## DISCUSSION PAPER SERIES

# Atmospheric Pollution and Child Health in Late Nineteenth Century Britain

Roy E. Bailey
Timothy J. Hatton
Kris Inwood

## DISCUSSION PAPER SERIES

# Atmospheric Pollution and Child Health in Late Nineteenth Century Britain

**Roy E. Bailey**
*University of Essex*

**Kris Inwood**
*University of Guelph*

**Timothy J. Hatton**
*University of Essex, Australian National University and IZA*

# ABSTRACT

## Atmospheric Pollution and Child Health in Late Nineteenth Century Britain*

Atmospheric pollution was an important side effect of coal-fired industrialisation in the nineteenth century. In Britain emissions of black smoke were on the order of fifty times as high as they were a century later. In this paper we examine the effects of these emissions on child development by analysing the heights on enlistment during the First World War of men born in England and Wales in the 1890s. We use the occupational structure to measure the coal intensity of the districts in which these men were observed as children in the 1901 census. We find strong negative effects of coal intensity on height, which amounts to difference of almost an inch between the most and least polluted localities. These results are robust to a variety of specification tests and they are consistent with the notion that the key channel of influence on height was via respiratory infection. The subsequent reduction of emissions from coal combustion is one factor contributing to the improvement in health (and the increase in height) during the twentieth century.

**Corresponding author:**
Tim Hatton
Department of Economics
University of Essex
Colchester, CO4 3SQ
United Kingdom
Email: hatton@essex.ac.uk

## 1. Introduction

There has been growing interest in effects of the local environment on the health and growth of children in a variety of different countries and periods. Most of the focus has been on mortality rates of infants and children and on the health status of those who survived, either during childhood or later in life. Here we examine the effects of the local environment on the heights of men who were born in England and Wales in the 1890s. These data come from the records of army servicemen who enlisted in the First World War. Previous studies have suggested that although the socioeconomic status and demographic structure of the household were significant influences, conditions in the locality seem to have mattered even more.

Here we investigate one key element of the local environment: atmospheric pollution. In the age of dark satanic mills coal combustion was the major cause of atmospheric pollution. Emissions of black smoke were fifty times higher than present day levels. As studies of more recent times have identified negative health effects of polluted air we should expect to find even greater effects in the past. Recently, Beach and Hanlon (2016) have found a strong link between local coal use and mortality rates across districts in the mid-nineteenth century. Here we adopt a similar approach to the measurement of atmospheric pollution and we use it to explore the effects on adult heights of exposure to pollution during childhood. Our approach has the advantage that we can identify the cumulative effects during childhood, as represented by final height of those that survived. By controlling for household-level characteristics we can separately identify individual and locality effects, avoiding the hazards of inferring individual effects from group-level analysis.

The paper proceeds as follows. We first summarise recent findings on the effects of atmospheric pollution on health before turning to conditions in late nineteenth century Britain and their likely effects. We then introduce our data on the height of First World War servicemen and our proxy measure of coal intensity in the districts in which they grew up. Direct estimates of local coal intensity on height produce significant negative effects that are robust to a number of challenges. We also find strong links between individual height and district death rates from respiratory diseases, which are interpreted to reflect the local health environment. This provides evidence of the channel of influence running from atmospheric pollution to height, which is supported by instrumental variables analysis. We conclude that the large historical literature that focuses on height as the outcome of nutrition and disease should pay more attention to the effects of atmospheric pollution as one of the key sources of health insults during childhood.

## 2. Atmospheric pollution and child health

*2.1 Evidence on the links between pollution and health*

A large medical literature has identified the threats to health arising from atmospheric pollution. Black smoke is emitted from combustion of fossil fuels such as coal as well as petroleum oil, natural gas and other biomass fuels. It contains carbon dioxide ($CO_2$), sulphur dioxide ($SO_2$) and nitrogen oxides (NOx). These combine with moisture and solid material in the atmosphere, through primary and secondary reactions, to form suspended particulate matter with different degrees of fineness. Fine particulate matter is particularly damaging as it also contains substances such as acid sulphates and trace metals and it can more easily be absorbed into the respiratory system. The effects are particularly strong during childhood, as children are more vulnerable during the process of physical development and they may have greater exposure. The resulting inflammation of the lungs and oxidative stress leads to bronchitis and asthma, the effects of which may be greater for those with pre-existing illnesses (Schwartz 2004; Bateson and Schwartz 2008; Anderson 2009).

Recent research on the effects of atmospheric pollution has identified significant negative effects on the health of children, both concurrently and in later life (for a useful survey see Currie et al. 2013). Most of the existing literature focuses on the modern era (since 1950). Such studies have investigated the effects on health of a variety of pollutants, notably vehicle emissions, atmospheric lead and the effects of the Chernobyl nuclear disaster; examples include Almond, et al. (2009), Nilsson (2009) and Currie and Neidell (2009). These pollutants have been associated with negative effects on a range of child outcomes, such as infant and child mortality rates, cognitive ability, educational attainment, as well as adult outcomes such as productivity, earnings and later-life morbidity.

Here the focus is on coal combustion, which when concentrated, produces a particularly toxic form of pollution. Several studies have focused on its effects. Clay et al. (2016) studied the effects of the installation of coal-fired electricity generating plants in the United States between 1930 and 1962, finding that this increased local infant mortality within a 30-mile radius by 5-7 percent. Ransom and Pope (1995) examined the local effect of the shutting down and reopening of a steel mill in Utah in the 1980s. They found that hospital admissions of pre-school children on account of asthma and bronchitis were halved when the plant was not in operation. So the effects are large and immediate, even when not resulting in death. There is also evidence that the introduction of clean air acts in the US reduced infant mortality as well

as earnings in later life (Chay and Greenstone 2003; Isen et al. 2016). Even low (post-legislation) levels of emissions have discernible effects on deaths from respiratory and cardiovascular disease in the UK (Janke et al. 2009).

Emissions in developing countries are more comparable with historical levels, and the evidence suggests that these have substantial effects on mortality. Jayachandran (2009) finds that smoke from massive 1997 wildfires in Indonesia increased the mortality rate of under-threes by 20 percent. Similarly Tanaka (2015) finds that regulation placed on the emissions of coal fired powers stations in China in 1997 reduced infant mortality by 20 percent in the prefectures in which the policy was applied. In Turkey the progressive replacement of coal by natural gas for cooking and heating also has substantial effects, reducing infant mortality by around 30 percent over the decade to 2011 (Cesur et al. 2016).

Most studies relate to relatively immediate outcomes such as mortality but there are also cumulative effects of exposure on those that survive. Medical studies identify the effects of particulate matter ($PM_{10}$ or $PM_{2.5}$) on pulmonary function and lung development through oxidative stress, which leads to inflammation and often to acute and repeated respiratory symptoms (Heinrich and Slama, 2007; Rückerl et al., 2011). Inhalation of ultrafine particles that penetrate the surface lining (epithelium) and enter the bloodstream enable the translocation of particles to other organs. These pathways retard growth by diverting the body's energy to fighting disease rather than building bone and muscle.

Studies of developing countries find that stunting is much more common among young children in households that rely on biofuels than among those in households using cleaner fuels (Mishra and Retherford, 2007; Kyu et al., 2009; Ghosh et al., 2010). The effects of ambient air pollution on height are more difficult to assess. One study found that children residing in the highly polluted centre of Krakow grew by 1.5cm less over two years than those less exposed (Jedrycowski et al., 2002). For the UK Bobak et al. (2004) examined heights in a cohort study of children born in one week in 1946, a time when coal consumption was still relatively high.[1] They found a negative relationship between coal consumption and height at years of age between 2 and 7. This result was robust to the inclusion of birthweight, suggesting that the

---

[1] Clean Air Acts were introduced in the UK in 1956 and 1968, initially in response to the great London smog of 1952.

effects were not mainly *in utero*. In the light of these results one might expect substantial effects of atmospheric pollution on health in late nineteenth century Britain.[2]

*2.2 Victorian Britain*

Nineteenth century England is famous for the "dark satanic mills" that were spawned by the industrial revolution. By mid-century smoke pollution in industrial towns had risen to unprecedented levels. Engels (1845, p. 40) reported of Bradford that "on week-days the town is enveloped in a grey cloud of coal smoke, but on a fine Sunday it offers a superb picture, when viewed from the surrounding heights (1892, p. 40). In 1888 a correspondent of the *Manchester Guardian* noted that: "Our city, which ought to be one of the finest in the kingdom is to be entirely at the mercy of the smoke and noxious vapour producer, with the result that the atmosphere is poisonous, mortality at its highest rate, and altogether produces the most depressing effect upon the spirits of those unfortunate inhabitants who are compelled to live within its boundaries" (quoted in Mosely, 2008, p. 15). Smoke pollution was pervasive even in less heavily industrialised cities like York, where emissions from railway works and the local iron foundry blackened and eroded the walls of the Minster (Brimblecombe and Bowler 1992).[3]

The smoke emitted by industry was exacerbated by the use of coal for domestic heating and cooking, which accounted for about a fifth of inland coal use (Mitchell 1984, p. 12). Its effects were compounded by the crowded conditions in the industrial towns with their back-to-back terraces and crowded tenements, often without trees. These conditions are exemplified by the periodic smogs that occurred typically in cold, humid and still conditions, associated with anticyclones and temperature inversions. The famous London smogs produced excess mortality especially from bronchitis and pneumonia (Clay and Troesken 2011). Such conditions were common to most towns but, for a given population density, they were made much worse by the emissions from concentrations of heavy industry.

Total inland coal consumption increased more than tenfold during the nineteenth century. In Manchester the number of factory chimneys increased from 500 in the 1840s to around 2000 at the end of the century (Mosely 2008, p. 129, 147). Pollution worsened in step with the growth

---

[2] Indeed, Hanlon and Tian (2015) suggest that the mortality effects of industrial emissions in nineteenth century Britain were greater than those in China's most polluted cities in 2000.

