use in upwind *versus* downwind districts, despite the fact that adult mortality in a variety of disease categories through the long-run effects of early-life exposure. Overall, the results in Table 2 suggest that assessing the effect of upwind relative to downwind coal use will allow us to recover the causal effect of coal use on mortality.

Having documented the validity of this approach, Table 3 presents results for the infant mortality rate. Column (1) presents the baseline specification, while column (2) includes population within 25 kilometres to ensure that the upwind and downwind coal use variables are not simply reflecting overall population in nearby districts. Columns (3)–(5) progressively add district-level controls, with column (5) including our child NPR mortality control. As in Table 2, the key statistic is the difference between the upwind and downwind coal use coefficients. This difference, which is generally significant at either just above or just below the 5% level, is remarkably stable across specifications: a log point increase in upwind coal use (relative to downwind coal use) would increase infant mortality by about 2.9 deaths for every 1,000 births.

In Table 4 we consider an alternative identification strategy. There we interact coal use with the hilliness of the topography in each district (or more precisely, the SD of

Table 4

Effect of Coal Use Interacted with Hilly Topography

	DV: Infant mortality rate		
	(1)	(2)	(3)
Ln(Coal use)	11.80***	12.45***	9.462***
	(1.364)	(2.480)	(2.275)
Ln(Coal use) × hilliness	3.535***	3.783***	4.239***
	(1.091)	(0.961)	(0.845)
Ln(Population density)	6.739***	5.618***	−1.129
	(1.576)	(1.434)	(1.454)
Ln(District population)		−1.842	−1.210
		(2.543)	(2.262)
Mean altitude	1.149	−0.737	0.768
	(1.730)	(1.744)	(1.672)
Hilliness	−9.334***	−9.496***	−7.834***
	(1.761)	(1.888)	(1.961)
Other controls		Yes	Yes
Child NPR mortality control			Yes
Observations	581	581	581
R^2	0.527	0.549	0.645

Notes. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors, in parenthesis, allow spatial correlation between any pair of districts within 50 kilometres of each other. Hilliness is the standard deviation of altitude within a district. The additional controls include female labour force participation, agricultural suitability, a seaport indicator, and seaport tonnage. Pollution measures are based on each district's industrial composition in 1851. The coal use variables are all standardised.

altitude). This identification approach is based on the idea that in hilly areas coal smoke was more likely to be trapped in the valleys, increasing exposure. Other than by trapping in pollution, local hilliness tended to decrease infant mortality, most likely because it was associated with cleaner water due to faster running streams (Antman, 2016). These results suggest that district hilliness was associated with greater impact of industrial coal use on infant mortality. These effects are quite large; a 1 SD increase in district hilliness raised the impact of local industrial coal use by about one third. These results provide an alternative check on our finding that air pollution associated with coal use substantially raised infant mortality during the period that we study.

3.4. *Coal Use and the Urban Mortality Penalty*

Because wind patterns are plausibly exogenous to other factors affecting mortality, we view the results in Table 3 as providing our cleanest identification. Yet in order to consider the role of industrial pollution in explaining the urban mortality penalty, we need reliable estimates of the within-district effect of pollution, such as those shown in Table 1. We bridge this gap by bringing in additional information about the diffusion rate of pollution across space in England and Wales. Using this, together with the effects of coal use on mortality in upwind and downwind districts, we can generate back-of-the-envelope estimates of the within-district pollution effect, which can then be compared to the within-district estimates shown in Table 1. Here we briefly describe this approach but a more detailed description is available in online Appendix B.3.

The key input needed to relate estimated upwind/downwind coal use effects to the impact of coal use within a district is an estimate of how the concentration of coal-based pollution diminishes with distance. We generate this using the ADMS 5 pollution dispersion modelling software package.[30] We feed into the ADMS 5 software parameters describing a typical nineteenth century industrial smokestack.[31] The software then uses these parameters together with 10-year average meteorological data for four regions of Britain in order to model the pattern of pollution diffusion from our hypothetical smokestack.[32] The output provides ground-level pollution concentrations for 50×50 metres^2 cells across a 20×20 kilometres^2 grid with the smokestack at the centre.[33]

These pollution diffusion patterns allow us to estimate the average concentration of pollutants that coal burning generates within its own district, in a representative upwind district, and in a representative downwind district. These estimates suggest that, for a fixed amount of coal use, the average concentration experienced in the source district is four times higher than the average concentration experienced in a

[30] This software package, which has also been used by Heblich *et al.* (2016) to model the dispersion of coal smoke in nineteenth century Britain, was developed by Cambridge Economic Research Consultants.

