The Long-Term Effects of Early Life Pollution Exposure: Evidence from the London Smog

Stephanie von Hinke
Emil N Sørensen
Discussion Paper 22/757

18 February 2022

School of Economics

University of Bristol
Priory Road Complex
Bristol
BS8 1TU
United Kingdom

University of BRISTOL
School of Economics

The Long-Term Effects of Early-Life Pollution Exposure: Evidence from the London Smog

Stephanie von Hinke* Emil N. Sørensen†

February 16, 2022‡

Abstract

This paper uses a large UK cohort to investigate the impact of early-life pollution exposure on individuals' human capital and health outcomes in older age. We compare individuals who were exposed to the London smog in December 1952 whilst *in utero* or in infancy to those born after the smog and those born at the same time but in unaffected areas. We find that those exposed to the smog have substantially lower fluid intelligence and worse respiratory health, with some evidence of a reduction in years of schooling.

Keywords: London fog; Developmental origins; Heterogeneity; Social science genetics

JEL Classifications: I14, I18, I24, C21

*School of Economics, University of Bristol; Erasmus School of Economics, Erasmus University Rotterdam; Institute for Fiscal Studies. E-mail: S.vonHinke@bristol.ac.uk

†School of Economics, University of Bristol. E-mail: E.Sorensen@bristol.ac.uk

‡We would like to thank Samuel Baker, Pietro Biroli, Jason Fletcher, Hans van Kippersluis, Niels Rietveld, Nicolai Vitt, and seminar participants from the University of Bologna and the University of Wisconsin-Madison for valuable comments. This work is based on data provided through www.visionofbritain.org.uk and uses historical material which is copyright of the Great Britain Historical GIS Project and the University of Portsmouth. We gratefully acknowledge financial support from NORFACE DIAL (Grant Reference 462-16-100) and the European Research Council (Starting Grant Reference 851725).

1. INTRODUCTION

There is a growing literature on the contemporaneous effects of exposure to air pollution on individuals' human capital and health outcomes (for a review see e.g., Graff Zivin and Neidell, 2013). There is relatively little empirical evidence, however, on the much longer-term and cumulative effects of early-life pollution exposure, despite the fact that the early-life environment has been shown to be crucial in shaping individuals' health and economic outcomes in older age. The literature on the so-called "Developmental Origins of Health and Disease" (DOHaD) hypothesis – proposing that circumstances early in life can have life-long, potentially irreversible impacts on individuals' health and well-being – explores the importance of the prenatal as well as early childhood environment and has mainly focused on the longer-term effects of (adverse) nutritional, health and economic environments (for a review, see e.g., Almond and Currie, 2011a; Almond and Currie, 2011b; Almond et al., 2018; Conti et al., 2019).[1]

The relative lack of studies on the very long-term effects of early life pollution exposure is likely to be at least partially driven by a general absence of high quality historical pollution data. Indeed, most studies that explore the effects of early-life pollution exposure investigate the *immediate* effects on child birth outcomes (see e.g., Chay and Greenstone, 2003; Currie and Neidell, 2005; Almond et al., 2009; Currie, 2009; Jayachandran, 2009; Currie and Walker, 2011; Knittel et al., 2016; Sanders and Stoecker, 2015; Arceo et al., 2016; Hanlon, 2018; Jia and Ku, 2019; Rangel and Vogl, 2019), with only few exploring potential effects in childhood or early adulthood (see e.g., Reyes, 2007; Bharadwaj et al., 2017; Almond et al., 2009; Sanders, 2012; Black et al., 2013; Isen et al., 2017) and even fewer focusing on outcomes in older age (Bharadwaj et al., 2016; Ball, 2018a). As such, ignoring potential long-term effects

[1]For example, research has explored the importance of maternal physical health (Behrman and Rosenzweig, 2004; Almond, 2006; Almond and Mazumder, 2005), maternal mental health (von Hinke et al., 2019), maternal health behaviours (Nilsson, 2017; von Hinke et al., 2014), maternal nutrition (van den Berg et al., 2021), the economic environment (Van den Berg et al., 2006; Banerjee et al., 2010), the early life health environment (Bleakley, 2007; Case and Paxson, 2009; Cattan et al., 2021), or the home environment (Carneiro et al., 2015).

of pollution may lead to a substantial underestimation of the total welfare effects caused by exposure to environmental toxins.

We overcome the lack of historical pollution data by relying on reduced form analyses. More specifically, we examine the effect of early life exposure to the London smog: a severe pollution event that affected London residents between 5 and 9 December 1952. Although pollution levels in London are currently much lower than in the 1950s, the high levels recorded at the time are similar to the levels currently reported in industrialising economies such as India and China, so our study is relevant in particular to those settings. During the smog, pollution from residential and industrial chimneys, vehicle exhausts and coal burning became trapped under a layer of warm air due to a thermal inversion, which caused a thick smog to form over London. We investigate the long-term effects of exposure to this smog event by studying individuals' human capital and health outcomes in older age. The data we use is the UK Biobank: a large population-based cohort of approximately 500,000 individuals living in the United Kingdom. It includes rich data on individuals' later life health and economic outcomes, linked to administrative records. Using participants' eastings and northings of birth, our identification strategy exploits spatio-temporal variation in exposure to the London smog across birth dates and locations using a difference-in-difference approach. In other words, we compare individuals who were exposed to the smog in early life to those living in unaffected regions as well as to those conceived *after* the smog, whilst controlling for local area-specific trends in the outcomes of interest across birth cohorts.

This paper has three main contributions. First, most of the literature that explores the effects of pollution exposure focuses on child *birth* outcomes. Whilst it is important to better understand the effects on, e.g., infant mortality, it is one of the most extreme consequences of exposure to environmental toxins. Indeed, those who survive pollution exposure may be affected in other ways, with potential scarring reducing individuals' human capital and health potential. We investigate the effects on years of schooling, fluid intelligence, respiratory disease, and COVID-19 hospitalisations/mortality. Finding strong evidence of such longer-

term effects would therefore indicate that any pollution impacts are much larger than what would be suggested by the literature focusing on the effects of pollution on birth outcomes, such as infant mortality.

Second, we provide new empirical evidence in support of the DOHaD hypothesis. There is a large and growing literature in economics estimating the *causal* developmental origins of later life economic and health outcomes. These have shown the consequences of many adverse circumstances, but generally lack evidence on the longer-term effects of pollution exposure.[2] We contribute to this literature by exploring the very long-term effects of early life exposure to pollution, investigating individuals' outcomes at age ∼60. Our identification is similar to Bharadwaj et al. (2016) and Ball (2018a), who also focus on the long-term effects of the London smog on asthma and employment outcomes respectively. In fact, to our knowledge, these are the only two studies that investigate the effects of early-life pollution exposure on individuals in older age. An additional advantage of our data and setting is that it allows us to identify the gestational ages that are most sensitive to pollution.[3] This builds on studies examining the long-term effects of other relatively short-term events, such as the Ramadan (see e.g. Almond and Mazumder, 2011) and the Dutch Hunger Winter (Lumey et al., 2011; Bijwaard et al., 2021). In addition, our setting allows us to provide evidence on the human capital and health effects of pollution in a high pollution setting that is similar to current pollution levels in several industrialising countries, where evidence on the effects of pollution

[2]The literature focusing on the contemporaneous effects of pollution exposure mainly shows large impacts on respiratory and cardiovascular disease as well as mortality, but also on brain health and cognitive decline (see e.g., Zhang et al., 2018; Bishop et al., 2018). The literature suggests that high pollution concentrations affect lung function and cause irritation and inflammation of the respiratory system. Small pollution particles can penetrate deeply into the lung tissue and interfere directly with the transfer of oxygen to the blood. Both the elderly and the young are at increased risk of air pollution; the latter because their organs are still developing (EPA, 2021) and because they inhale more air per body mass than adults (Laskin, 2006). In addition, because small particles can be passed through the placenta to the developing foetus, this can directly affect the oxygen available to the foetus, and with that, its development.

[3]With just 42 individuals exposed *in utero* and 15 in infancy in Bharadwaj et al. (2016), sample sizes do not allow for the analysis of trimester-specific effects. Although Ball (2018a) uses large samples, the analyses use individuals' *year* of birth, implying it is not possible to look at gestation effects. An advantage of our data is that we have large samples as well as information on the year *and month* of birth. This means we have more power to detect even relatively small effects on later life outcomes, and we are able to explore the importance of exposure at different gestational ages.

remains limited (Greenstone and Jack, 2015).

Our third contribution is that we explore heterogeneity of treatment effects with respect to three important sources of variation. First, we build on a recent literature in social science genetics, directly modelling human capital and health outcomes as a function not only of individuals' environments ('nuture'), but also of their genetic predisposition to these outcomes ('nature') as well as the 'nature-nurture' interaction. This acknowledges the major role that genetic variation has been shown to play in shaping individuals' life outcomes (see e.g. Turkheimer, 2000; Polderman et al., 2015), and allows nature and nurture to interact and *jointly* contribute to individuals' human capital and health formation, as highlighted in the medical (see e.g., Rutter, 2006) as well as economics and social science literature (see e.g., Cunha and Heckman, 2007). Indeed, finding evidence of such 'gene-environment interplay' provides a strong argument against ideas of genetic (or environmental) determinism. As such, we examine whether – and to what extent – one's genetic variation can protect against, or exacerbate, the effects of such adverse events. There is a large literature investigating the importance of gene-by-environment interactions ($G \times E$), with relatively recent contributions from economics and social science (see, for example, Biroli, 2015; Bierut et al., 2018; Barth et al., 2020; Ronda, 2020). However, most existing studies tend to use endogenous environments, where it is not always clear how to interpret the main effects as well as the $G \times E$ interaction effect (Biroli et al., 2021).[4] We address this issue by exploiting the London smog as a natural experiment, ensuring that the environment is orthogonal to observed and unobserved individual characteristics. Hence, we add to only a handful of relatively recent studies that exploit *exogenous* variation in the environment within a $G \times E$ setting, allowing us to identify the causal environmental impact within a $G \times E$ framework.[5]

[4]Indeed, the coefficient on the genetic component may partially capture environmental circumstances due to 'genetic nurture' (that is: parental genotypes can shape the offspring environment, and since the offspring's genetic variation is inherited from the parents, this may partially capture such environments; see e.g., Belsky et al., 2018; Kong et al., 2018), and the coefficient on the environmental circumstances may pick up variation driven by genetics due to gene-environment correlation (that is: the fact that individuals with a genetic predisposition to a specific trait can be more commonly found in certain environments).

[5]Other studies that exploit exogenous variation in the environment include, e.g., Fletcher (e.g., 2012), Schmitz and Conley (2016a), Schmitz and Conley (2016b), Fletcher (2018), Barcellos et al. (2018), Pereira

The second source of heterogeneity we explore is gender. Since the literature suggests that male foetuses are generally frailer than female foetuses, we explore whether the long-term effects of pollution differ by gender due to either scarring or differential selection. Finally, we investigate whether there is a social gradient in the effect of pollution exposure. For this, we characterise the local area that individuals are born in with respect to the social class and run our analysis separately for individuals in high and low social class areas.

Our findings indicate large effects of both prenatal and childhood smog exposure on later life fluid intelligence and – to a slightly lesser extent – years of education. We also find a robust increase in the probability of being diagnosed with respiratory disease, but no differences in rates of COVID-19 hospitalization or mortality. Furthermore, these effects are generally larger for individuals exposed in the first and second trimester of pregnancy, with overall reduced effect sizes for those exposed in the last trimester.