[3] One symptom of rise and fall of industrial pollution is the history of the peppered moth (*biston betularia*). Initially this moth was predominantly white with black spots but during the late nineteenth century a predominantly black form (*carbonaria*) was observed with increasing frequency, beginning in the north east and spreading to London by the 1890s. By 1950, 90 percent of peppered moths took this form but over the following 50 years its range contracted and the original form returned to dominance (Cook 2003).

of coal consumption and it met with little effective regulation until the end of the century (Wohl 1983, pp. 220-4; Thorsheim 2006, pp. 114-124; Mosely 2008, pp. 173-180; Clay and Troesken 2011).[4] Coal consumption peaked in the 1950 and has since fallen by nearly 80 percent, while increasing regulation has reduced emissions even more steeply. In 1922 emissions of black smoke were 50 times greater than they were 80 years later (Air Quality Expert Group 2005, p. 263). Indirect measurement for London and Glasgow suggest even greater reductions since the late nineteenth century (Harrison 2006; Aplin 2012).[5]

Despite the seemingly obvious importance of atmospheric pollution, the debate over environmental health in the nineteenth century has focused chiefly on the inadequacies of water supply and sewage systems in densely populated towns. A number of studies have found that sanitary reforms were important in reducing infant mortality in the United States (Cain and Rotella, 2001; Troesken, 2002; Cutler and Miller 2005). Examining 36 English towns between 1871 and 1905, Millward and Bell (1998) found that while growth in incomes and improvements in housing had substantial influences on the reduction in mortality, investment in sanitation had only modest effects. Most recently Chapman (2016) estimates the effect of investment in sanitary infrastructure on mortality rates in 381 registration districts from 1861-70 to 1891-1900. He finds that these improvements significantly reduced mortality rates, especially in the under-five age group. Growth in the stock of sanitation-related capital accounts for between 15 and 25 percent of the overall mortality decline of about one fifth between 1861-70 and 1891-1900.

By contrast with the debate over the effects of better nutrition versus sanitary improvements, the effects of atmospheric pollution have received much less attention. As noted by Szreter (2005, p. 126) "clean air was one obvious omission from the late nineteenth century sanitary arsenal." Partly as a result of this there are few direct measurements of air pollution to compare with health outcomes. However, in a recent paper Beach and Hanlon (2016) develop a proxy for industrial pollution by using the industrial structure of employment in different registration

---

[4] The Public Health Act of 1875 (p. 32) defined emissions as one form of nuisance: "Any fireplace or furnace which does not as far as practicable consume the smoke arising from the combustible used therein, and which is used for working engines by steam, or in any mill, factory, dyehouse, brewery, bakehouse or gaswork, or in any manufacturing or trade process whatsoever; and any chimney (not being the chimney of a private dwelling- house) sending forth black smoke in such quantity as to be a nuisance,." Such offences were punishable by fines only if not mitigated "as far as practicable". The Public Health Act of 1891 (applying to London) included the same clause. In addition, prosecution could only be sustained in court if some injury could be proved (reply to a parliamentary question: House of Commons Debates, 10th May 1877 vol. 234 c620).

[5] These measurements are based on the electrical conductivity of the air, which is reduced by the removal of ions as they become attached to smoke particles.

districts in order estimate coal use. They find that this measure of coal use has a strong negative correlation with mortality for a cross section of registration districts in the decade 1851-60, and especially with mortality from diseases of the respiratory system. On this estimate pollution accounts for about one third of the urban mortality penalty, with the greatest effects on children under five years old.

The findings of Beach and Hanlon suggest that atmospheric pollution was a major health hazard but, as with much of the debate over sanitation, the focus is on death rates and the causes of death. Yet, as noted above, exposure to atmospheric pollution also has long term consequences for those that survive. In what follows we explore the question of whether, and to what extent, coal use affected growth during childhood, as reflected in the adult heights of those born in the 1890s. Our focus on individuals also has the advantage that it avoids the potential ecological fallacy of inferring individual effects from group-level analysis.

## 3. Data

*3.1 Army service records*

To examine the effects of atmospheric pollution on the health and growth of children in nineteenth century Britain we use a sample of the heights of servicemen recorded on enlistment in the British army during the First World War. By using the heights of young adults we are able to assess the cumulative effect of exposure to pollutants throughout the individual's childhood. The sample from the First World War ameliorates possible concerns about selection into the army.[6] Recruitment to the British army was not very selective and it included a large proportion of the relevant cohorts of adult males. It is estimated that 46 percent of males aged 15-49 in England and Wales in 1911 were recruited into the armed forces (Winter, 2004, p. 28). We focus on those born in the 1890s of whom about two-thirds would have been recruited into the army.[7] The sample includes both volunteers, most of whom enlisted in the early years of the war, and those that entered the army after the introduction of conscription in March 1916. For these reasons alone, those that enlisted would have been fairly representative of this cohort of men.

One particular concern in samples of army recruits is the minimum height requirement. The height standard, which was five feet three inches at the beginning of the war was progressively

---

[6] These concerns have recently been highlighted by Bodenhorn et al. (2013, 2015).
[7] The male population aged 15-24 in in 1911 was 3.2 million and the number who served (in all the forces) was approximately 2.1 million (Winter 2004, p. 82).

relaxed,[8] beginning in November 1914 with the first Bantam regiments, for which the minimum height was four feet ten inches. Even though recruits were subject to medical examinations, these were notoriously superficial and the rejection rate declined as the pressure for recruits mounted. By the last year of the War, when a large proportion of those fit for service had already enlisted, just 10 per cent of those examined were rejected as unfit for any kind of service (Winter 2004, p. 57). Only a very small proportion of these were rejected on the grounds of 'poor physique', although the medical grounds for rejection could well have been correlated with height.[9] So it is likely that those with very poor health, and perhaps with unusually short stature, were underrepresented. On the other hand the army service records do not include those that entered the army as commissioned officers. These were selected from the middle and upper classes and were typically taller than average. Officers accounted for about five percent of the army, although many of these would have risen from the ranks. Overall it is likely that the army service records somewhat under-represent the extreme tails of the underlying height distribution.

We select our sample from the army service records that have been made available on *Ancestry.co.uk*, largely for the benefit of genealogists. The army service records originate from the file WO363 held in the National Archives, which contains the records of around two million servicemen who completed their army service between 1914 and the early 1920s.[10] These records comprise a (not necessarily complete) set of forms for each serviceman. The information that we use is taken from the attestation form that was filled in upon enlistment and the medical inspection report form. The attestation form recorded the serviceman's name, age, address, and usually his birthplace and details of his next of kin. The medical inspection report records various physical characteristics relating to fitness for service and most importantly height in inches. A sample was taken of the first 2 per cent of each three-letter

---

[8] With the surge of volunteers in the first few months of the War the height standard was raised to five feet six inches in September 1914 then lowered to five feet four in October and back to five feet three in November. It was further reduced to five feet two inches in May 1915. For the background to these policies, see Grieves (1988), Ch. 1, and Simkins (1988), Ch. 4.

[9] See Ministry of National Service (1920), which provides a compilation of statistics from the examinations conducted by the National Service Medical Boards that were established in 1917. The most common cause of rejection was tuberculosis, which accounted for about a quarter of all rejections.

[10] These records are sometimes know at the 'burnt documents.' As a result of fire at the War Office building at Arnside Street, London in September 1940, approximately three fifths of the original 5 million records were destroyed or damaged.

combination for about one third of the letter combinations, the principal focus being on those born in the years 1892 to 1897.[11] In total we have records for 2956 servicemen.

The individual's childhood circumstances are captured by finding each serviceman in the 1901 census of England and Wales. We discard those that were resident in Scotland or Ireland and those for whom key information was missing from the service record. We search by name and age-range and then use other information in the army service records to discriminate between alternative potential matches.[12] One important aid to identifying the correct household is that most of the servicemen were unmarried at the time of enlistment and so their next of kin was usually a parent who would be observed in the same household in 1901. Cases for which no match was found, or where there was no basis to choose among alternatives, were discarded. Also discarded were those who, on closer inspection, turned out not to be born in the 1890s and a few that were living in institutions containing more than 15 people. Our final matched sample consists of 2235 individuals. This is 85.3 percent of those for whom we searched, which is a much higher success rate than is typically obtained when relying solely on a computer algorithm.

The distribution of heights in our sample of servicemen is shown in Figure 1. The average height is five feet six inches, although the modal height is five feet seven. Although normality can be rejected the tails of the distribution are fairly symmetrical, with 9.5 per cent of the sample shorter than five feet three inches and 10.8 per cent taller than five feet nine inches.

Some mean characteristics of these individuals are reported in the left hand panel of Table 1. The average age of the servicemen was just over 20 on enlistment and the average date of attestation was halfway through 1915. The households in which they were observed as children in 1901 were mostly headed by manual workers, with only 13 percent in white-collar occupations. As there was no occupation-based social classification in the 1901 census, this is based on applying the 1921 classification to the occupational titles. The overwhelming majority of the servicemen grew up in male-headed households with the remainder mainly headed by widows. On average they had just over three siblings when observed in 1901. Fertility was still relatively high at the turn of the century and completed sibships would have been larger still.

---

[11] On closer inspection some of these were found to have been born earlier or later than the year reported in the summary in *Ancestry*.
[12] For this purpose we used the genealogical website *Findmypast*. This was preferred to *Ancestry* as search engine is rather more discriminating and also because the household summary is presented in a way that lends itself more easily to data entry.

We also have details of the number of rooms occupied which is capped at 5. Nearly half were living in households with more than one person per room.

## 3.2 Locality characteristics

The key variable for our purpose is coal use in the locality where these servicemen grew up. From the census we can identify the registration district in which the household was located. In 1901 there were 635 registration districts in England and Wales with an average population size of 51,000. The registration districts, based on poor law unions, exhibit a wide diversity of local conditions, much of which would be averaged away at the county level. We follow Beach and Hanlon (2016) in constructing a measure of coal intensity for each registration district by combining coal use per employee from the 1907 industrial census with the occupational structure of the registration district in the 1901 population census. Our measure of coal intensity for a district can be represented as follows:

$$Coal\ Intensity_r = \sum_{i=1}^{n} c_i o_{ir} / \sum_{i=1}^{n} o_{ir} \tag{1}$$

Where $c_i$ is annual coal consumption per employee in occupation $i$, $o_{ir}$, is the number employed in occupation $i$ in district $r$, and occupations are indexed from 1 to $n$. Thus the ratio is the predicted industrial coal consumption per occupied worker in the registration district.