[31] As in Heblich *et al.* (2016), we consider a 25 metres tall smokestack with a 1.5 metres diameter, an exit velocity of 4 m/s and a temperature of 120 °C.

[32] The meteorological data come from the Met and are from the 1980s or early 1990s.

[33] The original output is provided for four regions of England using the meteorological data specific to each region. Our analysis averages across the four different regional grids. The concentrations generated by the model should be thought of as representing a generic type of air pollution that is similar to total suspended particulates.

representative downwind district, while the concentration in the downwind district is 53% higher than in the representative upwind district.

Next, we relate the difference between the estimated effect of a given amount of coal use in upwind and downwind districts (without standardising the variables) to the difference in pollution concentration that we expect in these districts. This implies a relationship between the change in pollution concentration and the difference in infant mortality rates. We use this relationship, together with the difference in expected pollution concentrations in the downwind and source districts, to back out the expected impact of the same amount of coal use on mortality in the source district.

The results of this exercise suggest that increasing a source district's industrial coal use by one log-point would raise infant mortality in the source district by 7.26–9.67 deaths per 1,000 live births. For the sake of comparability with earlier results, this implies that a 1 SD increase in district log coal use would raise infant mortality by 8.12–10.82 deaths per 1,000 live births, or 6–8%. These values are just below the within-district estimated coal use effects shown in Tables 1 and in the top row of Table 3. This suggests that, while the within-district effects estimated in Tables 1 and 3 may be biased upwards, this bias is not likely to be large.[34]

3.5. *Effects Across All Age Categories*

Thus far we have focused on infant mortality, where concerns about selective migration occurring between pollution exposure and death are minimised. Next, we examine how local industrial coal use relates to mortality across all age categories.[35] To generate these results, we run regressions mirroring those reported in Table 3, which include upwind and downwind coal use, but with the mortality rate in different age groups as the dependent variable. These mortality rates are calculated using denominators based on the average population in the district during the sample decade, which were produced by the Census using data from the 1851 and 1861 Censuses.[36]

Table 5 describes the coefficients on both upwind and downwind coal use variables for age-group regressions, as well as F-statistics testing for the difference between upwind and downwind effects. For comparison, at the bottom of the Table we list the overall mortality rate for each age group in our data. All of the regressions in Table 5 include our full set of controls for district characteristics, including district coal use, population density, population etc. The results indicate that coal use in upwind districts increased mortality among those under five years old. There is also evidence of

[34] Also note that the values obtained using the upwind/downwind results are well within the bounds obtained using the approach from Oster (2016) in Table 1.

[35] These estimates should be thought of as a lower bound. This is because pollution can increase mortality through both acute effects, such as heart attacks due to exposure to high levels of pollution at a particular point in time, and chronic effects, such as lung diseases related to long-term pollution exposure. Migration will cause us to miss some of the chronic effects of pollution exposure because people who are exposed to pollution in one location may die somewhere else.

[36] One might be concerned about whether inaccurate population denominators are affecting our results. In online Appendix Table B10 we calculate additional results using only deaths in 1851–3 compared to population denominators based on the 1851 census. Because the mortality observations are closer to the census year, concerns about the population denominator should be less important in these regressions. These regressions deliver results that are very similar to those shown in Table 5.

Table 5

Upwind Versus Downwind Coal Use Effects By Age Group

DV: Mortality rate in each age category (per 1,000 persons)									
Under 5	5–9	10–14	15–19	20–24	25–34	35–44	45–54	55–64	65 up
Coefficient on coal use in upwind districts									
0.681	0.102*	−0.0002	0.0804	0.168	0.0806	−0.224**	−0.173	−0.136	−0.161
(0.416)	(0.0547)	(0.0433)	(0.0643)	(0.103)	(0.0828)	(0.112)	(0.128)	(0.170)	(0.360)
Coefficient on coal use in downwind districts									
−0.756	−0.0188	−0.0504	−0.0782	−0.0943	−0.139*	−0.248*	−0.200	−0.348**	−0.403
(0.477)	(0.0595)	(0.0411)	(0.0741)	(0.0864)	(0.0761)	(0.128)	(0.141)	(0.173)	(0.314)
Difference between upwind and downwind coefficients									
1.437	0.1208	0.0502	0.1586	0.2623	0.2196	0.024	0.027	0.212	0.242
F-test for significance of difference between upwind and downwind effects (F-stat & p-value)									
6.14	2.71	0.88	4.8	5.62	8.41	0.05	0.05	1.28	0.53
0.0136	0.1004	0.3495	0.029	0.0183	0.0039	0.8176	0.8284	0.2588	0.4652
Overall national mortality rate by age group									
68.23	8.52	4.93	6.96	8.58	9.74	12.46	16.85	29.36	91.62