Our heterogeneity analysis shows that the negative effects of being exposed to the smog prenatally and in early childhood are generally stronger for those with a high genetic predisposition to the outcome. For respiratory disease, for example, this suggests that the respiratory health of individuals who are genetically predisposed is more vulnerable to severe pollution events compared to the health of individuals how are not genetically predisposed. Furthermore, we show that the effect on years of schooling is driven by women, both for prenatal and early childhood exposure, whereas there is no clear gender-difference in the longer-term effects on fluid intelligence or respiratory disease. Using the gender ratio as the outcome, we find no evidence of gender differences in survival, suggesting that the differential gender effects are driven by scarring rather than selection.

Finally, we find a strong social gradient in long-term pollution effects, with individuals born in lower social class areas (as proxied by a high proportion of the population being either in semi-skilled or unskilled occupations) being substantially more affected; similar to e.g.,

et al. (2020), and Biroli and Zünd (2021), with Muslimova et al. (2020) exploiting exogenous variation in *both* genetic variation and environmental circumstances. See Pereira et al. (2021) for a recent review of this literature.

Jans et al. (2018). This in turn suggests either that the higher social classes were better able to avoid highly polluted areas, or that the health stock of lower class individuals is simply more vulnerable to adverse early life shocks. As Londoners at the time were not aware of the potential health risks of (severe) pollution, and there is little evidence of avoidance behaviour in the early 1950s, the former is perhaps less plausible, though we cannot say this with certainty.

The rest of the paper is structured as follows. Section 2 provides the background to the London smog and Section 3 describes the data used in our analysis. We set out the empirical strategy in Section 4, and discuss the results in Section 5. We explore the sensitivity of our findings in Section 6 and conclude in Section 7.

2. BACKGROUND: THE LONDON SMOG

On 4 December 1952, an anticyclone led to a temperature inversion over London, causing the cold air to be trapped under a layer of warm air. The resulting fog, in combination with higher than usual coal smoke (due to the slightly colder temperature at the time) from residential and industrial chimneys, the pollution from vehicle exhausts (e.g. steam locomotives, diesel-fuelled buses) and other pollution (e.g. coal-fired power stations), formed a thick smog.[6] With very little wind, it was not dispersed and led to an unprecedented accumulation of pollutants over the next five days, from 5–9 December 1952.

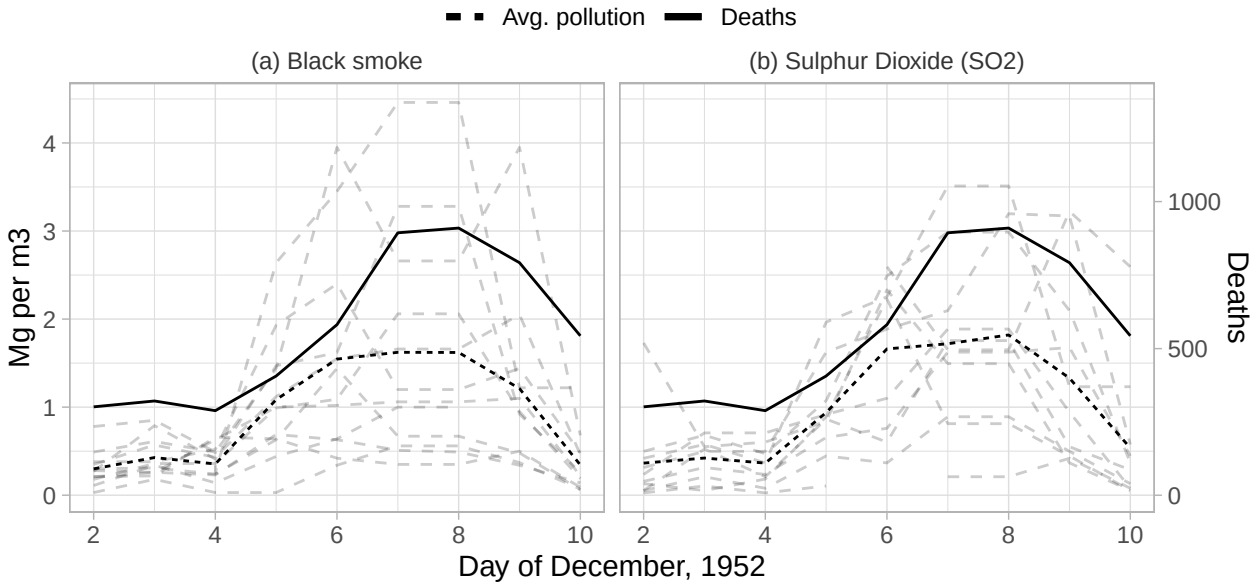
Wilkins (1954) discusses the severity of the London smog in terms of changes in concentrations of black smoke and sulphur dioxide (SO_2), with the historical measurements from that paper presented in Figure 1.[7] This shows two interesting features. First, there is a rapid rise in both black smoke and SO_2 concentrations between 5 and 9 December, with average concentrations rising to three to four times their usual level, after which they returned to

[6]The coal that was used domestically immediately after the war was of poor quality, with increased amounts of sulphur dioxide compared to the better quality coals that were mainly exported to pay off World War II debts.

[7]Both black smoke and SO_2 are released into the atmosphere via fuel combustion, such as coal burning.

pre-smog levels. Second, there is substantial regional variation in pollution within London, indicated by the grey dashed lines, each representing a different measurement station. Note, however, that the black smoke concentrations shown here are likely to be underestimated. Indeed, the smoke filters that measured the pollution were so overloaded that concentrations were more likely to be around 7-8 mg/m^3 in the worst polluted areas of London (Warren Spring Laboratory, 1967).

Figure 1: Pollution and mortality during the London smog of December 1952.



Historical measurements of pollution (black smoke and SO_2) from stations in London in December 1952. Each of the gray dashed lines represents the pollution measurements by a specific station. The dotted black line indicates the daily mean across all stations. The number of deaths in the Greater London area is overlaid with a solid black line. Pollution is digitised from Table I in Wilkins (1954) while deaths are digitised from Table VIII in Logan et al. (1953).

Levels of smoke and sulphur dioxide were measured at the time. However, as discussed in Wilkins (1954), it is likely that there were increases in tar, carbon monoxide (due to severe traffic congestion), carbon dioxide (due to a strong correlation with sulphur dioxide) and sulphuric acid (due to the oxidation of sulphuric dioxide).

Although Londoners were used to such smogs, the one in December 1952 was worse than any event Londoners had experienced before. Due to the dramatically reduced visibility, all

public transport other than the London Underground was suspended, most flights to London Airport were diverted, ambulance services stopped, and – with its penetration into indoor areas – concerts, theatres and cinema screenings were cancelled. Outdoor sporting events were also cancelled (see, e.g., BBC, 1952).

Despite this, Londoners got on with everyday life, potentially since the health consequences of extreme pollution were unknown. However, medical statistics that were published in the following weeks showed a substantial increase in mortality, with an estimated 4,000 deaths caused by the smog. Indeed, the right vertical axis of Figure 1 presents the daily number of deaths over the period of the smog, depicted as the solid black line. This shows around 300 daily deaths before the smog, increasing to ∼900 at its peak, after which is reduced; a similar inverse U-shaped pattern as the pollution data (Logan et al., 1953).

Subsequent calculations showed that 90% of the excess deaths were among those aged 45 and over (Ministry of Health, 1954). There was also an increase in mortality among newborns and infants, as well as foetal loss (Hanlon, 2018; Ball, 2018b), but these capture a relatively small proportion of the total increase. However, also in the months after the London smog, mortality exceeded normal levels.[8] About half of all excess deaths were attributed to bronchitis or pneumonia, with other increases observed in respiratory tuberculosis, lung cancer, coronary disease, myocardial degeneration and other respiratory disease (Logan et al., 1953).

3. DATA

Our primary dataset is the UK Biobank, a prospective, population-based cohort that contains detailed information on the health and well-being of approximately 500,000 individuals living in the United Kingdom. Recruitment and collection of baseline information occurred between

[8]Although an initial government report suggested these deaths were caused by influenza, there was no influenza outbreak in 1952, and Bell et al. (2004) find that only an extremely severe influenza epidemic could account for the excess deaths during this period. More recent analysis indeed suggests that the smog caused up to 12,000 deaths (Bell and Davis, 2001).

2006 and 2010, when participants were 40–69 years old. The data include information on demographics, physical and mental health, health behaviours, cognition, and economic outcomes, obtained via questionnaires, interviews, and measurement taken by nurses. It has also been linked to GP and hospital records, as well as the National Death Registry. Furthermore, samples of blood, urine and saliva have been collected, and all individuals have been genotyped. Bycroft et al. (2018) give a detailed description of the sample.

We are interested in the long term economic and health consequences of short term variation in the early-life pollution environment. We focus on a range of outcomes, informed by previous literature on the effects of pollution. First, we build on the literature that shows medium-to-long term effects of early life pollution exposure on economic outcomes (see e.g., Almond et al., 2009; Ball, 2018a), investigating the effects on educational attainment and fluid intelligence. Educational attainment is defined based on individuals' qualifications[9] , and fluid intelligence is a score based on problem solving questions that require logic and reasoning ability, independent of acquired knowledge.

Next, we build on the literature showing pollution effects on individuals' health (see e.g., Currie and Walker, 2011), and explore the effects on respiratory disease. The contemporaneous effects of air pollution on respiratory disease are well-known. Less is known, however, about the potential long term effects of early life exposure. Indeed, since air pollution disproportionally affects individuals with compromised lung function, and much of the burden in adulthood is believed to be due to poor development (rather than accelerated decline) in lung function (see e.g. Lancet, 2019), early life pollution is a natural exposure to consider in the development of respiratory disease. We create a dummy variable to indicate whether the individual has been diagnosed with respiratory disease from the administrative hospitalisation data and mortality records that have been merged into the UK Biobank, distinguishing between chronic and acute respiratory conditions.[10] Furthermore, due to the links between

[9]Table A.1 in Appendix A shows the mapping between qualifications and years of education, using a similar definition as in, e.g., Rietveld et al. (2013), Okbay et al. (2016), and Lee (2018).

[10]The hospitalisation data include all diagnoses in ICD-10 coding. We use ICD-10 J00-J99 to identify respiratory disease as diagnosis or cause of death. ICD-10 [J40–J47] and [J09, J1, J20–J22] are used to

respiratory disease and severe COVID-19 (Aveyard et al., 2021), we additionally use a binary indicator for being hospitalised with, or having died from, COVID-19.[11]

Using participants' eastings and northings of birth, we assign each individual one of the 1472 Local Government Districts of birth across England and Wales.[12] This spatial information, in combination with temporal information on individuals' year-month of birth, allows us to identify individuals who were exposed to the smog at different time points during the intrauterine and early childhood period. We split our sample along the time dimension by considering whether the prenatal period precedes, overlaps, or follows the smog event on December 5-9th, 1952. This allows us to define three groups: (i) those exposed to the smog during childhood (i.e., those born before the smog), (ii) those exposed to the smog *in utero*, and (iii) those conceived after the smog event and therefore not exposed.[13]

We split our sample along the spatial dimension by identifying the geographical areas in and around London that were exposed to high pollution during the smog. To do so, we overlay the reduced visibility and sulphur dioxide measurements from Wilkins (1954) onto a district-level shapefile. This is shown in Figure 2, where the solid black outlines indicate the areas with high and low reduced visibility and the dotted outline indicates the area with high sulphur dioxide measurements.[14] We define "high exposure" districts as

identify chronic and acute respiratory conditions, respectively.