In order to construct this index, we first link each occupation recorded in the 1901 census to one of the industries that was included in the census of production. These industries represent about 60 percent of inland coal use. The census of production includes 13 industry groups and 139 subgroups for each of which coal consumption per employee is reported.[13] We then match the occupations in the 1901 census with these industrial subgroups and give each occupation the coal consumption per employee in the matched subgroup. Only 434 of the 796 unique occupations are matched and the remainder are given a coal consumption value of zero. We then obtain average coal consumption in each registration district by applying this measure of coal consumption by occupation to the occupational composition of the district. It is important to stress that this is a measure of industrial coal use only. It does not include coal consumed in domestic heating, nor does it cover coal used in railways or other modes of transport. As domestic coal consumption is ubiquitous it should not matter very much for the comparison

---

[13] The 13 industry groups are: (1) mines and quarries; (2) iron and steel, engineering and shipbuilding; (3); metal trades other than iron and steel; (4) textiles, (5) clothing; (6) food, drink and tobacco; (7) chemical and allied trades; (8) paper, printing and allied trades; (9) leather, canvas and india-rubber; (10) timber; (11) clay, stone, building and contracting (12) miscellaneous; (13) public utility.

across districts. And while railways are more concentrated the atmospheric pollution that they generate would be more dispersed.[14]

Figure 2 provides a map of the registration districts where darker shading represents higher quartiles of coal intensity. The map highlights the geography of the industrial revolution with a broad swathe of darker shading covering the textile regions of Lancashire and Yorkshire, as well as the manufacturing districts of the midlands the northwest and coal mining areas of the northeast and South Wales. The location of the cotton industry was largely determined by original access water power and to ports and markets (Crafts and Wolf, 2014), while heavy industries were located in proximity to minerals, and above all to coalfields (Fernihough and O'Rourke, 2014; Beach and Hanlon, 2016). Thus the distribution of industry owed most to these geographical features and should not be the result of differences in physical stature.

We also calculate an alternative measure for heavy coal use only. This is represented as:

$$\text{Heavy coal use}_r = \sum_{i=1}^{n} z_i o_{ir} / \sum_{i=1}^{n} o_{ir} \tag{2}$$

Where $z_i$ is an indicator that takes the value 1 if the particular occupation is in an industrial subgroup for which coal consumption, $c_i$, exceeds 10 tons per employee per annum. Heavy coal use concentrates attention on those industries for which the pollution effect is likely to be most severe. It also avoids potential measurement error, which may arise for lower levels of coal intensity.

In our sample we observe individuals growing up in 472 out of the 635 registration districts. Not surprisingly, these randomly selected individuals are often from the larger registration districts. As Table 2 shows, the average serviceman came from a registration district with a population of nearly 150,000 and with a population density of 21 persons per acre. The average coal intensity of the districts in which these individuals lived was 2.5 tons per worker with 4.4 percent of the labour force engaged in heavy coal use industries. Finally, the child mortality rate in the districts where the servicemen grew up is 5.8 percent.

---

[14] For coal mines the relevant value relates only to coal consumed, not coal produced and the same applies to coking ovens. Coke consumption is not included here although the results reported below differ little if coke consumption is added (ton for ton) to coal consumption.

## 4. The effect of coal combustion on height

*4.1 Height and coal intensity*

Did coal intensity affect the growth of children in the late nineteenth century and therefore influence their heights as adults? Figure 3 plots height against coal intensity. This is entered in log form to account for the diminishing effect of more intensive coal use. Not surprisingly there is a wide scatter of heights with evidence of a mild downward slope.

Table 3 reports regression results where the dependent variable is height in inches and the standard errors are clustered by registration district. The regressions include dummy variables for the 11 regions of England and Wales distinguished by the Registrar General (not reported in the table).[15] These capture regional differences in climate, such as rainfall, average temperature and insolation and possibly also economic differences such as wage and income levels. Thus our estimates are within-region effects.

The coefficients on individual and household variables are consistent with those reported in Bailey et al. (2016) and can be summarised briefly. The dummy variables for year of age show height advancing with age up to 20 after which there is no further increase. The number of siblings has a negative effect on height, which is consistent with the idea of a trade-off between the quality and the quantity of children (a point explored more fully in Hatton, 2016). Those that grew up in female-headed households were shorter that those in male-headed households but not significantly so. Not surprisingly servicemen that grew up in middle class households were about half an inch taller than those in households headed by manual workers. And those growing up in larger dwellings were also taller, by around a third of an inch. The latter two variables, in particular, are strong indicators of income and living standards.

Turning to the principal variables of interest, column (1) shows that our coal intensity variable takes a significant negative coefficient. The coefficient implies that the difference between the 10th and 90th percentile of coal use across registration districts is associated with a difference of almost an inch in height. Column (2) shows that using instead the proportion of the labour force engaged in heavy coal use occupations also gives a significantly negative coefficient. It implies that increasing the share of those engaged in heavy coal use occupations from the 10th to 90th percentiles is associated with a height reduction of about quarter of an inch. But as shown in column (3), when both measures are included in the regression, coal intensity

---

[15] The regions are; London, South East, South Midland, East, South West, West Midland, North Midland, North West, Yorkshire, North and Wales and Monmouthshire.

dominates and the coefficient on heavy coal use becomes insignificant. Column (4) adds the population density of the registration district in 1901. This variable, above all, is often associated with poor health and high infant mortality. But it may also capture the effects of smoke pollution from domestic fires. Population density takes a negative coefficient, as would be expected, but it has very little effect on the size and significance of the coefficient on coal intensity. These results are robust to different levels of clustering and to different functional forms, as reported in Tables A1 and A2 of the appendix.

*4.2 The incidence of atmospheric pollution effects*

If the coal intensity variable captures atmospheric pollution, and especially the presence of particulate matter, then it may have diffused beyond the immediate vicinity and into neighbouring districts. As the wind direction is predominantly south-westerly, pollution from outside the district should come from the southwest and not from the northeast. In order to test this hypothesis we calculate the average coal intensity (as defined in equation (1)) of districts that fall within the southwest quadrant up to 25km from a central point in the district. The central point in a district is defined either as the major concentration of population (usually a metropolitan district) or, in the case of rural districts, as the approximate geographical centre.[16] The same procedure was used to calculate coal intensity of districts to the northeast and similar variables were calculated for heavy coal use (as defined in equation (2)).

Table 4 shows the result of adding the variables for coal intensity in districts to the southwest and to the northeast. All of the individual and household variables that appeared in Table 3 are included but not reported. In column (1) coal intensity to the southwest has a significant negative effect on height while coal intensity to the northeast has a small and insignificant effect. Note that the coefficient on coal intensity within the district remains negative and significant although it is somewhat smaller than in column (4) of Table 3. This is not surprising in light of the clustering of coal intensity illustrated in Figure 2. Diffusion across districts might be especially strong where coal use is particularly heavy. Accordingly, column (2) includes the measure of heavy coal use in districts to the southwest and the northeast. As before, heavy coal use to the southwest takes a negative and significant coefficient while heavy coal use to the northeast does not. This is an important test of the hypothesis that the coal use

---

[16] These locations were identified from the maps available at: www.visionofbritain.org.uk/, and the coordinates were then found using: www.latlong.net/. This is similar to the procedure used by Beach and Hanlon (2016); we are grateful to Walker Hanlon for advice on this point.

variables are capturing the effect of atmospheric pollution and not some other effect associated with coal use within each district.

It is possible that these results would be stronger still if predominantly rural and agricultural districts, spread over large areas and distant from industrial conurbations, are omitted. In order to test this hypothesis, we exclude districts that were predominantly agricultural.[17] Column (3) shows that, when agricultural districts are excluded, the coefficients on coal intensity and coal intensity to southwest become slightly larger, as would be predicted. A further possibility is that the results depend on including the 28 districts of London, which are small in area and in close proximity. The result in column (4) of Table 4 indicates that the coefficient on coal intensity to the southwest becomes somewhat weaker when London is excluded, London although it is still significant at the 10 percent level. It is also notable that the coefficient on population density increases in size.

The effect of coal intensity on health and height could be influenced by topography and climate. Column (1) of Table 5 includes a dummy variable for districts with some (non-estuary) coastline and average wind speed. The wind speed variable used here is average wind speed (in metres per second) over the years 2000-10 at a height of 25 metres from the ground (a little higher than most Victorian factory chimneys). This is from the NOABL database of the UK Department of Energy and Climate Change. These variables might have been expected to lead to greater dispersal of coal emissions, and hence positive coefficients. However, both coefficients are negative, neither is significant and the coefficient on coal intensity remains strongly significant.

Column (2) includes average temperature, measured as average annual growing degree days (divided by 365) for 1961 to 1990. This is taken from the UK Meteorological Office's 5km gridded database. One interpretation of the positive coefficient is that higher temperatures and greater atmospheric convection is associated with greater dispersal of smoke pollution. Column (3) introduces average rainfall, measured as average daily precipitation (greater than 1mm) for 1961-90, also from the Met Office's 5km gridded database. Rainfall could help either to suspend or to disperse particulate matter and the coefficient is negative but insignificant. In columns (2) and (3) the coefficient on coal intensity remains negative and significant.

---

[17] This classification is taken from the Registrar General (1907), p. viii.

Column (4) of Table 5 includes a measure of terrain ruggedness. This is measured as the average slope in the area of the registration district using elevation data from the Shuttle Radar Topography Mission gridded elevation dataset. As towns and cities tend to be in valleys rather than on hills, greater ruggedness is likely to trap atmospheric pollution leading to slower dispersal. The coefficient on average ruggedness takes a negative and significant coefficient, which is consistent with this hypothesis. The coefficient on coal intensity is slightly reduced in size but it remains significant. As shown in Appendix Table A3, the influences of climate and topography on height in Table 5 are robust to the addition of coal intensity in adjacent districts, as in Table 4.