Notes. ***p < 0.01, **p < 0.05, *p < 0.1. Standard errors, in parenthesis, allow spatial correlation between any pair of districts within 50 kilometres of each other. All regressions include controls for district coal use, population density, district population, population of other districts within 25 kilometres, altitude, hilliness, female labour force participation, agricultural suitability, a seaport indicator, and seaport tonnage. All regressions use observations for 422 districts with non-zero upwind and downwind coal use values.

smaller increases among the working age and older populations. The increases in mortality observed among the working age population may be due either to migration or because there are channels specific to this age group through which pollution affects mortality. One potential channel is tuberculosis, a respiratory disease, which was a particularly important killer among this age group. A second potential channel is that pollution may have increased death in childbirth.[37]

Given the substantial impact of coal use on mortality for children under 5 shown in Table 5, it is useful to use the upwind/downwind effects estimated in this Table to validate estimates of the within-district effects, as we did for infant mortality. Given the rate of pollution diffusion in Britain obtained from the ADMS 5 model, the estimated upwind and downwind effects for children under 5 shown in Table 5 imply that a one log point increase in source district coal use raises under-5 mortality in that district by 5.45 deaths per 1,000. This suggests that a 1 SD increase in log coal use in a district generates an increase in under 5 mortality of 6.1 deaths per 1,000. Direct estimates of the relationship between coal use and under 5 mortality (see online Appendix Table B11) suggest that a 1 SD increase in log coal use raises district mortality by 4.5–8.27 (8–15% of the under 5 mortality rate). Thus, results obtained by extrapolating from the effects observed in upwind and downwind districts fit near the middle of this range, suggesting that, as with the infant mortality analysis, any bias in the within-district estimates is likely to be small.

4. Discussion

Because we do not have direct measures of pollution levels, it is difficult to compare the effects we have estimated to the existing pollution literature. However, one way to put our findings into perspective is to calculate the impact on life expectancy implied by our results. Our basis for calculating life expectancy is a set of mortality rates obtained from the Registrar General's reports for five-year age bins up to age 75. Given these coarse values, our estimates are necessarily rough, so they are intended only to provide a sense of magnitude of the effects that we document.

The results above suggest that a 1 SD increase in district log coal use raised infant mortality by 8.12–13.68 deaths per 1,000 births, where the lower value is based on extrapolating from the effects observed in upwind and downwind districts and the higher value is from column (4) of Table 1. Applying these effects to mortality for those under one year of age suggests a reduction in life expectancy ranging from 0.33–0.56 years, which is 0.86–1.44% of total life expectancy (up to age 75) based on the mortality rates observed in 1851. If we conduct a similar exercise using the estimated impact on those under age 5 (from online Appendix Table B11) the results suggest that a 1 SD increase in local industrial coal use raised mortality by 4.37–8.27 deaths per 1,000 persons. These estimates suggest a decrease in life expectancy ranging from 0.84–1.58 years, which is equal to 2.19–4.11% of total life expectancy during this period.[38]

[37] Additional results, available upon request, suggest that local industrial coal use was associated with increased mortality in both of these categories; however, we leave explanation of these channels for future work since our identification strategies do not allow us to remove the role of migration.

[38] As a point of comparison, Chen *et al.* (2013) estimate that increased coal use due to a Chinese policy that provided free winter heating lowered life expectancy by 8%.

Another way to put our results into perspective is to use recent findings relating pollution levels to mortality, together with our estimated pollution impacts, to back out the implied difference in pollution levels between high and low coal-using regions in nineteenth century Britain. Specifically, we draw on the work of Arceo *et al.* (2016), which estimated the effect of PM10 exposure in Mexico and then compared the impact to estimates of the exposure in the US, where pollution levels were much lower. Their results suggest that the relationship between PM10 and infant mortality is close to linear. Assuming linearity, we can then use the estimated relationship between infant mortality and PM10 concentrations from their paper to generate implied differences in PM10 concentrations across the districts in our data.