[11]We use ICD-10 emergency codes U071 and U072 to identify COVID-19 related hospitalisations and deaths.

[12]Our districts are defined based on the 1951 shapefiles from Vision of Britain (Southall and Aucott, 2009).

[13]Note that we do not observe gestational age at birth. Hence, we assume that the prenatal period cover the nine months before the year-month of birth. The exact birth date cutoffs are as follows. Exposed in childhood: 1950-Dec to 1952-Nov. Exposed *in utero*: 1952-Dec to 1953-Aug. Not exposed: 1953-Sep to 1956-Dec. We drop those born after December 1956 for two reasons. First, depending on their month of birth in 1957, individuals may have been directly affected by an educational reform – the raising of the school leaving age – which has been shown to have affected individuals' longer-term education as well as health outcomes (see e.g. Harmon and Walker, 1995; Davies et al., 2018), though note that the evidence on the health effects are more mixed (see e.g., Clark and Royer, 2013). Second, the first Clean Air Act allowed local authorities to create Smoke Control Areas; areas that prohibited all smoke emissions. The first orders of such Smoke Control Areas were announced in 1957 (Fukushima, 2021). By dropping all births in 1957 onwards from our analysis, we avoid our estimates potentially capturing reductions in pollution due to the Smoke Control Areas.

[14]The sulphur dioxide boundary, based on Wilkins (1954), shows measurements from different stations with limited geographical coverage, resulting in a boundary with a sharp border,while the visibility boundary is based on observations recorded by the Meteorological Office at 9am and 6pm throughout the smog event

those that experienced severe reductions in visibility (i.e., overlap with the two inner solid boundaries in Figure 2) and/or experienced high sulphur dioxide measurements (i.e., overlap with the dotted boundary in Figure 2). Districts that only overlap with the outer solid boundary, indicating the mildest reduction in visibility, are classified as "low exposure". In our main analysis, we do not distinguish between the high and low exposure districts but instead refer to them jointly as "treated" districts. We compare these treated districts to a set of "control" districts that are defined as other urban districts in England and Wales with a population density exceeding 400 individuals per km^2. In our robustness checks, we explore the sensitivity of our results to control districts with different population densities, to excluding the "low exposure" districts, to assigning exposure based on individuals' reported birth *locations*, as well as by defining control districts as other major cities in England and Wales.[15]

Our sample selection process is as follows: we only consider the subsample of individuals born in the years 1950 to 1956, and restrict the sample to those born in either treated or control districts. Furthermore, we follow the (genetics) literature and restrict our sample to those of white European ancestry.[16] This leaves us with between 26,805–65,060 participants for the main analysis, depending on the outcome of interest.

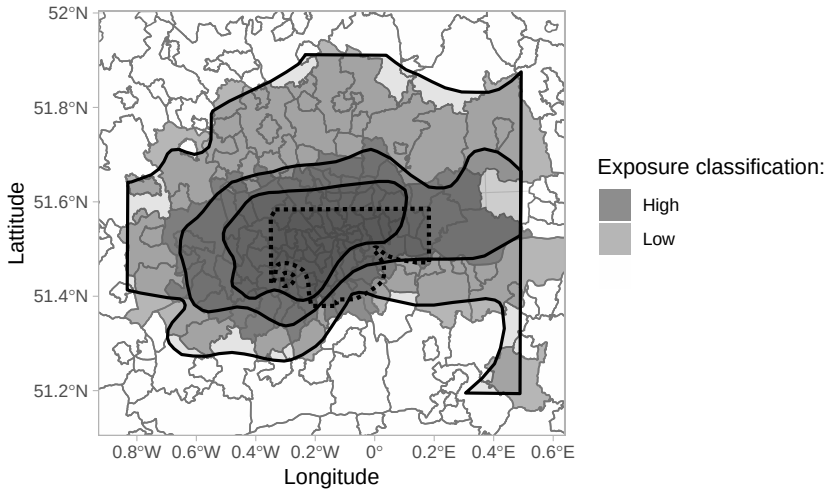
Given the potential importance of the weather for exposure to smog, we merge in an auxiliary dataset on ambient temperature, sunshine, and rainfall. These data are available from the MET Office in the form of an interpolated grid of measurements (MET Office, 2022). We use a grid resolution of 25km and assign measurements to individuals by linking their location of birth, as measured in eastings and northings, to its nearest grid point. We merge in the weather data at individuals' birth location for the period of the smog.

(Wilkins, 1954).

[15]The control districts used in the main analysis are shown on a map in Figure A.1, Appendix A, and colourised according to their population density. The data on population density is from Vision of Britain (Southall, 2011). Ball (2018a) argues that the only other city with unusually high pollution at the time of the London smog was Leeds. We therefore drop Leeds in all our analyses. All city boundaries are defined according to 1951 district shapefile.

[16]Because genetic variation differs by ancestry, this accounts for population stratification; a form of genetic confounding. We discuss the genetic data as well as its interpretation in more detail in Appendix D.

Figure 2: Visibility and pollution measurements during the London smog.



The geographic boundaries of the London smog based on the maps in Wilkins (1954). The solid black outlines show the areas with reduced visibility. The inner boundaries experienced a more severe reduction in visibility. The dotted outline shows the area with high sulphur dioxide measurements. The map classifies the districts into 'high exposure' (dark gray), 'low exposure' (light gray), and 'unexposed' (white) districts. City of London is approximately at the center of the map.

For ambient temperature, we assign the minimum temperature measured during the smog, while for sunshine and rainfall, we use the average. The weather measurements capture additional local conditions that, linked to individuals' eastings and northings of birth, vary *within* districts.[17]

Table 1 presents the descriptive statistics, showing that 44% of the sample is male and individuals, on average, have 13.3 years of education. Individuals' fluid intelligence is a continuous score, standardised to have mean zero and unit variance. 9.2% of our sample has been diagnosed with respiratory illness; for 7.6%, this is an acute condition, and for 1.5%, it is chronic. Finally, 0.7% of our sample has been either hospitalized or has died with COVID-19 as primary or secondary cause.

[17]Figure A.2 in Appendix A shows the monthly time series of weather conditions (temperature, sunshine and rainfall) for our sample of interest, distinguishing between the districts that are exposed to the London smog (treated) and the control districts. As mentioned in the introduction, this shows a slightly lower temperature in exposed districts at the time of the smog. However, the difference between treated and control districts is minimal (less than 0.5°C). Hence, these do not suggest any notable differences in weather conditions between the district types.

Table 1: Descriptive statistics for the main outcomes and variables.

	(1) Mean	(2) Std. dev.	(3) Obs.
Male	0.443	0.497	65,081
Educational attainment	13.318	2.274	64,702
Fluid intelligence	0.000	1.000	26,934
Respiratory disease	0.092	0.290	64,944
– Acute	0.076	0.266	64,941
– Chronic	0.015	0.120	64,941
COVID-19	0.007	0.081	65,081

Columns: (1) sample mean, (2) sample standard deviation, (3) number of observations. The availability of the variables varies and hence also the number of observations in column (3).

4. EMPIRICAL STRATEGY

To investigate the long term effects on human capital and health outcomes of early life pollution exposure, we exploit spatio-temporal variation in the exposure to the London smog across birth dates and locations using a difference-in-difference approach. We distinguish between those born inside and outside of the exposed London area (spatial variation) while also considering the timing of birth relative to the smog event (temporal variation). Our main specification is:

$$Y_{ijt} = \alpha_j + \gamma_t + \tau_k t + \beta_{IU} E_i^{IU} \times L_i + \beta_{CH} E_i^{CH} \times L_i + \delta \mathbf{X}_i + \epsilon_{ijt}, \tag{1}$$

where Y_{ijt} denotes the outcome of interest for individual i, born in district j at year t. Thus, α_j denote district fixed effects, and γ_t are year of birth fixed effect. We additionally include administrative-county-specific time (year-month) trends, denoted by $\tau_k t$.[18] The vector $\mathbf{X}_i$ in-

[18] We do not include district-specific trends in our main analysis, as with over 1400 districts, some only include few individuals. Instead, with ∼230 administrative counties, we observe a larger number of individuals in each administrative county-year and we include trends specific to these geographical regions. In our sensitivity analysis in Appendix B, however, we show that our results are generally robust to including administrative county-specific year (as opposed to year-month) trends, district-specific year or year-month trends, and not including any trends.

cludes weather conditions during the smog, gender, and month-of-birth dummies to account for weather effects, gender differences, and seasonality in the outcome. The indicators E_i^{IU} and E_i^{CH} are dummy variables that are equal to one for individuals who are exposed to the London smog *in utero* and in early childhood (i.e., <age 2) respectively, while L_i is a binary variable that indicates whether the individual is born in an area of London that was exposed to the smog (i.e., in treated districts). Hence, our identification strategy compares individuals' outcomes Y_{ijt} for those exposed at different ages (i.e., *in utero* and in early childhood) to those conceived after the London smog in treated districts, relative to others born at the same time, but in districts unaffected by the smog. The parameters of interest are therefore β_{IU} and β_{CH} which parameterise, respectively, the long-term effects of being exposed to the smog *in utero* and in early childhood, relative to non-exposed cohorts, accounting for any administrative county-specific trends in Y_{ijt}, time and district fixed effects. We use robust standard errors, clustered by district throughout.

An important issue in the above specification is potential foetal selection. Indeed, there is evidence of increased mortality among newborns and infants (Ministry of Health, 1954), as well as foetal loss (Hanlon, 2018; Ball, 2018b). Although these effects were relatively small as the smog mainly affected deaths among the elderly (Ministry of Health, 1954), they do affect our analysis and interpretation. More specifically, assuming that the smog increased mortality among relatively frail infants, leaving the stronger ones to survive, this may have led to an improvement in average cohort-level human capital and health outcomes for those exposed in the affected districts. This, in turn, suggests that our estimates may be underestimates of the effects of interest.

Our specification implicitly assumes that individuals who were born in districts that were affected by the London smog would have had similar trends in their outcomes of interest in the absence of the smog compared to those born in districts that were not affected. We explore this common trend assumption empirically in Appendix C. Since those born *prior to* the smog may have been exposed in early childhood, we should find common trends among

those conceived *after* the smog in treated and control districts. Indeed, we find no evidence to suggest that those born treated districts have differential trends in our outcomes of interest compared to those born control districts.

The fact that the London smog only lasted five days additionally allows us to explore the gestational ages that are most sensitive to pollution exposure. This identification exploits the fact that the period of pollution exposure is substantially shorter than the length of gestation. To do this, we replace E_i^{IU} in Equation 1 with three binary variables indicating the trimester in which the individual was exposed to the London smog.

5. RESULTS

5.1. The London smog

We start by examining the long-term impact of pollution exposure on human capital outcomes. Columns (1) and (2) of Table 2 show the estimates from Equation 1 for years of education and fluid intelligence respectively. We find no strong differences in the outcomes among individuals exposed to the London smog prenatally or in early childhood in control districts compared to those conceived afterwards. However, within treated districts, we find a reduction in fluid intelligence for those exposed *in utero* as well as in childhood. The latter show the largest negative effects, with 0.16 standard deviations lower fluid intelligence score compared to those born at the same time in control districts. Being exposed *in utero* reduces fluid intelligence by 0.11 standard deviations. Looking at the estimates for years of education, we find that they are of a similar magnitude, but they are not significantly different from zero at conventional levels.