*4.3 Selection and mobility*

One possible concern is that, in spite of the inclusion of a set of household-level variables, coal intensity may be capturing heterogeneity among households that were engaged in different industries. Thus the effect of coal intensity may reflect differences in the composition of household types rather than pure district effects. As pointed out by Higgs (1991) the Registrar General's office was much concerned with examining the links between occupations and health outcomes. Occupation was associated not only with the medical consequences of working with certain materials but also with the individual's moral and behavioural characteristics (Higgs 1991, p. 473).[18] Thus, if the effect of coal intensity on height is to be interpreted as reflecting atmospheric pollution, then it should not be affected by the coal intensity of the household head's occupation. In order to test this hypothesis we have linked coal intensity to the occupations of the heads of households of the individual servicemen that we observe as children in the 1901 census. Coal intensity is zero for 52 percent of our households and so we add one before taking the log.

Column (1) of Table 6 shows that the coal intensity of the occupation of the head of household takes a coefficient that is negative but small and insignificant. Column (2) includes a dummy variable taking the value 1 if the head of household was employed in an occupation with heavy coal use (as defined above). This too is negative and insignificant. These results provide little evidence that those that grew up in households with heads that worked in coal intensive industries suffered worse health than others in the same district. They suggest that the influence of local coal intensity is not simply a composition effect but a genuine locality effect. This is

---

[18] According to Higgs (1991, p. 472), "The organizing concept underlying the occupational classification systems of the mid-nineteenth century was, in fact, one based on the materials being worked up in productive processes."

all the more striking as households with heads engaged in coal intensive industries are likely to have lived nearer to the immediate location of coal combustion than others in the same district.

It is often observed that middle class households attempted to avoid the worst of atmospheric pollution by moving, even within districts, to locations further from the immediate sources of emissions (Wohl 1983 p. 215; Thorsheim 2006, p. 43; Mosely 2008, p.31-2). Heblich et al. (2016) show that low-skilled workers are more concentrated on the east side of Victorian cities. If so, this is another source of possible heterogeneity in the effects of local pollution on child health. In order to test this hypothesis columns (3) and (4) include an interaction between the log of coal intensity in the district and the dummy variable for white collar households. If middle class households chose locations within the district that enabled them to escape some of the effects of atmospheric pollution then the coefficient should be positive. But in both regressions it is negative and insignificant. However, as noted above, there are relatively few of these households in the data and they are typically in lower middle classes occupations such as clerks and other office workers, many of whom would be employed in factories.

There was considerable geographic mobility in late nineteenth century (Baines, 1982; Boyer and Hatton, 1997; Long, 2005). This raises a potential concern that the effects of coal intensity on height may reflect selective migration of households into districts with different levels of coal intensity. It is sometimes argued that migration to the towns and cities selected the fittest. If migrants tended to be taller than average, genetically or otherwise, then the same might apply to their children. On the other hand it is possible that districts with heavy coal use recruited those with below average health, whose children might be shorter than average.[19] If so then the negative effect of coal intensity on the height of servicemen in our sample could reflect the selective in-migration of their parents.

In order to identify migrant families we have examined the birthplaces of the heads of the households in which the individual servicemen lived as children in 1901. We use the information on birthplaces reported in the census in order to find the county and registration

---

[19] This point was recognised the Registrar General's extensive analysis of the link between occupation and mortality: "it is apparent that there are selective tendencies at work. In the first place, some occupations are ordinarily recruited from sections or classes of the community which exhibit a special health standard of their own: agricultural workers, for example are usually recruited from the more healthy rural population. But, generally speaking, the selection process operates by reference to the physical aptitude of the individual worker, whether it takes effect through his individual choice or change of occupation or, or through his unwilling exclusion or expulsion (Registrar General, 1920, p. ii).

district of birth of the head of household, normally the individual's father.[20] Among all the household heads, 59 percent were living in a registration district different from that in which they were born and 36 percent were living in a different county. Column (1) of Table 7 shows that a dummy variable taking the value 1 for inter-district in-migrant household heads takes a positive coefficient, which is significant at the 10 percent level. A dummy restricted to inter-county migrants takes a larger and more significant positive coefficient in column (2). This is consistent with what is sometimes known as the 'healthy migrant effect', but in this case for the second generation. Not surprisingly, the positive selection is stronger for longer distance moves. However, including this variable has a little effect on the coefficient on coal intensity.

Although the children of migrants were taller as adults than the children of non-migrants, that alone does not identify selection into districts by coal intensity. In order to assess this we attach the 1901 coal intensity to the district of birth of the head of household. This is only possible for the 94 percent of household heads that were born in England and Wales, for whom we can measure coal intensity in the district of birth. On average, migrant household heads moved to districts where coal intensity was just 8 percent higher than in the districts where they were born. Column (3) of Table 7 includes the difference in log coal intensity between the district of residence and the district of birth. The coefficient is positive and significant at the 10 percent level, which suggests that the children of migrant household heads that moved into districts with greater coal intensity tended to be taller than the average for the district. Column (4) shows, however, that there is no equivalent effect for the children of heads that moved into districts where heavy coal use was higher than in the district of birth. As shown in columns (3) and (4) the coefficient on coal intensity in the district of residence is not much affected by the selective migration of households into districts with greater atmospheric pollution from coal. So the effect that we estimate seems to be despite, rather than because of, selective in-migration.

The mobility of the servicemen during childhood is another potential source of concern. For those that moved sometime before the 1901 census, childhood exposure to coal pollution would arise partly from the district of residence and partly from the district of birth. This might be especially important in the early years of childhood. In order to examine this issue, we take the

---

[20] Birthplaces were reported by the householder to the census enumerator. These are sometimes a parish but often a district, city, town, village or specific locality. In order to identify the corresponding registration district (in 1901) we have used the lists of parishes, registration districts and the online gazetteer at: http://www.genuki.org.uk. Cases where the individual was born outside of England and Wales (mostly Scotland or Ireland) are coded as a change of county and a change of registration district.

individual's birthplace as recorded in the census and, as with the household heads, we use this to identify the registration district and county of birth. Among the individual servicemen, 21 percent had moved registration district sometime before they were observed as children in the 1901 census and 10 percent had moved county. As with the household heads, this reflects the relatively high geographic mobility of Victorian households. The average age of the servicemen in 1901 is 5.5 years, so it implies that about 4 percent of households moved districts in a year, and 2 percent also moved county. The first and second columns of Table 8 display the results of including the dummy variable for inter-district and inter-county movers respectively. As children would not have been the prime movers, it is not surprising that these coefficients are insignificant.

As with the heads of household, we have attached the measure of coal intensity to the individual's district of birth. The third column includes the difference between the log of coal intensity in the district of residence and the district of birth. If coal intensity in the district of birth contributes to height then the coefficient should be positive, representing the contribution of conditions in the district of birth.[21] Although the coefficient is positive as predicted, it is insignificant. One possible reason is that, in the absence of information on exactly when the individual moved, it is not possible to assign weights for exposure in the district of birth and in the district of residence. Alternatively, in column (4) those that moved district are dropped from the regression, with only a modest effect on the coefficient on log coal intensity. As some of these individuals would have moved during the latter years of their childhood (after 1901), there might be some attenuation, although the effects of pollution on height are likely to have been greatest during early childhood.

## 5. Height and child mortality

*5.1 Causes of death*

Studies of height often include infant or child mortality in the locality as an explanatory variable in place of, or in addition to, other characteristics of the locality. As is often observed, the mortality rate among infants or among those under five is a sensitive indicator of the local disease environment. The effect of infant mortality on height can be interpreted as the balance of two forces (Bozzoli et al. 2009). One is the selection effect, where mortality selects those

---

[21] The effect of the weighted average of coal intensity can be written as: $\beta(\alpha CI_R + (1 - \alpha)CI_B)$, where $CI_R$, and $CI_B$ denote coal intensity in the districts of residence and birth respectively and $\alpha$ is the weight or duration in the district of residence. This can be rearranged to give: $\beta CI_R - \beta(1 - \alpha)(CI_R - CI_B)$, so the coefficient on the difference term should take opposite sign to the main effect and is scaled by the weight on the district of birth.

with weaker constitutions, who would have been shorter than average if they had survived. In this case the higher the mortality rate, the greater will be the selection effect, and the taller will be the survivors. The other is the scarring effect, where infant mortality stands as a proxy for elements of the disease environment in the locality that inhibit growth among those that do survive. It therefore has a negative effect on height.

In most historical studies the scarring effect dominates so that the net effect is a negative correlation between height and mortality. Indeed, one attempt to assess these effects separately for children in interwar Britain finds no evidence of a selection effect (Hatton 2011). But such studies often focus on the overall mortality rate without distinguishing the separate causes of mortality. As far as the scarring effect is concerned, different causes of mortality will be associated with different elements of the disease environment. This should help to identify the epidemiological channels through which the local environment affects height. In particular, if atmospheric pollution is an important influence then we should expect to see a link between height and the incidence of death from respiratory disease.

Table 9 shows deaths rates under the age of five, from specific causes and in total, for England and Wales over the decade 1891-1900. Death rates per thousand are calculated as the annual average of deaths divided by the average population under five years of age. The well-known infectious diseases, measles, scarlet fever, diphtheria and whooping cough, accounted for 15 percent of all deaths, with another 5 percent accounted for by tuberculosis. Diarrhoea and dysentery, together with other diseases of the digestive system, accounted for nearly 20 percent while respiratory diseases, mainly bronchitis and pneumonia, accounted for 21 percent. Of course the classification of diseases was somewhat imprecise even though it had improved markedly in the preceding decades (Hardy 1994). Of particular note is the large category of other causes, where the cause of death was either not recorded or was impossible to fit into the standard classification.