Their study finds that a 1 $\mu g/m^3$ increase in PM10 raised infant mortality by 0.123 deaths per 1,000 births during a year. We find that a one log point increase in coal use increases infant mortality by between 7.26 to 9.67 deaths per 1,000 births.[39] Using these facts, we consider the implied difference in PM10 levels in a district lying at the 25th percentile of the coal use distribution in our data compared to a district lying at either the 75th or 90th percentile. Specifically, moving from the 25th to the 75th percentile of coal use in our data involves a 1.3 log point increase in coal use, which is associated with an increase in infant mortality of 9.4–12.6 additional deaths per 1,000 births. Dividing these effects by 0.12 (the marginal effect of PM10 on infant mortality) implies that moving from the 25th to 75th percentiles in coal use would increase pollution concentrations by 78–103 $\mu g/m^3$. Moving from the 25th to the 90th percentile implies an increase in PM10 concentrations ranging from 137 to 183 mg/m^3.[40] As a point of comparison, rough estimates from Greenstone and Hanna (2014) suggest mean PM10 concentrations in the US, India, and China from 1990–5 of 25, 133 and 178 mg/m^3 respectively.[41] This comparison suggests that pollution levels in British cities in the 19th century were likely to have been comparable to the levels experienced in some of the most polluted modern cities.

5. Conclusions

Due in part to lack of evidence, recent surveys of health and mortality during the nineteenth century have largely ignored the role of pollution. Cutler *et al.* (2006), for example, discusses the health effects of urbanisation in Britain but never directly addresses pollution. In Deaton (2013), pollution merits only a passing remark. Szreter (2005) spends just one out of four hundred pages discussing the role of pollution and draws primarily on anecdotal evidence. Focusing on the US, Costa (2015) describes

[39] These results are from the exercise (online Appendix B.3) where we aggregate our estimates relying on wind variation to estimate the effect of exogenous coal use (within a district) on infant mortality.

[40] If we instead use the results from Chay and Greenstone (2003) provided by Arceo *et al.* (2016) (their table 6, column 6, row 5), we calculate somewhat larger implied PM10 differences; a movement from the 25th to the 75th percentile of coal use in our data is associated with an increase in PM10 concentration of 101–135 $\mu g/m^3$, while a movement from the 25th to the 90th percentile is associated with an increase ranging from 179 to 238.

[41] To obtain estimated PM10 concentrations we convert the TSP values reported by Greenstone and Hanna (2014) using the formula PM10 = 0.55 TSP, following Arceo *et al.* (2016) and Knittel *et al.* (2016). The WHO air quality standard limits annual PM10 averages to below 20 $\mu g/m^3$. Arceo *et al.* (2016) report mean PM10 levels of 67 in Mexico using data from 1997 to 2006.

Pittsburgh skies darkened by pollution but argues that the lack of reliable particulate data limit our ability to measure the impact of pollution or to assess the benefits generated as air quality improved. This study fills this gap in the literature by providing broad-based and well-identified evidence of the impact of industrial pollution on mortality in the middle of the nineteenth century.

Our results indicate that local industrial pollution had a powerful impact on mortality during this period. Raising local industrial coal use by 1 SD from the mean increased infant mortality by roughly 6–8% and under-5 mortality by 8–15%. In terms of life expectancy, given mortality patterns in the 1851–60 decade, the impact of a 1 SD increase in local industrial coal use on under-5 mortality is associated with a reduction in life expectancy at birth of 0.84–1.58 years or 2.19–4.11%. In the most heavily polluted cities, such as Sheffield, Manchester, or Birmingham, where coal use was more than two standard deviations above the national mean, the effects could have been very large.

College of William & Mary
UCLA & NBER

Accepted: 1 March 2017

Additional Supporting Information may be found in the online version of this article:

Appendix A. Data.
Appendix B. Analysis.
Data S1.