We now turn to the long-term consequences of pollution on health outcomes. Table 3, Column (1) presents the estimates from Equation 1 for respiratory disease. This shows an increase in the prevalence of respiratory disease for those exposed to the smog prenatally. More specifically, those *in utero* during the smog and born in the exposed areas are 2 percentage points more likely to be diagnosed with respiratory disease compared to those conceived

Table 2: Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Dependent variable:	
	(1) Educational attainment	(2) Fluid intelligence
Treated × In utero	−0.100	−0.112**
	(0.089)	(0.051)
Treated × Childhood	−0.135	−0.158**
	(0.088)	(0.068)
In utero	−0.003	0.008
	(0.054)	(0.036)
Childhood	−0.014	−0.051
	(0.120)	(0.074)
Observations	64,681	26,877
R^2	0.08	0.067

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

after the smog. Considering that the overall incidence of respiratory disease is 9% in the sample, this is a large effect, similar to a 22% increase. Columns (2) and (3) distinguish between respiratory hospitalisations due to, respectively, acute and chronic causes. We find that acute conditions are the main drivers of hospitalisations, while there is no effect of in-trauterine pollution exposure on chronic respiratory conditions. Next, we examine whether the negative effect on respiratory disease translates into COVID-related deaths or hospitalizations in Column (4). With coefficients that are close to zero and with relatively large standard errors, we find no evidence of increased COVID-related morbidity or mortality.

Next, we examine the impact of the timing of exposure relative to individuals' gestational age. For brevity, we here only report the estimates for outcomes that indicated some sug-

Table 3: Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Dependent variable:			
	(1) Respiratory, any	(2) Respiratory, acute	(3) Respiratory, chronic	(4) COVID-19
Treated × In utero	0.020*	0.019**	−0.002	0.000
	(0.011)	(0.009)	(0.004)	(0.003)
Treated × Childhood	−0.007	−0.004	−0.005	−0.001
	(0.012)	(0.011)	(0.005)	(0.003)
In utero	−0.005	−0.009	0.004	0.000
	(0.006)	(0.006)	(0.003)	(0.002)
Childhood	−0.008	−0.015	0.008	−0.001
	(0.014)	(0.012)	(0.006)	(0.004)
Observations	64,923	64,920	64,920	65,060
R^2	0.018	0.018	0.013	0.013

Columns: (1) ever experienced a (primary) respiratory hospitalisation, (2)-(3) splits (1) into acute and chronic causes of respiratory hospitalisation, (4) occurence of hospitalisation or death due to COVID-19. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

gestion of intrauterine effects. Table 4 shows the estimates for Equation 1, where we replace the indicator for *in utero* exposure, E_i^{IU}, with three indicators for the relevant trimesters. For years of education, this suggests that second trimester exposure is the most important and reduces education by 0.2 years on average, with third trimester and childhood exposure also showing negative effects, though these are not significantly different from zero. For fluid intelligence, we find the largest effects for first trimester exposure, reducing slightly with gestational age. Similarly, the intrauterine effect on respiratory disease, shown in columns (3) and (4), is mainly driven by exposure in the first and, to a lesser extent, the second trimester.

Table 4: Trimester effects. Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Dependent variable:				
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any	(4) Respiratory, acute	(5) Respiratory, chronic
Treated × In utero, 1. tri.	0.064	−0.147**	0.032*	0.032**	0.000
	(0.134)	(0.073)	(0.018)	(0.014)	(0.007)
Treated × In utero, 2. tri.	−0.220*	−0.120*	0.019	0.022	−0.006
	(0.118)	(0.068)	(0.016)	(0.014)	(0.006)
Treated × In utero, 3. tri.	−0.144	−0.068	0.009	0.003	0.000
	(0.121)	(0.089)	(0.015)	(0.014)	(0.005)
Treated × Childhood	−0.143	−0.155**	−0.008	−0.005	−0.005
	(0.088)	(0.069)	(0.012)	(0.011)	(0.005)
In utero 1. tri.	−0.007	−0.009	−0.004	−0.008	0.005
	(0.063)	(0.047)	(0.008)	(0.007)	(0.003)
In utero 2. tri.	−0.049	0.006	0.003	−0.004	0.006
	(0.061)	(0.040)	(0.009)	(0.008)	(0.004)
In utero 3. tri.	0.084	0.042	−0.023***	−0.020***	−0.002
	(0.081)	(0.051)	(0.008)	(0.007)	(0.003)
Childhood	0.066	−0.008	−0.027*	−0.029**	0.002
	(0.134)	(0.080)	(0.015)	(0.013)	(0.007)
Observations	64,681	26,877	64,923	64,920	64,920
R^2	0.08	0.067	0.018	0.018	0.013

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation, (4)-(5) splits (3) into acute and chronic causes of respiratory hospitalisation. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

5.2. Treatment effect heterogeneity

We next explore potential heterogeneity of treatment effects. For this, we investigate three sources of variation: heterogeneity with respect to individuals' genetic predisposition, gender, and socio-economic status. We discuss each in turn.

5.2.1. Genetic heterogeneity

To directly incorporate the genetic component into the analysis, we construct variables measuring individuals' 'genetic predisposition' to the outcomes of interest. We do this by running our own tailor-made Genome-Wide Association Study (GWAS) for each of the outcomes on UK Biobank participants born in the years *outside* our analysis sample (i.e., 1934–1949 and

1957–1970), as well as those born in districts that are not defined as either treated or control districts during the study years 1950-1956.[19] We use the summary statistics from this GWAS to construct so-called polygenic scores (also known as polygenic indices), measuring one's 'genetic predisposition' to the relevant outcome, for those in the (independent) analysis sample covering the birth cohorts 1950–1956. We do the latter using LDpred2, a Bayesian genetic risk prediction method (Vilhjálmsson et al., 2015; Privé et al., 2020). For ease of interpretation, all polygenic scores are standardised to have mean zero and unit variance in the analysis sample.

To investigate the extent to which one's genetic variation may protect or exacerbate the effects of early life pollution exposure, Table 5 estimates the main difference-in-difference analysis distinguishing between individuals with a high versus low genetic predisposition to the outcome, defined as having a polygenic score above or below the median, shown in Panel (a) and (b) respectively. Note that the polygenic score is specific to the outcome of interest. For example, the polygenic score in Columns (1) and (2) of Table 5 is the best linear genetic predictor for education and fluid intelligence, respectively. This shows that the zero effect of the smog on educational attainment conceals substantial genetic heterogeneity. More precisely, the negative effect is substantially larger for those with a high polygenic score for education, for both prenatal and childhood exposure, with the estimates being close to zero for those with a polygenic score below the median.

To explore what may be driving the negative effect on educational attainment for those with a high polygenic score, Table A.2 in Appendix A examines the effects of pollution exposure on the probability of reaching different levels of qualifications. This shows that the negative effect of pre- and post-natal pollution exposure for those with a high polygenic score is driven by a reduction in the probability of obtaining an upper secondary qualification (i.e., A-levels, university/college degrees, and professional qualifications), and – correspondingly – a higher probability of exiting the education system with lower secondary qualifications

[19] See Appendix D for an introduction to genetics, an explanation of the genetic terms used here, as well as more detail on the construction of the 'genetic scores'.

(O-levels, CSEs, GCSEs), or no qualifications. This suggests that pollution exposure reduces the human capital potential, in particular among those with a high *genetic* potential.

There is little difference between the estimates for individuals with high and low polygenic scores for fluid intelligence, with both showing a negative effect of smog exposure, though with the smaller sample sizes, they are not significantly different from zero. Furthermore, we find that the effects of prenatal smog exposure on respiratory disease, in particular *acute* respiratory conditions, are larger for those with a high polygenic score, suggesting that the respiratory health of individuals who are genetically predisposed is more vulnerable to severe pollution events.[20]

5.2.2. Heterogeneity by gender

We next investigate whether the effect of exposure to the smog is similar for men and women. Table 6 presents the estimates for our main outcomes of years of education, fluid intelligence and respiratory disease, with Panel (a) and (b) presenting the estimates for women and men, respectively. This shows that the negative effect of smog exposure on years of education in the full sample is largely driven by women, with much smaller effect estimates for men. In particular, women who were exposed to the smog prenatally and in childhood have 0.12 and 0.24 fewer years of education respectively, compared to those not exposed and relative to women born in control areas. For fluid intelligence and respiratory disease, we do not see large differences between men and women, though the estimates are not always significantly different from zero due to the reduced sample sizes and with that, larger standard errors.[21]

5.2.3. Heterogeneity by socio-economic status

Finally, we explore potential treatment effect heterogeneity with respect to socio-economic status. Although the UK Biobank does not include data on individuals' (or parental) socio-

[20] Our genetic heterogeneity analysis is robust to the use of polygenic scores constructed from an alternative GWAS, obtained from the polygenic index repository (Becker et al., 2021).

[21] In Table A.3, we model the gender ratio as the outcome of interest to explore whether the smog caused differential mortality by gender. We find no gender differences, suggesting that the differential effects in Table 6 are driven by scarring rather than selection.

Table 5: Heterogeneity across genetics. Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Dependent variable:				
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any	(4) Respiratory, acute	(5) Respiratory, chronic
Panel (a) – High polygenic score					
Treated × In utero	−0.249**	−0.090	0.019	0.034**	−0.007
	(0.107)	(0.073)	(0.017)	(0.015)	(0.007)
Treated × Childhood	−0.197*	−0.153	−0.015	0.002	−0.011
	(0.118)	(0.095)	(0.019)	(0.018)	(0.008)
In utero	−0.010	0.020	−0.016*	−0.016*	0.009**
	(0.076)	(0.050)	(0.010)	(0.008)	(0.004)
Childhood	−0.168	−0.054	−0.007	−0.012	0.011
	(0.164)	(0.095)	(0.020)	(0.020)	(0.010)
Panel (b) – Low polygenic score					
Treated × In utero	0.096	−0.087	0.022	0.005	0.003
	(0.152)	(0.084)	(0.014)	(0.013)	(0.005)
Treated × Childhood	−0.010	−0.113	0.007	−0.004	0.001
	(0.161)	(0.104)	(0.016)	(0.014)	(0.005)
In utero	−0.016	−0.027	0.007	−0.002	−0.002
	(0.068)	(0.047)	(0.008)	(0.008)	(0.004)
Childhood	0.111	−0.037	−0.006	−0.015	0.006
	(0.152)	(0.113)	(0.018)	(0.016)	(0.006)

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation, (4)-(5) splits (3) into acute and chronic causes of respiratory hospitalisation Panels: (a) subsample with above-median polygenic score. (b) subsample with below-median polygenic score. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

economic position at birth, and because individuals' socio-economic position in adulthood is endogenous to the smog exposure, we merge the 1951 UK Census to the UK Biobank, allowing us to characterize the local area of birth in terms of its socio-economic composition relative to other areas in England and Wales. As such, we create measures of social class at the district level, indicating the share of different social classes defined according to individuals' occupation. This allows us to classify individuals as being born into different social class environments. To identify 'high social class' districts, we focus on districts with an above-median share of social classes I and II (including professional, managerial, and

Table 6: Heterogeneity across sex. Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Dependent variable:		
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any
Panel (a) – Female			
Treated × In utero	−0.118	−0.074	0.020
	(0.128)	(0.072)	(0.013)
Treated × Childhood	−0.238*	−0.152*	−0.010
	(0.130)	(0.083)	(0.015)
In utero	0.077	−0.028	−0.004
	(0.074)	(0.048)	(0.008)
Childhood	0.127	−0.064	−0.012
	(0.149)	(0.114)	(0.018)
Panel (b) – Male			
Treated × In utero	−0.084	−0.163*	0.022
	(0.113)	(0.085)	(0.017)
Treated × Childhood	−0.022	−0.185*	−0.003
	(0.135)	(0.105)	(0.019)
In utero	−0.080	0.055	−0.007
	(0.077)	(0.051)	(0.010)
Childhood	−0.159	−0.026	−0.004
	(0.175)	(0.103)	(0.023)

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation. Panels: (a) female subsample, (b) male subsample. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

intermediate occupations). Similarly, we use districts with above-median shares of social classes IV and V to identify 'low social class' districts (including partly skilled and unskilled occupations).