The relationship between height and district mortality rates is explored in Table 10. As previously, the coefficients of individual and household variables are included but not reported. The first column demonstrates the strong negative relationship between individual height and the overall death rate in the district of childhood residence. Clearly, any selection effect is overwhelmed by the scarring effect. That in turn suggests that specific causes of death will capture different attributes of the underlying disease environment. Column (2) includes only the death rate from 'diseases of the respiratory system'. This gives an equally strong negative

coefficient, which reflects the high correlation in the sample of 0.94 between mortality from respiratory diseases and total mortality.[22]

Column (3) adds variables for death rates from two more causes of death. Gastrointestinal disease includes diarrhoea and dysentery as well as diseases of the digestive system, while infectious disease is the sum of deaths from measles, scarlet fever, diphtheria and whooping cough. Gastrointestinal disease evidently has little additional effect on height, whereas infectious disease has a negative effect on height that is significant at the 10 percent level. In the presence of these additional variables respiratory disease retains significance. In the last column all identified causes of death other than respiratory disease are added together as one group. This takes an insignificant coefficient while death from respiratory disease remains just significant. This should not be taken to imply that growth was inhibited exclusively by respiratory diseases but rather than these best represent the local conditions associated with restricted growth.[23] As shown in Appendix Table A4, in the presence of total mortality or respiratory mortality, coal intensity is insignificant, which suggests that the effect of coal intensity on height works through its influence on the disease environment, and specifically though respiratory infections.

*5.2 Instrumental variables analysis*

As noted earlier the effect of death rates on height is interpreted as the influence of the disease environment. Here we explore the effect of coal intensity on height running explicitly through the disease channel by means of instrumental variables regressions. Table 11 reports first and second stage regressions. As before, the individual and household variables are included in the regressions but the coefficients are not reported. In columns (1) and (2) coal intensity is the only excluded instrument at the second stage while in columns (3) and (4) both coal intensity and population density are excluded instruments. The first two columns show that log coal intensity and population density both have highly significant positive effects on child mortality from all causes and on deaths from respiratory diseases respectively. Not surprisingly the F-statistics for excluded instruments indicate that these are highly relevant determinants of mortality rates.

---

[22] This is not a quirk of the sample of servicemen; the correlation across the 635 registration districts is 0.93.

[23] It is likely that air pollution, as an additional stress factor, also contributed to deaths from other causes (Anderson, 2009).

The second stage coefficients are reported in the lower panel of Table (11). Both the overall death rate and the death rate from respiratory disease take significantly negative coefficients in columns (1) and (2) respectively. Population density, which is not excluded from the second stage, is insignificant. This suggests that much of its effect is due to the concentration of population in heavily polluted environments. In columns (3) and (4), where population density is an excluded instrument, the over-identification test indicates that this restriction is valid. The coefficients on the overall mortality and the death rate from respiratory diseases remain negative and strongly significant. The coefficients are slightly more negative than the OLS equivalents in columns (1) and (2) of Table 10. But it is worth noting also that the effect of coal intensity on height is consistent with the reduced form estimates presented earlier. The coefficient estimate in column (4) of Table 3 is $-0.403$ whereas the comparable estimate from column (2) of Table 10 is $-1.437 \times 0.280 = -0.402$. Thus while death rates may be a crude proxy for the contribution of atmospheric pollution to the disease environment the results are consistent with this channel of influence.

## 6. Conclusion

In this paper we have investigated the effects atmospheric pollution on the health and development of children growing up amid the environmental consequences of the industrial revolution that was founded upon coal. While a growing literature has identified the adverse health outcomes of atmospheric pollution in recent times, much less attention has been paid to the much heavier pollution of the past. Indeed the focus of studies of the nineteenth and early twentieth century has been almost exclusively on the effects of investment in water and sewage systems. And while these studies have estimated the effects of local conditions on mortality, much less attention has been paid to the ongoing health outcomes, as represented by the heights of those that survived polluted environments during childhood.

In this paper we used the occupational structure of a diverse range of localities in Britain at the turn of the twentieth century to derive an indicator for local coal combustion. The results presented here suggest that local atmospheric pollution had significant negative effects on the adult heights of those that grew up in crowded industrial cities where the presence of black smoke was orders of magnitude greater than today. Although our measure of industrial pollution is indirect, the results are plausible and they survive a number of robustness checks. In particular we show that the results are not simply due to occupational selection or geographic

mobility. In addition we find evidence consistent with the channel of influence running from coal intensity to respiratory illness and thence to height.

In the last century industrial pollution has declined massively as the result of the switch away from coal, efficiency gains in fossil fuel use, and increasingly strict regulation of emissions. This trend is likely to have contributed to the substantial gains in health and in stature that have been observed, although its precise contribution remains elusive. From the cohort of men born in the late 1890s to the cohort of the late 1970s adult male heights in Britain increased by 2.75 inches (Hatton and Bray 2010). If the difference between the most and the least polluted areas in 1901 is any guide to the possible contribution over time, then the decline in atmospheric pollution could have accounted for as much as one third of the long-term increase in height.

## References

Air Quality Expert Group (2005), *Particulate Matter in the United Kingdom*, London: UK Department of Environment, Food and Rural Affairs.

Almond, D., Edlund, L. and Palme, M. (2009), "Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden," *Quarterly Journal of Economics,* 124, pp. 1729-72

Anderson, H. R. (2009), "Air Pollution and Mortality: A History," *Atmospheric Environment*, 43, pp. 142-152.

Aplin, K. L. (2012), "Smoke Emissions from Industrial Western Scotland in 1859 Inferred from Lord Kelvin's Atmospheric Electricity Measurements," *Atmospheric Environment,* 50, pp. 373-376.

Bailey, R. E., Hatton, T. J. and Inwood, K. (2016), "Health, Height and the Household at the Turn of the Twentieth Century," *Economic History Review*, 69, pp 35-53.

Bateson, T. F. and Schwartz, J. (2008), "Children's Response to Air Pollutants," *Journal of Toxicology and Environmental Health*, 71, pp. 238-243.

Beach, B. and Hanlon, W. W. (2016), "Coal Smoke and Mortality in an Early Industrial Economy," University of California, Los Angeles: unpublished paper.

Bobak, M., Richards, M. and Wadsworth, M. (2004), "Relation between Children's Height and Outdoor Air Pollution from Coal-burning Sources in the British 1946 birth cohort," *International Archives of Occupational Environmental Health*, 77, pp. 383–386.

Bodenhorn, H., Guinnane, T. W. and Mroz, T. A. (2013), "Problems of Sample-selection Bias in the Historical Heights Literature: A Theoretical and Econometric Analysis," Yale Economics Working Paper No. 114.

Bodenborn, H, T W Guinnane, and T Mroz (2015), "Sample-selection Biases and the 'Industrialization Puzzle'", NBER Working Paper No. 21249.

Boyer G. R. and Hatton, T. J. (1997), "Migration and Labour Market Integration in Late Nineteenth-Century England and Wales," *Economic History Review*, 50, pp. 697–734.

Bozzoli, C., Deaton, A. S. and Quintana-Domeque, C. (2009), "Adult Height and Childhood Disease," *Demography,* 76, pp. 647-669.

Brimblecombe, P. and Bowler, C. (1992), "The History of Air Pollution in York, England," *Journal of the Air and Waste Management Association*, 42, pp. 1562-1566.

Burnett, R. T., Pope C. A. III, Ezzati, M., Olives, C., Lim, S. S. et al. (2014), "An Integrated Risk Function for Estimating the Global Burden of Disease Attributable to Ambient Fine Particulate Matter Exposure," *Environmental Health Perspectives,* 122, pp. 397-403

Cain, L. P. and Rotella, E. (2001), "Death and Spending: Urban Mortality and Municipal Expenditure on Sanitation," *Annales de Démographie Historique,* 101, pp. 139-154.

Chapman, J. (2016), "Local Government Investment in Urban Infrastructure and Mortality Decline in England and Wales, 1861-1900," Paper presented to the Economic History Society Conference, Cambridge.

Cesur, R., Tekin, E. and Ulker, A. (2016), "Air Pollution and Infant Mortality: Evidence from the Expansion of Natural Gas Infrastructure," *Economic Journal* (online first).

Chay, K. Y. and Greenstone, M. (2003), "Air Quality, Infant Mortality, and the Clean Air Act of 1970," NBER Working Paper No. 10053.

Clay, K. and Troesken, W. (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*, Chicago: University of Chicago Press.

Clay, K., Lewis J. and Severnini, E. (2016), "Canary in a Coal Mine: Impact of Mid-20th Century Air Pollution on Infant Mortality and Property Values," NBER Working Paper 22155

Cohen, A. J., Anderson, H. R., Ostro, B., Pandey, K. D., Krzyzanowski, M. et al. (2004), "Urban Air Pollution," in M. Ezzati, A. D. Lopez, A. Rodgers and C. J. L. Murray (eds.) *Comparative Quantification of Health Risks,* Geneva: World Health Organization.

Cook, L. M. (2003), "The Rise and Fall of the *Carbonaria* form of the Peppered Moth," *Quarterly Review of Biology*, 78, pp. 399-417.

Crafts, N. and Wolf, N. (2014), "The Location of the UK Cotton Textiles Industry in 1838: A Quantitative Analysis," *Journal of Economic History,* 74, pp. 1103-1139.

Currie, J., Graff Zivin, J. S., Mullins, J. and Neidell, M. J. (2013), "What do we Know about Short and Long Term Effects of Early Life Exposure to Pollution? NBER Working Paper 19571.

Currie, J. and Neidell, M. (2005), "Air Pollution and Infant Health: What can we Learn from California's Recent Experience?" *Quarterly Journal of Economics*, 120, pp. 1003-1030.

Cutler, D. and Miller, G. (2005), "The Role of Public Health Improvements in Health Advances: The Twentieth Century United States," *Demography*, 42, pp. 1-22.

Engels, F. (1892), *The Condition of the Working Class in England in 1944*, London: Allen and Unwin.

Fernihough, A. and O'Rourke, K. H. (2014), "Coal and the European Industrial Revolution," University of Oxford, Economic and Social History Discussion Paper 124.

Ferrie, J. P. and Troesken, W. (2008), "Water and Chicago's Mortality Transition, 1850–1925," *Explorations in Economic History*, 45, pp. 1–16.

Ghosh, R., Amirian, E., Dostal, M., Sram, R. J. and Hertz-Piccioto, I. (2011), "Indoor Coal Use and Early Childhood Growth," *Archives of Pediatric and Adolescent Medicine*, 165, pp. 492-497.