References

Almond, D. (2006). 'Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population', *Journal of Political Economy*, vol. 114(4), pp. 672–712.
Alsan, M. and Goldin, C. (2014). 'Watersheds in infant mortality: the role of effective water and sewerage infrastructure, 1880–1915', Working Paper No. 21263, NBER.
Antman, F. (2016). 'For want of a cup: the rise of tea in England and the impact of water quality on economic development', mimeo, UC Boulder Working paper.
Arceo, E., Hanna, R. and Oliva, P. (2016). 'Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City', Economic Journal, vol. 126(591), pp. 257–80.
Ball, A. (2015). 'Air pollution, foetal mortality and long-term health: evidence from the Great London Smog', Working Paper No. 63229, MPRA.
Barreca, A., Clay, K. and Tarr, J. (2014). 'Coal, smoke, and death: bituminous coal and American home heating', Working Paper No. 19881, NBER.
Beach, B., Ferrie, J., Saavedra, M. and Troesken, W. (2016). 'Typhoid fever, water quality, and human capital formation', *Journal of Economic History*, vol. 76(01), pp. 41–75.
Bleakley, H. (2007). 'Disease and development: evidence from hookworm eradication in the American South', *Quarterly Journal of Economics*, vol. 122(1), pp. 73–117.
Brimblecombe, P. (1987). *The Big Smoke: A History of Air Pollution in London Since Medieval Times*, Methuen, MA: Methuen & co ltd.
Brunekreef, B. and Holgate, S. (2002). 'Air pollution and health', *Lancet*, vol. 360(9341), pp. 1233–42.
Cain, L. and Hong, S. (2009). 'Survival in 19th century cities: the larger the city, the smaller your chances', *Explorations in Economic History*, vol. 46(4), pp. 450–63.
Chay, K. and Greenstone, M. (2003). 'The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession', *Quarterly Journal of Economics*, vol. 118 (3), pp. 1121–67.

Chen, Y., Ebenstein, A., Greenstone, M. and Li, H. (2013). 'Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy', *Proceedings of the National Academy of Sciences*, vol. 110(32), pp. 12936–41.
Clay, K., Lewis, J. and Severnini, E. (2015). 'Pollution, infectious disease, and mortality: evidence from the 1918 Spanish influenza pandemic', Working Paper No. 21635, NBER.
Clay, K., Lewis, J. and Severnini, E. (2016). '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 No. 22115, NBER.
Conley, T. (1999). 'Gmm estimation with cross sectional dependence', *Journal of Econometrics*, vol. 92(1), pp. 1–45.
Costa, D. L. (2015). 'Health and the Economy in the United States from 1750 to the Present', *Journal of Economic Literature*, vol. 53(3), pp. 503–70.
Currie, J. (2013). 'Pollution and infant health', *Child Development Perspectives*, vol. 7(4), pp. 237–42.
Currie, J. and Neidell, M. (2005). 'Air pollution and infant health: what can we learn from California's recent experience?', *Quarterly Journal of Economics*, vol. 120(3), pp. 1003–30.
Currie, J., Neidell, M. and Schmieder, J. (2009). 'Air pollution and infant health: lessons from New Jersey', *Journal of Health Economics*, vol. 28(3), pp. 688–703.
Cutler, D. and Miller, G. (2005). 'The role of public health improvements in health advances: the twentieth-century United States', *Demography*, vol. 42(1), pp. 1–22.
Cutler, D., Lleras-Muney, A. and Deaton, A. (2006). 'The determinants of mortality', *Journal of Economic Perspectives*, vol. 20(3), pp. 97–120.
Deaton, A. (2013). *The Great Escape*, Princeton, NJ: Princeton University Press.
Ferrie, J. and Troesken, W. (2008). 'Water and Chicago's mortality transition, 1850–1925', *Explorations in Economic History*, vol. 45(1), pp. 1–16.
Fogel, R. (2004). *The Escape from Hunger and Premature Death, 1700–2100: Europe, America, and the Third World*, vol. 38, Cambridge: Cambridge University Press.
Fogel, R. and Costa, D. (1997). 'A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs', *Demography*, vol. 34(1), pp. 49–66.
Fouquet, R. (2012). 'The demand for environmental quality in driving transitions to low-polluting energy sources', *Energy Policy*, vol. 50(November), pp. 138–49.
Graff Zivin, J. and Neidell, M. (2013). 'Environment, health, and human capital', *Journal of Economic Literature*, vol. 51(3), pp. 689–730.
Greenburg, L., Jacobs, M., Drolette, B., Field, F. and Braverman, M. (1962). 'Report of an air pollution incident in New York City, November 1953', *Public Health Reports*, vol. 77(1), pp. 7.
Greenstone, M. and Hanna, R. (2014). 'Environmental regulations, air and water pollution, and infant mortality in India', *American Economic Review*, vol. 104(10), pp. 3038–72.
Hanlon, W. (2016). 'Coal smoke and the costs of the Industrial Revolution', Working Paper No. 22921, NBER.
Hanlon, W. and Miscio, A. (2017). 'Agglomeration: a long-run panel data approach', *Journal of Urban Economics*, vol. 99(May), pp. 1–14.
Hanlon, W. and Tian, Y. (2015). 'Killer cities: past and present', *American Economic Review, Papers and Proceedings*, vol. 105(5), pp. 669–86.
Heblich, S., Trew, A. and Zylberberg, Y. (2016). 'East Side Story: historical pollution and neighborhood segregation', CESifo Working Paper Series No. 6166.
Isen, A., Rossin-Slater, M. and Walker, W. (2017). 'Every breath you take – every dollar you make: the long-term consequences of the Clean Air Act of 1970', *Journal of Political Economy*, vol. 125(3), pp. 848–902.
Kesztenbaum, L. and Rosenthal, J.L. (2011). 'The health cost of living in a city: the case of France at the end of the 19th century', *Explorations in Economic History*, vol. 48(2), pp. 207–25.
Kesztenbaum, L. and Rosenthal, J.L. (2013). 'The democratization of longevity: how the poor became old in Paris 1870–1940', in (D. Ramiro-Fariñas and M. Oris, eds.), *New Approaches to Death in Cities during the Health Transition*, pp. 137–54, Springer Publishing.
Knittel, C.R., Miller, D.L. and Sanders, N.J. (2016). 'Caution, drivers! Children present: traffic, pollution, and infant health', *Review of Economics and Statistics*, vol. 98(2), pp. 350–66.
Logan, W. (1953). 'Mortality in the London fog incident, 1952', *Lancet*, vol. 261(6755), pp. 336–8.
Long, J. and Ferrie, J. (2003). 'Labor mobility', in (J. Mokyr, ed.), *Oxford Encyclopedia of Economic History*, pp. 248–50, New York: Oxford University Press.
Long, J. and Ferrie, J. (2013). 'Intergenerational occupational mobility in Great Britain and the United States since 1850', *American Economic Review*, vol. 103(4), pp. 1109–37.
McKeown, T. (1976). *The Modern Rise of Population*, New York: Academic Press.
Mitchell, B. (1984). *Economic Development of the British Coal Industry 1800–1914*, Cambridge: Cambridge University Press.
Mitchell, B. (1988). *British Historical Statistics*, Cambridge: Cambridge University Press.
Mosley, S. (2001). *The Chimney of the World*, Cambridge: The White Horse Press.