We estimate the main specification for these subsamples and report the results in Table 7. Panels (a)-(b) show the estimates from the high social class subsample, while Panels (c)-(d) report those for the low social class subsample. Overall, we find adverse effects of smog exposure across almost all groups, but with the reduced sample sizes, the standard errors are larger and they are not always significantly different from zero. However, the larger estimates among the low social classes suggest that the impact is disproportionally felt among those born in districts characterized by a lower socio-economic status.

For educational attainment in Column (1), we find large negative estimates among the lower social classes, and significantly so for childhood exposure, while the effects among the high social classes are estimated closer to zero. This suggests that the negative, but insignificant estimate in the main analysis is driven primarily by the lower social classes. For fluid intelligence in Column (2), we find negative estimates of pre- and postnatal exposure in all subgroups, though with variation in magnitude and statistical significance. Notably, the estimates for childhood exposure are larger in low relative to high social class districts, and they are significantly different from zero in panel (d). Similar to the findings for educational attainment, these results suggest larger cognitive impacts of pollution for individuals born in districts characterized by a larger proportion of lower social classes.

Turning to respiratory disease in Column (3), we find, consistent with the main analysis, that intrauterine exposure increases the incidence of respiratory disease. Furthermore, the estimates are almost double the size for low compared to high social classes, suggesting a stronger impact of pollution exposure on the former.

Table 7: Heterogeneity across SES groups. Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Dependent variable:		
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any
Panel (a) – High share of social class I (very high social class)			
Treated × In utero	−0.034	−0.137*	0.018
	(0.115)	(0.078)	(0.013)
Treated × Childhood	−0.035	−0.164	−0.001
	(0.105)	(0.101)	(0.014)
In utero	−0.070	−0.062	−0.002
	(0.094)	(0.068)	(0.011)
Childhood	0.080	−0.027	0.003
	(0.194)	(0.136)	(0.024)
Panel (b) – High share of social class I and II (high social class)			
Treated × In utero	−0.039	−0.099	0.020
	(0.132)	(0.097)	(0.015)
Treated × Childhood	0.019	−0.122	−0.009
	(0.130)	(0.117)	(0.016)
In utero	−0.103	−0.126	0.005
	(0.128)	(0.084)	(0.015)
Childhood	0.026	−0.150	0.044
	(0.238)	(0.163)	(0.031)
Panel (c) – High share of social class IV and V (low social class)			
Treated × In utero	−0.260	−0.066	0.047
	(0.217)	(0.137)	(0.033)
Treated × Childhood	−0.334*	−0.226	0.003
	(0.189)	(0.184)	(0.019)
In utero	0.042	0.044	−0.010
	(0.085)	(0.055)	(0.010)
Childhood	0.072	−0.006	−0.021
	(0.168)	(0.089)	(0.023)
Panel (d) – High share of social class V (very low social class)			
Treated × In utero	−0.142	−0.130**	0.038**
	(0.127)	(0.063)	(0.017)
Treated × Childhood	−0.208*	−0.280***	0.014
	(0.114)	(0.086)	(0.019)
In utero	−0.036	0.050	−0.009
	(0.059)	(0.038)	(0.007)
Childhood	−0.124	−0.001	−0.022
	(0.139)	(0.081)	(0.016)

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation. Panels: (a)-(b) subsamples with individuals born in districts with high shares of high social classes. (c)-(d) subsamples with individuals born in districts with high shares of low social classes. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

6. ROBUSTNESS ANALYSIS

We next present a range of sensitivity analyses to explore the robustness of our main findings. First, we investigate the sensitivity of our estimates to the definition of the treatment and control group. Second, we explore whether our estimates are robust to different definitions of the reference group (i.e., those conceived after the smog). Third, we examine whether the exposure effects differ for exposure in infancy versus early childhood. And fourth, we investigate the robustness to different specifications of the time trend controls. In all robustness checks, we run our analyses only on the three main outcomes above: educational attainment, fluid intelligence and respiratory disease.

6.1. Definition of treated and control districts

We start by exploring the sensitivity of our estimates to the definition of treated and control districts. For this, we examine (1) alternative definitions of treated districts, and (2) alternative definitions of control districts. First, we consider different definitions of the exposed districts. The results are reported in Table 8. In Panel (a), we start by dropping the districts classified as low exposure (defined in Section 3). Assuming these districts were less exposed compared to Central London, excluding them may increase our effect estimates. Panel (a) indeed shows that dropping low exposure districts results in estimates that are similar or slightly larger relative to those in the main analysis.

To reduce measurement error in the exposure classifications, Panel (b) exploits the actual birth locations of individuals (at a 1 km^2 resolution) and assigns exposure based on the individual's eastings and northings of birth relative to the pollution boundaries. This classification is illustrated graphically in Figure 3, showing the actual (rounded) locations of birth of UK Biobank participants within the London area. Table 8 shows that excluding individuals born outside the high exposure boundaries, but in districts that are (at least partially) exposed does not affect our estimates.

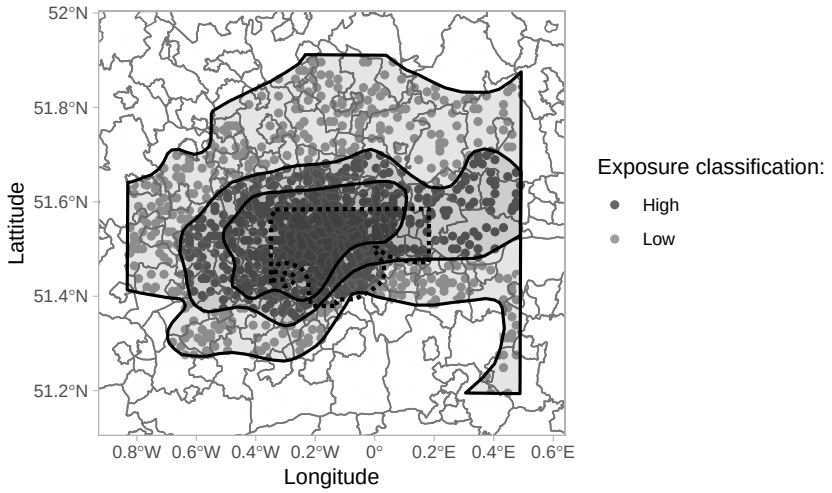
Second, our main analysis defines control districts as those with a population density of

Table 8: Definition of exposure. Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Sample:		
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any
Panel (a) – Districts, low exposure dropped			
Treated × In utero	−0.120	−0.105**	0.018
	(0.095)	(0.052)	(0.012)
Treated × Childhood	−0.147	−0.185***	−0.015
	(0.090)	(0.068)	(0.013)
In utero	−0.001	0.012	−0.005
	(0.054)	(0.036)	(0.006)
Childhood	−0.023	−0.033	−0.007
	(0.121)	(0.074)	(0.014)
Panel (b) – Birth location			
Treated × In utero	−0.099	−0.100**	0.020*
	(0.091)	(0.051)	(0.011)
Treated × Childhood	−0.130	−0.136**	−0.005
	(0.089)	(0.068)	(0.013)
In utero	−0.004	0.005	−0.005
	(0.053)	(0.035)	(0.006)
Childhood	−0.016	−0.057	−0.008
	(0.120)	(0.074)	(0.014)

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation. Panels: (a) exposed districts defined as districts overlapping with any pollution boundary but districts with low exposure have been dropped, (b) exposed individuals defined as individuals with birth location inside any pollution boundary. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

Figure 3: Visibility and pollution measurements during the London smog.

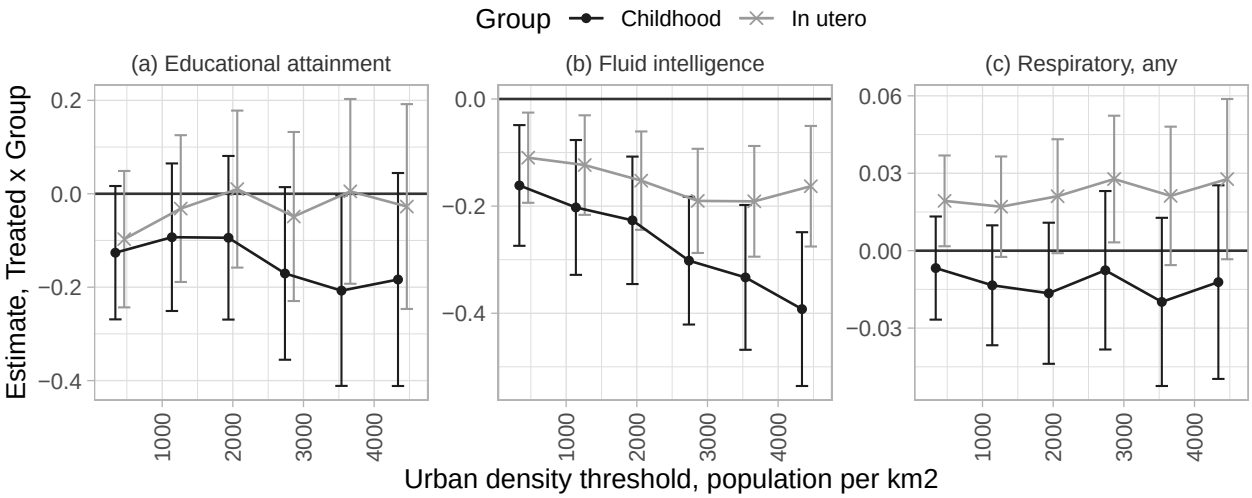


The geographic boundaries of the London smog based on the maps in Wilkins (1954). The solid black outlines mark the areas with reduction in visibilty. The inner boundaries experienced a more severe reduction in visibility. The dotted outline shows the area with high sulphur dioxide measurements. The map shows the classification of individuals' birth locations into 'high exposure' (dark gray), 'low exposure' (light gray), and 'unexposed' (white). City of London is approximately at the center of the map.

400 individuals per km^2. We here investigate the sensitivity of our results to using different population density thresholds. Figure 4 presents the estimates on the vertical axis obtained from regressions that define the control districts as those with different population densities, as indicated on the horizontal axis. The leftmost estimate corresponds to that of our main analysis. For educational attainment in Panel (a), we find insignificant negative estimates for exposure to the smog throughout. However, using more densely populated control districts, the estimates for prenatal exposure move closer to zero, while those for postnatal exposure increase in (absolute) magnitude, and the standard errors become larger due to the reduction in sample size. The estimates for fluid intelligence in Panel (b) increase in (absolute) magnitude as the control districts become more densely populated, particularly for childhood exposure. This suggests that the estimates vary somewhat depending on the definition of the control district. However, they are always negative and significantly different from zero. For Panel (c), the effect of prenatal exposure to the smog on the likelihood

of being diagnosed with respiratory disease is relatively robust to the use of more densely populated control districts.