Grieves, K. (1988), *The Politics of Manpower, 1914-1918*, Manchester: Manchester University Press.

Hanlon, W. W. (2014), "Endogenous City Disamenities: Lessons from Industrial Pollution in 19th Century Britain," Unpublished paper, University of California Los Angeles.

Hanlon, W. W. and Tian, Y. (2015), "Killer Cities: Past and Present," *American Economic Review, Papers & Proceedings*, 105, pp. 570–57.

Hardy, A. (1994), "'Death is the Cure of All Diseases': Using the General Register Office Cause of Death Statistics for 1837-1920," *Social History of Medicine*, 7, pp. 472-492.

Harrison, R. G. (2006), "Urban Smoke Concentrations at Kew, London, 1898-2004," *Atmospheric Environment,* 40, pp. 3327-3332.

Hatton, T. J. (2011), "Infant Mortality and the Health of Survivors: Britain, 1910-1940," *Economic History Review*, 64, pp. 951-972.

Hatton, T. J. (2016), "Stature and Sibship: Historical Evidence," *History of the Family* (online first).

Hatton, T. J. and Bray. B. E. (2010), "Long Run Trends in the Heights of European Men, 19th–20th centuries," *Economics and Human Biology*, 8, pp. 405-413.

Heblich, S., Trew, A. and Zylberberg, Y. (2016), "East Side Story: Historical Pollution and Persistent Neighbourhood Sorting," London School of Economics, SERC Discussion Paper 208.

Heinrich, J. and Slama, R. (2007), "Fine Particles, a Major Threat to Children," *International Journal of Hygiene and Environmental Health*, 210, pp. 617–622.

Isen, A., Rossin-Slater, M. and Walker, M. (2016), "Every Breath You Take – Every Dollar You'll Make: The Long-Term Consequences of the Clean Air Act of 1970," *Journal of Political Economy* (forthcoming).

Janke, K., Propper, C. and Henderson, J. (2009), "Do Current Levels of Air Pollution Kill? The Impact of Air Pollution on Population Mortality in England," *Health Economics*, 18, pp. 1031-1055.

Jayachandran, S. (2009), "Air Quality and Early-Life Mortality: Evidence from Indonesia's Wildfires," *Journal of Human Resources*, 44, pp. 916-954.

Jedrychowski, W., Maugeri, U. and Jedrychowska-Bianchi, I. (2002), "Body Growth Rate in Preadolescent Children and Outdoor Air Quality," *Environmental Research* (Section A), 90, pp. 12-20.

Kyu, H. H., Georgiades, K. and Boyle, M. H. (2009), "Maternal Smoking, Biofuel Smoke Exposure and Child Height-for-age in Seven Developing Countries," *International Journal of Epidemiology*, 38, pp. 1342–1350.

Long, J. (2005), "Rural-Urban Migration and Socioeconomic Mobility in Victorian Britain," *Journal of Economic History*, 65, pp. 1-3

Millward, R. and Bell, F. N. (1998), "Economic Factors in the Decline of Mortality in Late Nineteenth Century Britain," *European Review of Economic History*, 2, pp. 263-288.

Ministry of National Service (1920), *Report upon the Physical Examination of Men of Military Age by National Service Medical Boards from November 1st 1917-October 31st, 1918*, Vol. I, (Cmd. 504), London: HMSO.

Mishra, V. and Retherford, R. D. (2007), "Does Biofuel Smoke Contribute to Anaemia and Stunting in Early Childhood?" *International Journal of Epidemiology*, 36, pp. 117–129.

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

Mosley, S. (2008), *The Chimney of the World: A History of Smoke Pollution in Victorian and Edwardian Manchester,* Abingdon: Routledge.

Nilsson, P. J. (2009), "The Long-run effects of Early Childhood Lead Exposure: Evidence from the Phase-out of Leaded Gasoline," Unpublished paper, Uppsala University.

Ransom, M. R. and Pope, C. A. (1995), "External Health Costs of a Steel Mill," *Contemporary Economic Policy*, 13, pp. 86-97.

Registrar General of England and Wales (1907), *Supplement to the Sixty-fifth Annual Report* London: HMSO.

Registrar General of England and Wales (1920), *Supplement to the Seventy-Fifth Annual Report*, Part IV, London: HMSO.

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*, 23, pp. 555–592.

Schwartz, J, (2004), "Air Pollution and Children's Health," *Pediatrics*, 113 pp. 1037-43.

Simkins, P. (1988), *Kitchener's Army: The Raising of the New Armies, 1914-1916*, London: Pen and Sword.

Szreter, S. (2005), *Health and Wealth: Studies in History and Policy*, Rochester NY: Rochester University Press.

Tanaka, S. (2015), "Environmental Regulations on Air Pollution in China and their Impact on Infant Mortality," *Journal of Health Economics*, 42, pp. 90-105.

Thorsheim, P. (2006), *Inventing Pollution: Coal, Smoke, and Culture in Britain since 1800,* Athens OH: University of Ohio Press.

Troesken, W. (2002), "The Limits of Jim Crow: Race and the Provision of Water and Sewerage Services in American Cities," *Journal of Economic History*, 62, pp. 734-773.

Winter, J. M. (2004), *The Great War and the British People* (2nd Edn.), London: Palgrave Macmillan.

Wohl, A. S. (1983), *Endangered Lives: Public Health in Victorian Britain*, London: Methuen.

**Table 1: Individual, and household characteristics of the sample of servicemen**

| Individual characteristics | Mean | S.D | Household characteristics | Mean | S.D |
|---|---|---|---|---|---|
| Height (inches) | 66.05 | 2.57 | Social class 1 or 2 | 0.13 | 0.34 |
| Attestation date | 1915.5 | 1.36 | Female head | 0.06 | 0.24 |
| Age on attestation | 20.26 | 2.29 | Sibship size | 4.16 | 2.07 |
| | | | Rooms occupied | 4.14 | 1.06 |

Sources: see text.

**Table 2: Locality characteristics**

| Locality characteristics | Mean | S.D |
|---|---|---|
| Coal intensity (tons per annum) | 2.49 | 1.29 |
| Heavy coal use (percent of employed) | 4.37 | 5.28 |
| Population (000s) | 148.7 | 130.0 |
| Population density (per 100 acres) | 20.81 | 36.72 |
| Under 5 mortality (percent) | 5.79 | 1.63 |

Source: see text.

**Table 3: Height and district coal intensity**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Age < 18 | -0.816*** | -0.797*** | -0.816*** | -0.815*** |
|  | (-5.39) | (-5.21) | (-5.39) | (-5.42) |
| Age 18 | -0.351** | -0.346** | -0.351** | -0.363*** |
|  | (-2.51) | (-2.48) | (-2.51) | (-2.61) |
| Age 19 | -0.360** | -0.358** | -0.360** | -0.377*** |
|  | (-2.49) | (-2.47) | (-2.49) | (-2.60) |
| Sibship size | -0.071*** | -0.072*** | -0.071*** | -0.071*** |
|  | (-2.86) | (-2.88) | (-2.86) | (-2.81) |
| Female household head | -0.264 | -0.279 | -0.264 | -0.253 |
|  | (-1.08) | (-1.15) | (-1.08) | (-1.04) |
| Household head social class 1 or 2 | 0.554*** | 0.568*** | 0.554*** | 0.566*** |
|  | (3.59) | (3.67) | (3.58) | (3.68) |
| Number of rooms | 0.169*** | 0.170*** | 0.169*** | 0.153*** |
|  | (2.93) | (2.99) | (2.93) | (2.69) |
| Log district coal intensity | -0.392*** |  | -0.388*** | -0.403*** |
|  | (-3.16) |  | (-2.65) | (-3.26) |
| District heavy coal use |  | -0.026** | -0.001 |  |
|  |  | (-2.13) | (-0.05) |  |
| District population density in 1901 |  |  |  | -0.539** |
|  |  |  |  | (-2.58) |
| F-Statistic | 13.16 | 12.48 | 12.74 | 12.85 |
| R-squared | 0.072 | 0.069 | 0.072 | 0.074 |
| Observations | 2,235 | 2,235 | 2,235 | 2,235 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Reference categories: age 20 over; social classes 3, 4 and 5. Regional dummies are included but not reported.

**Table 4: Height and pollution diffusion**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log district coal intensity | -0.341** | -0.361*** | -0.491*** | -0.369*** |
|  | (-2.54) | (-2.76) | (-3.41) | (-2.70) |
| District population density in 1901 | -0.510** | -0.534** | -0.550*** | -1.257*** |
|  | (-2.50) | (-2.56) | (-2.75) | (-2.80) |
| Log coal intensity in districts to the southwest | -0.080** |  | -0.106** | -0.069* |
|  | (-2.18) |  | (-2.03) | (-1.86) |
| Log coal intensity in districts to the northeast | 0.007 |  | 0.023 | 0.015 |
|  | (0.18) |  | (0.52) | (0.41) |
| Heavy coal use in districts to the southwest |  | -0.036** |  |  |
|  |  | (-2.32) |  |  |
| Heavy coal use in districts to the northeast |  | 0.013 |  |  |
|  |  | (0.78) |  |  |
| F-statistic | 11.87 | 12.29 | 11.09 | 13.30 |
| R-squared | 0.075 | 0.076 | 0.081 | 0.088 |
| Observations | 2,235 | 2,235 | 1,915 | 1,897 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported. Column (3) excludes those residing in 167 agricultural districts; column (4) excludes all those in 30 London boroughs.

**Table 5: The effects of local climate and topography**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log district coal intensity | -0.434*** | -0.384*** | -0.386*** | -0.357*** |
|  | (-3.44) | (-3.20) | (-3.17) | (-2.95) |
| Population density | -0.537** | -0.681*** | -0.543** | -0.597*** |
|  | (-2.58) | (-3.15) | (-2.58) | (-2.90) |
| District on coast | -0.164 |  |  |  |
|  | (-0.93) |  |  |  |
| District average wind speed | -0.074 |  |  |  |
|  | (-0.68) |  |  |  |
| District temperature |  | 0.333** |  |  |
|  |  | (1.99) |  |  |
| District rainfall |  |  | -0.151 |  |
|  |  |  | (-1.53) |  |
| District ruggedness |  |  |  | -0.098** |
|  |  |  |  | (-2.28) |
| F-statistic | 11.59 | 12.75 | 12.81 | 13.24 |
| R-squared | 0.075 | 0.076 | 0.075 | 0.076 |
| Observations | 2,235 | 2,235 | 2,235 | 2,335 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported.