Oster, E. (2016). 'Unobserved selection and coefficient stability: theory and evidence', *Journal of Business Economics and Statistics*, forthcoming.
Pope, C.A., Ezzati, M. and Dockery, D.W. (2009). 'Fine-particulate air pollution and life expectancy in the United States', *New England Journal of Medicine*, vol. 360(4), pp. 376–86.
Pope, C., Burnett, R., Turner, M., Cohen, A., Krewski, D., Jerrett, M., Gapstur, S. and Thun, M. (2011). 'Lung cancer and cardiovascular disease mortality associated with ambient air pollution and cigarette smoke: shape of the exposure-response relationships', *Environmental Health Perspectives*, vol. 119(11), pp. 1616–21.
Rückerl, R., Schneider, A., Breitner, S., Cyrys, J. and Peters, A. (2011). 'Health effects of particulate air pollution: a review of epidemiological evidence', *Inhalation Toxicology*, vol. 23(10), pp. 555–92.
Szreter, S. (2005). *Health and Wealth: Studies in History and Policy*, Rochester: University of Rochester Press.
Thorsheim, P. (2006). *Inventing Pollution*, Athens, OH: Ohio University Press.
Townsend, J. (1950). 'Investigation of the smog incident in Donora, Pa., and vicinity', *American Journal of Public Health*, vol. 40(2), pp. 183–9.
Troesken, W. (2002). 'The limits of Jim Crow: race and the provision of water and sewerage services in American cities, 1880–1925', *Journal of Economic History*, vol. 62(8), pp. 734–72.
Troesken, W. and 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 (G. Libecap and R.H. Steckel, eds.), *The Economics of Climate Change: Adaptations Past and Present*, pp. 281–309, Chicago: University of Chicago Press.
Woods, R. (1997). 'Causes of death in England and Wales, 1851–60 to 1891–1900: the decennial supplements', [computer file].
Woods, R. (2000). *The Demography of Victorian England and Wales*, Cambridge: Cambridge University Press.