Figure 4: Sensitivity of main estimates with respect to the urban density cutoff (population per km^2).



The vertical bars show 0.90 confidence intervals around the point estimates.

Finally, instead of using population density to define the control districts, we use the main major cities in England and Wales: Birmingham, Bristol, Cardiff, Leicester, Liverpool, Manchester, Newcastle, Nottingham and Sheffield. Indeed, one interpretation of our results is that the treated not only have a pollution shock, but also face an accumulation of pollution throughout their childhood, which may affect their later-life health. By using specific major cities *only* in the control group, we ensure that both the treated and control groups experience heightened pollution throughout their early lives. Table 9 shows that this reduces the sample size substantially and with that, increases the standard errors. Despite that, the magnitude of the estimates are very similar to those reported above. All together, this suggests that our findings are not very sensitive to the definition of treatment and control districts.

Table 9: Difference-in-Difference estimates comparing treated to control cities.

	Dependent variable:		
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any
Treated × In utero	−0.099	−0.177***	0.020*
	(0.105)	(0.050)	(0.012)
Treated × Childhood	−0.133	−0.269***	0.000
	(0.099)	(0.077)	(0.017)
In utero	−0.040	0.071	−0.007
	(0.083)	(0.053)	(0.008)
Childhood	−0.076	0.005	−0.016
	(0.212)	(0.130)	(0.021)
Observations	27,279	12,509	27,386
R^2	0.072	0.042	0.012

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation, (4)-(5) splits (2) into acute and chronic causes of respiratory hospitalisation. The 'control' cities are: Bristol, Cardiff, Leicester, Liverpool, Manchester, Newcastle, Nottingham, Sheffield, and Birmingham (defined according to 1951 districts). Includes fixed-effects for district, month of birth, and year of birth. Also controls for district-specific linear time trends. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

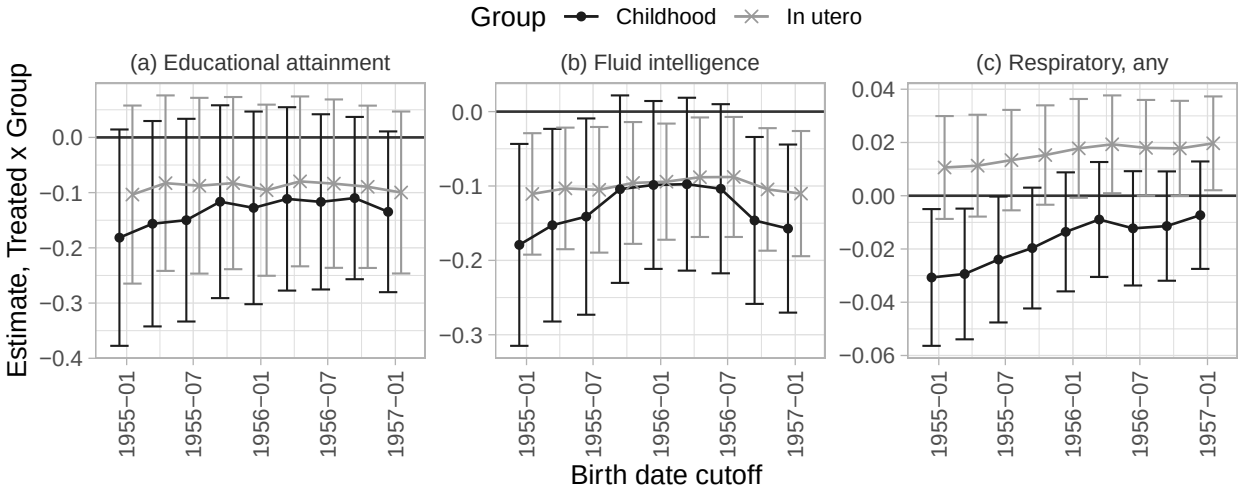
6.2. Definition of the reference group

Our main analysis compares those exposed to the smog *in utero* or in early childhood to those conceived *after* the smog. The latter (reference, or unexposed) group includes those born between September 1953 and December 1956. We next explore the sensitivity of this definition by restricting the year-months of birth to be closer to the smog, with that, reducing our sample size. Figure 5 shows the estimates of interest on the vertical axis, with the horizontal axis showing the end date used to define those who are not exposed to the smog (the right-most estimate, December 1956, is our main estimate above).

This shows that our results are relatively stable across the different definitions of the reference (unexposed) group. As we restrict the reference group to be born closer to the smog

event in panel (a) (i.e., as we move to the left on the horizontal axis), the effect of childhood exposure to the smog increases in (absolute) terms, but with the larger standard errors, its confidence intervals always overlap with zero. The estimate for prenatal exposure remains around -0.1, but it is insignificantly different from zero throughout. For fluid intelligence in Panel (b), the estimates for prenatal and childhood exposure to the smog are negative, and are almost always significantly different from zero, though with slightly more variation in the estimate for childhood exposure. Finally, in Panel (c), the estimate for the effect of childhood exposure on the probability of being diagnosed with respiratory disease increases in (absolute) magnitude and the effect of prenatal exposure reduces as we restrict the size of the reference group.

Figure 5: Sensitivity of main estimates with respect to the birth date cutoff that defines the reference group.



The vertical bars show 0.90 confidence intervals around the point estimates.

6.3. Childhood exposure

We next explore whether the childhood exposure effect differs for exposure in infancy (age 0) or later (age 1). Table 10 presents the estimates that distinguish between the two ages in early childhood, showing that the effect on fluid intelligence is driven mainly by exposure in infancy. We also find larger effects on years of education for exposure in infancy. Al-

though the effect of exposure at age 1 remains negative on education and intelligence, it is insignificantly different from zero for both outcomes.

Table 10: Childhood effects at age 0 and 1. Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Dependent variable:		
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any
Treated × In utero	−0.093	−0.094*	0.022**
	(0.090)	(0.051)	(0.011)
Treated × Childhood, age 0	−0.142	−0.172**	−0.010
	(0.088)	(0.067)	(0.012)
Treated × Childhood, age 1	−0.107	−0.089	0.005
	(0.114)	(0.082)	(0.015)
In utero	−0.012	0.005	−0.005
	(0.056)	(0.036)	(0.006)
Childhood, age 0	−0.050	−0.043	−0.005
	(0.126)	(0.079)	(0.015)
Childhood, age 1	−0.181	−0.044	0.001
	(0.181)	(0.122)	(0.022)
Observations	64,681	26,877	64,923
R^2	0.08	0.067	0.018

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

7. CONCLUSIONS

There is a substantial literature documenting the *contemporaneous* effects of exposure to pollution on individuals' human capital and health outcomes. Much less is known, however, about the potential longer-term effects of early-life pollution exposure, despite the fact that the intrauterine and early childhood environment are crucial for shaping individuals' out-

comes in older age. Indeed, most of the literature that investigates the effects of early-life pollution focuses on short-term effects, such as outcomes at birth, finding largely negative impacts. A lack of historical pollution data with good coverage of geographical locations means it is often not possible to use actual pollution measurements and relate those to later-life outcomes. To shed light on the longer-term effects, research instead has to rely on reduced form analysis and natural experiments. That is exactly what we do in this paper. Indeed, we are among the first to estimate the very long-term effects of being exposed to a severe pollution event on a range of outcomes measured at age $\sim$60.

With that, we present new evidence of the very long-term effects of an early-life pollution shock. The London smog affected Londoners between 5–9 December 1952, when a thermal inversion trapped pollution over London, which – due to weather conditions at the time – was not dispersed. We focus on the long-term human capital and health effects of exposure to the smog. We compare individuals exposed to the smog in London in either the intrauterine or infancy period to those born in other urban areas, as well as to those conceived after the smog. Our difference-in-difference analysis shows that those exposed to the smog have lower fluid intelligence scores, with some suggestive evidence that they also have fewer years of education, though this is not always sufficiently precisely estimated. We find that exposure in infancy has slightly larger effects compared to exposure *in utero*. Investigating the long-term *health* effects, we find a large increase in the probability of being diagnosed with respiratory disease due to intrauterine exposure, which is driven by acute respiratory conditions.

We next study potential differential effects of *in utero* exposure, distinguishing between exposure in the first, second, and third trimester. We find some evidence for differential gestational effects for fluid intelligence, with larger effects for exposure at early gestational ages. Similarly, we find that the increase in respiratory conditions is driven by first trimester exposure.

We then model the heterogeneity of our effect estimates with respect to three sources of (predetermined) variation: individuals' genetic predisposition, gender and socio-economic

status at birth. We estimate the extent to which individuals' genetic variation can moderate the effects of exogenous early-life pollution exposure. Indeed, individuals with a high 'genetic predisposition' for education may be able to overcome such adverse early-life environments. Our findings, however, show that the negative effects of smog exposure on educational attainment are driven by individuals with a high polygenic score for education, and that the increase in respiratory disease due to prenatal smog exposure is larger for individuals with a high polygenic score, suggesting that a higher polygenic score increases individuals' vulnerability with respect to respiratory conditions following a severe pollution event. This highlights the *joint* role that 'nature' and 'nurture' play in shaping individuals' outcomes, and presents clear evidence *against* genetic (or environmental) determinism; the belief that one's outcomes are solely affected by genetic variation (or environmental characteristics). Similarly, we find that the negative effect on years of schooling is driven by women, and that the worsening of human capital as well as respiratory health is driven by those in low socio-economic status environments.

Our estimates are quantitatively and qualitatively important, but there are three key points regarding their interpretation. First, they estimate the effects of exposure to a *severe* pollution event. Indeed, London nowadays experiences pollution levels that are still high, but nowhere near those observed in 1952. Hence, our results cannot be extrapolated to smog events occurring in London or most other cities in Europe nowadays. Despite that, they *do* compare to smog events that are happening each year in industrialising economies such as India and China. Hence, our findings are relevant to those settings, suggesting that such extreme pollution events do not only affect contemporaneous outcomes, but also have longer term adverse effects.

Second, our analysis compares pollution during the smog to 'standard' pollution levels in control districts as well as in the years immediately after the smog. Although these levels of pollution are indeed lower than those in inner London, they are not comparable to current levels of pollution. Hence, our estimates capture the effect of being exposed to

a large pollution shock, relative to already high levels throughout early childhood. Again, the results can therefore better be extrapolated to industrialising economies with higher pollution levels in general, as well as larger pollution shocks.

Third, our estimates are likely to be a lower bound of the 'true' effect of the smog. This is the case for three reasons. First, the evidence suggests that the smog led to an increase in infant mortality and foetal loss. Assuming that those who died were more vulnerable and those who survived were stronger, this suggests that our estimates are likely to be a lower bound.[22] Second, and relatedly, since individuals in the UK Biobank were invited to participate in 2006-2010, we implicitly condition on survival until this time. In the presence of frailty selection, where fragile individuals are more likely to die prior to assessment leaving stronger survivors in the sample, our estimates are likely to be attenuated. The third reason why our estimates are likely to be downward biased is due to measurement error. For one, since we do not observe gestational age, we assume all individuals were *in utero* for nine months prior to their year-month of birth, and we assume individuals were born on the first of the month. In reality, some individuals would have had a shorter gestational period, potentially misclassifying them as being exposed to the smog. Two, we are reliant on publications from the 1950s, showing the extent of the smog as well as its variability across London. We therefore define individuals as either exposed or unexposed, but in reality, pollution would have shown more regional variability that we are unable to capture in our analyses. Three, related to this, we observe individuals' location of birth (eastings and northings) with a $1km^2$ resolution. Given that pollution changes across space, individuals who are born on the boundary of our exposed and unexposed districts may be misclassified, leading to additional measurement error.