**Table 6: Height and household head's occupation**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log district coal intensity | -0.366*** | -0.382*** | -0.383*** | -0.339*** |
|  | (-2.83) | (-2.96) | (-3.10) | (-2.60) |
| District population density in 1901 | -0.543** | -0.547*** | -0.540** | -0.544*** |
|  | (-2.58) | (-2.60) | (-2.58) | (-2.59) |
| Log household head's occupational coal intensity | -0.054 |  |  | -0.058 |
|  | (-1.06) |  |  | (-1.13) |
| Household head heavy coal use |  | -0.123 |  |  |
|  |  | (-0.71) |  |  |
| Log district coal intensity × w/collar household head |  |  | -0.189 | -0.224 |
|  |  |  | (-0.69) | (-0.80) |
| F-statistic | 12.44 | 12.45 | 12.20 | 11.84 |
| R-squared | 0.075 | 0.074 | 0.074 | 0.075 |
| Observations | 2,235 | 2,235 | 2,235 | 2,235 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported.

**Table 7: Height and migration of household head**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log district coal intensity | -0.402*** | -0.397*** | -0.500*** | -0.451*** |
|  | (-3.26) | (-3.27) | (-3.79) | (-3.52) |
| Population density | -0.549*** | -0.589*** | -0.447* | -0.445* |
|  | (-2.67) | (-2.86) | (-1.94) | (-1.94) |
| Head born in different district | 0.216* |  |  |  |
|  | (1.85) |  |  |  |
| Head born in Different county |  | 0.373*** |  |  |
|  |  | (2.97) |  |  |
| Difference in log coal intensity from birth district |  |  | 0.206* |  |
|  |  |  | (1.76) |  |
| Difference in heavy coal use from birth district |  |  |  | 0.009 |
|  |  |  |  | (0.66) |
| F-statistic | 12.51 | 12.88 | 12.52 | 12.49 |
| R-squared | 0.077 | 0.080 | 0.074 | 0.073 |
| Observations | 2,234 | 2,234 | 2,109 | 2,109 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported. Columns (3) and (4) exclude all individuals for whom the birthplace of the head of household was outside England and Wales.


**Table 8: Height and child migration**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log district coal intensity | -0.403*** | -0.402*** | -0.422*** | -0.336** |
|  | (-3.26) | (-3.25) | (-3.27) | (-2.51) |
| Population density | -0.539** | -0.540** | -0.556*** | -0.673*** |
|  | (-2.56) | (-2.59) | (-2.60) | (-2.71) |
| Individual born in different district | 0.009 |  |  |  |
|  | (0.06) |  |  |  |
| Individual born in Different county |  | 0.043 |  |  |
|  |  | (0.22) |  |  |
| Difference in log coal intensity from birth district |  |  | 0.111 |  |
|  |  |  | (0.46) |  |
| F-statistic | 12.36 | 12.24 | 11.98 | 8.42 |
| R-squared | 0.074 | 0.074 | 0.073 | 0.075 |
| Observations | 2,235 | 2,235 | 2,211 | 1,768 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported. Column (3) excludes all individuals not born in England and Wales; column (4) excludes all individuals not living in their district of birth.

**Table 9: Causes of death under five years, England and Wales 1891-1900**

| Cause of death | % of all < 5 deaths | Death rate per 1000 |
|---|---|---|
| Smallpox | 0.05 | 0.03 |
| Measles | 5.62 | 3.25 |
| Scarlet fever | 1.46 | 0.84 |
| Diphtheria | 2.36 | 1.36 |
| Whooping cough | 5.35 | 3.09 |
| Typhus | 0.00 | 0.00 |
| Enteric fever | 0.14 | 0.08 |
| Simple continued fever | 0.03 | 0.02 |
| Diarrhoea and dysentery | 9.24 | 5.34 |
| Cholera | 0.26 | 0.15 |
| Cancer | 0.05 | 0.03 |
| Tabes mesenterica | 2.53 | 1.46 |
| Phthisis | 0.72 | 0.41 |
| Other tuberculous and scrofulous diseases | 3.57 | 2.06 |
| Diseases of the nervous system | 12.30 | 7.10 |
| Diseases of the circulatory system | 0.34 | 0.20 |
| Diseases of the respiratory system | 21.35 | 12.33 |
| Diseases of the digestive system | 8.18 | 4.72 |
| Diseases of urinary system | 0.31 | 0.18 |
| Diseases of generative system | 0.03 | 0.02 |
| Violence | 2.34 | 1.35 |
| Other causes | 23.77 | 13.72 |
| Total | 100.00 | 57.74 |

Source: Calculated from Registrar General (1907) Table 7, p. 3.

**Table 10: Height and child mortality rates**

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Mortality rate under 5 years | -0.203*** (-4.22) | | | |
| Respiratory disease | | -0.797*** (-4.06) | -0.678*** (-2.66) | -0.606** (-1.98) |
| Gastrointestinal diseases | | | 0.149 (0.59) | |
| Infectious diseases | | | -0.476* (-1.83) | |
| All other identified diseases | | | | -0.090 (-0.70) |
| F-statistic | 13.33 | 13.75 | 12.83 | 12.94 |
| R-squared | 0.076 | 0.075 | 0.076 | 0.075 |
| Observations | 2,235 | 2,235 | 2,235 | 2,235 |

Note: $t$-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported.

**Table 11: Instrumental variables analysis of height and child mortality**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| *First stage* |  |  |  |  |
| Dependent variable | All cause death rate | Respiratory death rate | All cause death rate | Respiratory death rate |
| Log district coal intensity | 1.092*** | 0.280*** | 1.092*** | 0.280*** |
|  | (9.03) | (10.90) | (9.03) | (10.90) |
| District population density | 2.491*** | 0.658*** | 2.491*** | 0.658*** |
|  | (6.80) | (8.99) | (6.80) | (8.99) |
| R-squared | 0.659 | 0.714 | 0.659 | 0.714 |
| F—excluded instruments | 81.46 | 118.88 | 58.70 | 49.61 |
| *Second stage* |  |  |  |  |
| Dependent variable | Height | Height | Height | Height |
| All cause death rate | -0.369*** |  | -0.276*** |  |
|  | (-3.07) |  | (-4.20) |  |
| Respiratory death rate |  | -1.437*** |  | -1.054*** |
|  |  | (-3.13) |  | (-4.02) |
| District population density | 0.380 | 0.407 |  |  |
|  | (1.12) | (1.15) |  |  |
| R-squared | 0.071 | 0.071 | 0.075 | 0.074 |
| $X^2$—under-identification | 38.63 | 55.90 | 50.97 | 59.27 |
| $X^2$—over-identification |  |  | 1.27 | 1.31 |
| Observations | 2,235 | 2,235 | 2,235 | 2,235 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported. The F-test for excluded instruments is that of Angrist and Pischke; the under-identification test is the Kleinbergen-Paap LM statistic; and the over-identification test is the Hanson J statistic.
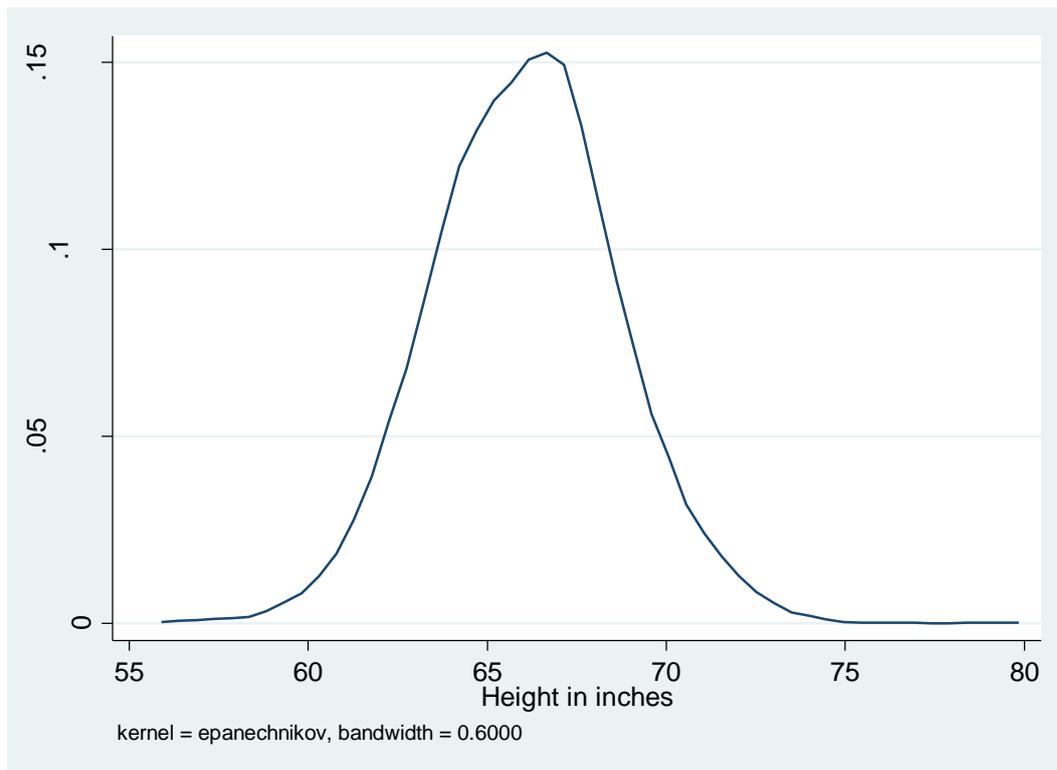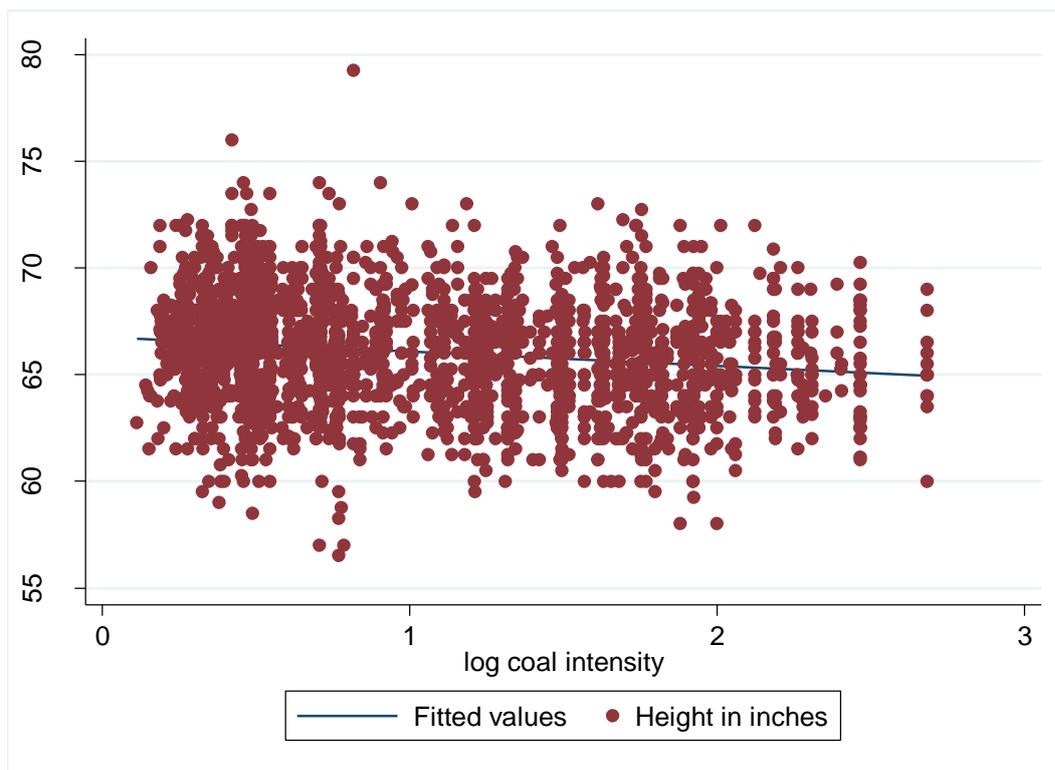
**Figure 1: Kernel density of height**