As with any research, our analysis comes with its limitations. First, we cannot identify

[22]To explore potential selective mortality with respect to gender, using the knowledge that male foetuses are generally frailer than female foetuses, we estimate the effect of the smog on the probability of being male. However, this analysis does not show significant effects on the gender ratio, suggesting there is no strong evidence of gender differences in survival. We do not report the results here, but they are available from the authors upon request.

which pollutants matter more for individuals' human capital and health outcomes. Indeed, the literature suggests that multiple pollutants increased during the smog, and we cannot determine whether one or more of these are driving the deterioration in later-life outcomes. Second, the UK Biobank is a large database of those aged 45–69 in the United Kingdom. However, it is not representative of the population, since recruitment into the study was voluntary. Indeed, UK Biobank participants are generally healthier and wealthier than the general UK population (Fry et al., 2017). Having said that, the data is unique in combining information from questionnaires, objective measurements, genetic information, and administrative data on a very large sample of UK residents, with information on year-month and location of birth, allowing us to identify which participants were exposed and which were not. With that, our paper highlights the very long-run pollution effects, identifying reductions in individuals' human capital and health outcomes up to ∼60 years after the actual exposure.

Finally, our findings have clear policy implications. They suggest that reducing pollution has large, long-term benefits for the population. From the population perspective, given the ease and low cost of pollution forecasting, the benefits of avoiding pollution are substantial. From a policy maker perspective, this should encourage the implementation of incentives or regulation that reduce pollution from e.g., residential homes, firms, or transportation. Indeed, we show that creating environments that improve individuals' human capital and health outcomes in the long-term starts with improving environments before they are born.

REFERENCES

Almond, D (2006). "Is the 1918 Influenza Pandemic Over? Long-Term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population". In: *Journal of Political Economy* 114 (4), pp. 672–712.

Almond, D and J Currie (2011a). "Human capital development before age 5". In: *Handbook of Labor Economics*. Elsevier, pp. 1315–1486.

— (2011b). "Killing Me Softly: The Fetal Origins Hypothesis". In: *The Journal of Economic Perspectives* 25 (3), pp. 153–172.

Almond, Douglas, Janet Currie, and Valentina Duque (2018). "Childhood circumstances and adult outcomes: Act II". In: *Journal of Economic Literature* 56.4, pp. 1360–1446.

Almond, Douglas, Lena Edlund, and Mårten Palme (2009). "Chernobyl's subclinical legacy: prenatal exposure to radioactive fallout and school outcomes in Sweden". In: *Quarterly journal of economics* 124.4, pp. 1729–1772.

Almond, Douglas and Bhashkar Mazumder (2005). "The 1918 influenza pandemic and subsequent health outcomes: an analysis of SIPP data". In: *American Economic Review* 95.2, pp. 258–262.

— (2011). "Health capital and the prenatal environment: the effect of Ramadan observance during pregnancy". In: *American Economic Journal: Applied Economics* 3.4, pp. 56–85.

Altshuler, David M. et al. (Sept. 2010). "Integrating common and rare genetic variation in diverse human populations". In: *Nature* 467.7311, pp. 52–58.

Arceo, Eva, Rema Hanna, and Paulina Oliva (2016). "Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City". In: *The Economic Journal* 126.591, pp. 257–280.

Aveyard, Paul et al. (2021). "Association between pre-existing respiratory disease and its treatment, and severe COVID-19: a population cohort study". In: *The lancet Respiratory medicine*.

Ball, A (2018a). "The long-term economic costs of the Great London Smog". In: *Birkbeck Working Paper 1814*.

Ball, Alastair (2018b). "Hidden costs of the Great London Smog: evidence from missing births". In.

Banerjee, Abhijit et al. (2010). "Long-run health impacts of income shocks: Wine and phylloxera in nineteenth-century France". In: *The Review of Economics and Statistics* 92.4, pp. 714–728.

Barcellos, Silvia H, Leandro S Carvalho, and Patrick Turley (Oct. 2018). "Education can reduce health differences related to genetic risk of obesity". In: *Proceedings of the National Academy of Sciences* 115.42, E9765–E9772. DOI: 10.1073/pnas.1802909115.

Barth, Daniel, Nicholas W Papageorge, and Kevin Thom (2020). "Genetic endowments and wealth inequality". In: *Journal of Political Economy* 128.4, pp. 1474–1522.

BBC (1952). *London fog clears after days of chaos*.

Becker, Joel et al. (2021). "Resource profile and user guide of the Polygenic Index Repository". In: *Nature human behaviour*, pp. 1–15.

Behrman, Jere R and Mark R Rosenzweig (2004). "Returns to birthweight". In: *Review of Economics and statistics* 86.2, pp. 586–601.

Bell, M L and D L Davis (2001). "Reassessment of the lethal London fog of 1952: Novel indicators of acute and chronic consequences of acute exposure to air pollution". In: *Environmental Health Perspectives* 109 (3), pp. 389–394.

Bell, M L, D L Davis, and T Fletcher (2004). "A retrospective assessment of mortality from the London smog episode of 1952: The role of influenza and pollution". In: *Environmental Health Perspectives* 112 (1), pp. 6–8.

Belsky, Daniel W et al. (2018). "Genetic analysis of social-class mobility in five longitudinal studies". In: *Proceedings of the National Academy of Sciences* 115.31, E7275–E7284.

Bharadwaj, Prashant et al. (2016). "Early-life exposure to the great smog of 1952 and the development of asthma". In: *American journal of respiratory and critical care medicine* 194.12, pp. 1475–1482.

Bharadwaj, Prashant et al. (2017). "Gray matters: Fetal pollution exposure and human capital formation". In: *Journal of the Association of Environmental and Resource Economists* 4.2, pp. 505–542.

Bierut, Laura et al. (2018). "Childhood socioeconomic status moderates genetic predisposition for peak smoking". In: *BioRxiv*, p. 336834.

Bijwaard, Govert E et al. (2021). "Severe Prenatal Shocks and Adolescent Health: Evidence from the Dutch Hunger Winter". In.

Biroli, Pietro (2015). "Genetic and economic interaction in health formation: The case of obesity". In: *Working Paper*.

Biroli, Pietro and Christian Zünd (2021). "Genes, Pubs, and Drinks: Gene-environment interplay and alcohol licensing policy in the UK". In: *Mimeo, University of Zurich*.

Biroli, P et al. (2021). *Using genetic data in economics: The interplay between 'nature' and 'nurture'*. Tech. rep. Mimeo.

Bishop, Kelly C, Jonathan D Ketcham, and Nicolai V Kuminoff (2018). *Hazed and confused: the effect of air pollution on dementia*. Tech. rep. National Bureau of Economic Research.

Black, Sandra E et al. (2013). *This is only a test? Long-run impacts of prenatal exposure to radioactive fallout*. Tech. rep. National Bureau of Economic Research.

Bleakley, Hoyt (2007). "Disease and development: evidence from hookworm eradication in the American South". In: *Quarterly Journal of Economics* 122.1, pp. 73–117.

Bycroft, Clare et al. (2018). "The UK Biobank resource with deep phenotyping and genomic data". In: *Nature* 562.7726, pp. 203–209.

Carneiro, Pedro, Katrine V Løken, and Kjell G Salvanes (2015). "A flying start? Maternity leave benefits and long-run outcomes of children". In: *Journal of Political Economy* 123.2, pp. 365–412.

Case, Anne and Christina Paxson (2009). "Early life health and cognitive function in old age". In: *American Economic Review* 99.2, pp. 104–09.

Cattan, Sarah et al. (2021). *The health effects of universal early childhood interventions: Evidence from Sure Start*. Tech. rep. IFS Working Paper.

Chay, Kenneth Y and Michael Greenstone (2003). "The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession". In: *Quarterly Journal of Economics* 118.3, pp. 1121–1167.

Clark, D and H Royer (2013). "The effect of education on adult mortality and health: Evidence from Britain". In: *American Economic Review* 103 (6), pp. 2087–2120.

Conti, Gabriella, Giacomo Mason, and Stavros Poupakis (2019). *Developmental origins of health inequality*. Tech. rep. IZA Discussion Papers.

Cunha, Flavio and James Heckman (2007). "The technology of skill formation". In: *American Economic Review* 97.2, pp. 31–47.

Currie, J (2009). "Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development". In: *Journal of Economic Literature* 47 (1), pp. 87–122.

Currie, Janet and Matthew Neidell (2005). "Air pollution and infant health: what can we learn from California's recent experience?" In: *Quarterly Journal of Economics* 120.3, pp. 1003–1030.

Currie, Janet and Reed Walker (2011). "Traffic congestion and infant health: Evidence from E-ZPass". In: *American Economic Journal: Applied Economics* 3.1, pp. 65–90.

Davies, N et al. (2018). "The causal effects of education in the UK Biobank". In: *Nature Human Behaviour*, pp. 117–125.

Elsworth, BL et al. (2019). *MRC IEU UK Biobank GWAS pipeline version 2*. MRC IEU, University of Bristol. URL: https://doi.org/10.5523/bris.pnoat8cxo0u52p6ynfaekeigi.

EPA (2021). *Sulfur Dioxide Basics*.

Fletcher, Jason M (Jan. 2012). "Why have tobacco control policies stalled? Using genetic moderation to examine policy impacts." In: *PloS One* 7.12, e50576. DOI: 10.1371/journal.pone.0050576.

— (2018). "Environmental Bottlenecks on Children's Genetic Potential for Adult Socioeconomic Attainments: Evidence from a Health Shock". In: *IZA discussion paper*.

Fry, Anna et al. (2017). "Comparison of sociodemographic and health-related characteristics of UK Biobank participants with those of the general population". In: *American journal of epidemiology* 186.9, pp. 1026–1034.

Fukushima, Nanna (2021). "The UK Clean Air Act, Black Smoke, and Infant Mortality". PhD thesis. PhD thesis.

Graff Zivin, Joshua and Matthew Neidell (2013). "Environment, health, and human capital". In: *Journal of Economic Literature* 51.3, pp. 689–730.

Greenstone, Michael and B Kelsey Jack (2015). "Envirodevonomics: A research agenda for an emerging field". In: *Journal of Economic Literature* 53.1, pp. 5–42.

Hanlon, W Walker (2018). *London fog: A century of pollution and mortality, 1866-1965*. Tech. rep. National Bureau of Economic Research.

Harmon, C and I Walker (1995). "Estimates of the economic return to schooling for the United Kingdom". In: *American Economic Review* 85 (5), pp. 1278–1286.

Isen, Adam, Maya Rossin-Slater, and W Reed Walker (2017). "Every breath you take—every dollar you'll make: The long-term consequences of the clean air act of 1970". In: *Journal of Political Economy* 125.3, pp. 848–902.

Jans, Jenny, Per Johansson, and J Peter Nilsson (2018). "Economic status, air quality, and child health: Evidence from inversion episodes". In: *Journal of health economics* 61, pp. 220–232.