kernel = epanechnikov, bandwidth = 0.6000

**Figure 2: Map of coal intensity shaded by quartiles (darker for higher coal intensity)**

**Figure 3: Scatterplot of height and log coal intensity**

## Appendix: Supplementary Regressions

If there are spatial correlations across registration districts then this could bias the standard errors. Column (1) of Table A1 reports the regression result with no clustering of standard errors. The Cook-Weisberg test indicates no heteroscedasticity. Not surprisingly the standard errors are similar to those in column (2), which clusters at the district level (reproducing col. (4) of Table 3). As there is typically more than one observation in each registration district it is not possible to use standard spatial regression to test for error correlations between nearby districts. Instead column (3) clusters at the county level. The main difference is that the significance of coal intensity increases while that on population density declines. Finally, clustering at the region level further increases the significance of coal intensity and reduces that of population density.

**Table A1: Heteroskedasticity and clustering level of standard errors**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log district coal intensity | -0.403*** | -0.403*** | -0.403*** | -0.403*** |
|  | (-3.37) | (-3.26) | (-3.80) | (-4.54) |
| Population density | -0.539** | -0.539** | -0.539 | -0.539 |
|  | (-2.43) | (-2.58) | (-1.63) | (-1.54) |
| Clustering level (number of clusters) | None | Registration district (472) | County (45) | Region (11) |
| Cook-Weisberg ($\chi^2_{(1)}$) | 0.12 |  |  |  |
| R-squared | 0.074 | 0.074 | 0.074 | 0.074 |
| Observations | 2,235 | 2,235 | 2,235 | 2,235 |

Note: *significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported.

Table A2 explores different functional forms for district coal intensity and population density. In column (1) coal intensity is entered in quadratic form, with significant negative and positive coefficients respectively. So the relationship appears to be nonlinear. When a quadratic in the log of coal intensity is used the squared term is insignificant, although the linear term also loses significance. This is consistent with the concave relationship found in much of the literature between mortality (and other negative health outcomes) and the concentration of pollution (e.g. Cohen et al., 2004; Burnett et al., 2014). For population density, the squared term in column (3) is marginally significant. When the log of density and its square are included (col. 4), the latter is insignificant. However, when the squared term is dropped, the log of density is significant only at the 10 percent level. The RESET test indicates that there is little to choose between these specifications. The RESET statistic for the basic specification in col. (4) of Table 3 is 0.41.

**Table A2: Functional form of coal intensity and population density**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| District coal intensity | -0.246*** |  |  |  |
|  | (-2.97) |  |  |  |
| District coal intensity squared | 0.014** |  |  |  |
|  | (2.57) |  |  |  |
| Log district coal intensity |  | -0.729* | -0.411*** | -0.414*** |
|  |  | (-1.82) | (-3.31) | (-3.02) |
| Log district coal intensity squared |  | 0.128 |  |  |
|  |  | (0.91) |  |  |
| Population density | -0.557*** | -0.525** | -1.336** |  |
|  | (-2.65) | (-2.56) | (-2.27) |  |
| Population density squared |  |  | 0.474* |  |
|  |  |  | (1.77) |  |
| Log population density |  |  |  | -0.298** |
|  |  |  |  | (-2.25) |
| Log population density squared |  |  |  | -0.030 |
|  |  |  |  | (-1.65) |
| RESET ($F_{3, 2211}$) | 0.39 | 0.54 | 0.62 | 0.80 |
| F-statistic | 12.43 | 12.34 | 12.90 | 12.50 |
| R-squared | 0.074 | 0.074 | 0.075 | 0.075 |
| Observations | 2,235 | 2,235 | 2,235 | 2,235 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported.

Table A3 reports the results combining the influences on height of coal intensity in districts to the southwest and districts to the northeast and those representing climate and topography that were included separately in Tables 4 and 5 respectively. The coefficients on coal intensity in the district of residence and in the districts 25km to the southwest remain significant and are similar in size to those in column (1) of Table 4. The coefficients on climatic and topographic variables are similar in size and significance to those in Table 5, with the exception that rainfall now becomes significantly negative. Not surprisingly dropping the insignificant variable for coal intensity to the northeast (not shown) has almost no effect on the size and significance of the other coefficients. These results reinforce the finding that climate and topography matters for health and height but does not undermine the basic results for coal intensity.

**Table A3: Climate, terrain and diffusion**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Log district coal intensity | -0.355*** | -0.307** | -0.309** | -0.281** |
|  | (-2.67) | (-2.36) | (-2.38) | (-2.17) |
| Population density | -0.496** | -0.657*** | -0.510** | -0.570*** |
|  | (-2.46) | (-3.12) | (-2.49) | (-2.84) |
| Log coal intensity in districts to the southwest | -0.092** | -0.087** | -0.090** | -0.089** |
|  | (-2.47) | (-2.39) | (-2.48) | (-2.43) |
| Log coal intensity in districts to the northeast | -0.013 | -0.000 | 0.004 | 0.004 |
|  | (-0.34) | (-0.01) | (0.11) | (0.12) |
| District on coast | -0.271 |  |  |  |
|  | (-1.51) |  |  |  |
| District average wind speed | -0.061 |  |  |  |
|  | (-0.57) |  |  |  |
| District temperature |  | 0.358** |  |  |
|  |  | (2.18) |  |  |
| District rainfall |  |  | -0.181* |  |
|  |  |  | (-1.91) |  |
| District ruggedness |  |  |  | -0.107** |
|  |  |  |  | (-2.56) |
| F-statistic | 10.82 | 11.83 | 12.13 | 12.53 |
| R-squared | 0.077 | 0.078 | 0.077 | 0.078 |
| Observations | 2,235 | 2,235 | 2,235 | 2,235 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported.

In Table A4, regressions for height include coal intensity as well as the overall under-five mortality rate and under-five the death rate from respiratory disease that appear in Table 10. If these death rates adequately capture the disease environment arising from atmospheric pollution then coal intensity should not have a separate effect. The first and second columns show that overall mortality rate and the respiratory mortality rate retain significance. By contrast coal intensity is insignificant although the coefficient remains negative. When population density is added as in columns (3) and (4) that too takes negative but insignificant coefficients, while the coefficients on the mortality variables retain their significance. This suggests that the effect of coal intensity is mediated through its influence on the disease environment, as reflected in the respective mortality rates. The Sobel-Goodman test suggests that mediation accounts for between two-fifths and half of the total effect. This point is explored more formally in the IV analysis in Table 11.

**Table A4: Height child mortality rates, coal intensity and population density**

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Mortality rate under 5 years | -0.175*** | | -0.159** | |
|  | (-3.17) | | (-2.48) | |
| Respiratory disease | | -0.664*** | | -0.581** |
|  | | (-2.96) | | (-2.06) |
| Log district coal intensity | -0.210 | -0.215 | -0.230 | -0.240 |
|  | (-1.47) | (-1.51) | (-1.53) | (-1.57) |
| Population density | | | -0.144 | -0.157 |
|  | | | (-0.66) | (-0.62) |
| F-statistic | 12.75 | 13.16 | 12.28 | 12.64 |
| R-squared | 0.077 | 0.076 | 0.078 | 0.076 |
| Observations | 2,235 | 2,235 | 2,235 | 2,235 |
| Sobel-Goodman mediation test (%) | 46.4 | 45.2 | 43.0 | 40.4 |

Note: *t*-statistics from robust standard errors clustered at the registration district level; significance levels 1% ***, 5% ** and 10% *. Individual and household variables and regional dummies are included but not reported.