Jayachandran, Seema (2009). "Air quality and early-life mortality evidence from Indonesia's wildfires". In: *Journal of Human resources* 44.4, pp. 916–954.

Jia, Ruixue and Hyejin Ku (2019). "Is China's pollution the culprit for the choking of South Korea? Evidence from the Asian dust". In: *The Economic Journal* 129.624, pp. 3154–3188.

Knittel, Christopher R, Douglas L Miller, and Nicholas J Sanders (2016). "Caution, drivers! Children present: Traffic, pollution, and infant health". In: *Review of Economics and Statistics* 98.2, pp. 350–366.

Kong, Augustine et al. (2018). "The nature of nurture: Effects of parental genotypes". In: *Science* 359.6374, pp. 424–428.

Lancet, The (2019). *Air pollution: a major threat to lung health.*

Laskin, D. (2006). "The great London smog". In: *Weatherwise* 59 (6), pp. 42–45.

Lee, J. et al (2018). "Gene discovery and polygenic prediction from a genome-wide association study of eduational attainment in 1.1 million individuals". In: *Nat Gen* 50, pp. 1112–1121.

Logan, William PD et al. (1953). "Mortality in the London fog incident, 1952." In: *Lancet*, pp. 336–8.

Loh, Po-Ru et al. (Mar. 2015). "Efficient Bayesian mixed-model analysis increases association power in large cohorts". In: *Nature Genetics* 47.3, pp. 284–290.

Lumey, Lambert H, Aryeh D Stein, and Ezra Susser (2011). "Prenatal famine and adult health". In: *Annual review of public health* 32, pp. 237–262.

MET Office (2022). *Datasets.*

Ministry of Health (1954). *Mortality and morbidity during the London fog of December 1952. Reports on public health and medical subjects.* 95.

Muslimova, D et al. (2020). "Dynamic complementarity in skill production: Evidence from genetic endowments and birth order". In.

Nilsson, J Peter (2017). "Alcohol availability, prenatal conditions, and long-term economic outcomes". In: *Journal of Political Economy* 125.4, pp. 1149–1207.

Novembre, John and Matthew Stephens (2008). "Interpreting principal component analyses of spatial population genetic variation". In: *Nature genetics* 40.5, pp. 646–649.

Okbay, Aysu et al. (May 2016). "Genome-wide association study identifies 74 loci associated with educational attainment". In: *Nature* 533.7604, pp. 539–542.

Pereira, Rita Dias, Cornelius A Rietveld, and Hans van Kippersluis (2020). "The interplay between maternal smoking and genes in offspring birth weight". In: *MedRxiv.*

Pereira, Rita Dias et al. (2021). "Gene-by-Environment Interplay". In: *mimeo.*

Polderman, Tinca JC et al. (2015). "Meta-analysis of the heritability of human traits based on fifty years of twin studies". In: *Nature genetics* 47.7, pp. 702–709.

Price, Alkes L et al. (2006). "Principal components analysis corrects for stratification in genome-wide association studies". In: *Nature genetics* 38.8, pp. 904–909.

Privé, Florian, Julyan Arbel, and Bjarni J Vilhjálmsson (2020). "LDpred2: better, faster, stronger". In: *Bioinformatics* 36.22-23, pp. 5424–5431.

Purcell, Shaun M. et al. (2009). "Common polygenic variation contributes to risk of schizophrenia and bipolar disorder". In: *Nature* 460.7256, pp. 748–752. DOI: 10.1038/nature08185.

Rangel, Marcos A and Tom S Vogl (2019). "Agricultural fires and health at birth". In: *Review of Economics and Statistics* 101.4, pp. 616–630.

Reyes, Jessica Wolpaw (2007). "Environmental policy as social policy? The impact of childhood lead exposure on crime". In: *The BE Journal of Economic Analysis & Policy* 7.1.

Rietveld, Cornelius A et al. (Nov. 2013). "GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment". In: *Science* 340.6139, pp. 1467–1471. DOI: 10.1257/jep.25.4.57.

Ronda, Victor (2020). "Family disadvantage, gender and the returns to genetic human capital". In.

Rutter, Michael (2006). *Genes and Behavior: Nature-Nurture Interplay Explained*. Oxford Blackwell Publishing.

Sanders, Nicholas J (2012). "What doesn't kill you makes you weaker prenatal pollution exposure and educational outcomes". In: *Journal of Human Resources* 47.3, pp. 826–850.

Sanders, Nicholas J and Charles Stoecker (2015). "Where have all the young men gone? Using sex ratios to measure fetal death rates". In: *Journal of health economics* 41, pp. 30–45.

Schmitz, Lauren L and Dalton C Conley (June 2016a). "The Impact of Late-Career Job Loss and Genotype on Body Mass Index". In: *NBER Working Paper* 22348. DOI: 10.3386/w22348.

— (Jan. 2016b). "The Long-Term Consequences of Vietnam-Era Conscription and Genotype on Smoking Behavior and Health". In: *Behavior Genetics* 46.1, pp. 43–58. DOI: 10.1007/s10519-015-9739-1.

Southall, Humphrey (2011). "Rebuilding the Great Britain Historical GIS, Part 1: Building an indefinitely scalable statistical database". In: *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 44.3, pp. 149–159.

Southall, Humphrey and Paula Aucott (2009). *A vision of Britain through time*. English. Department of Geography, University of Portsmouth. URL: http://www.visionofbritain.org.uk.

Turkheimer, Eric (2000). "Three laws of behavior genetics and what they mean". In: *Current directions in psychological science* 9.5, pp. 160–164.

van den Berg, G, S von Hinke, and A Wang (2021). "Prenatal Sugar Consumption and Late-Life Health: Analyses Based on Post-Wartime Rationing and Polygenic Scores". In: *mimeo, University of Bristol*.

Van den Berg, Gerard J, Maarten Lindeboom, and France Portrait (2006). "Economic conditions early in life and individual mortality". In: *American Economic Review* 96.1, pp. 290–302.

Vilhjálmsson, Bjarni J et al. (2015). "Modeling linkage disequilibrium increases accuracy of polygenic risk scores". In: *The american journal of human genetics* 97.4, pp. 576–592.

von Hinke, Stephanie, Nigel Rice, and Emma Tominey (2019). *Mental Health around Pregnancy and Child Development from Early Childhood to Adolescence*. Tech. rep.

von Hinke, Stephanie et al. (2014). "Alcohol exposure in utero and child academic achievement". In: *The Economic Journal* 124.576, pp. 634–667.

Warren Spring Laboratory (1967). *The investigation of Atmospheric Pollution 1958–1966: Thirty-second report*. Her Majesty's Stationery Office, London.

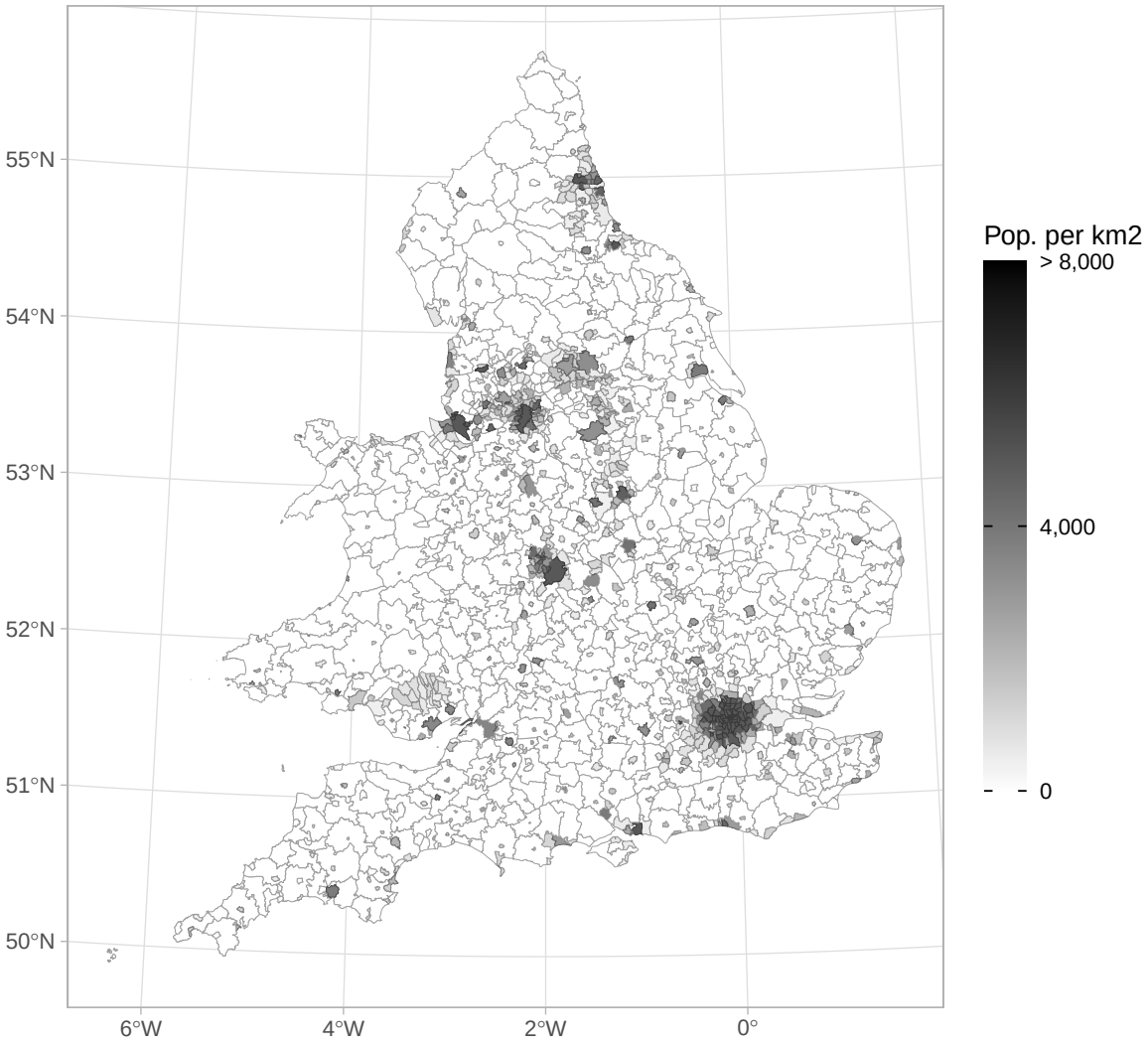
Wilkins, ET (1954). "Air pollution and the London fog of December, 1952". In: *Journal of the Royal Sanitary Institute* 74.1, pp. 1–21.

Zhang, Xin, Xi Chen, and Xiaobo Zhang (2018). "The impact of exposure to air pollution on cognitive performance". In: *Proceedings of the National Academy of Sciences* 115.37, pp. 9193–9197.

ONLINE APPENDIX

A. Additional Tables and Figures

Figure A.1: Population density (in population per km^2) at the district level in England and Wales.



The map uses data on population density and districts from Vision of Britain (Southall, 2011)

Table A.1: Mapping between qualifications and years of education.

Qualifications	Years of education
College or university degree	16
A/AS levels + NVQ/HND/HNC	14
A/AS levels + Other professional qualifications	15
NVQ/HND/HNC	13
Other professional qualifications	12
A/AS levels	13
CSEs, GCSEs, or O levels	11
No qualifications	10

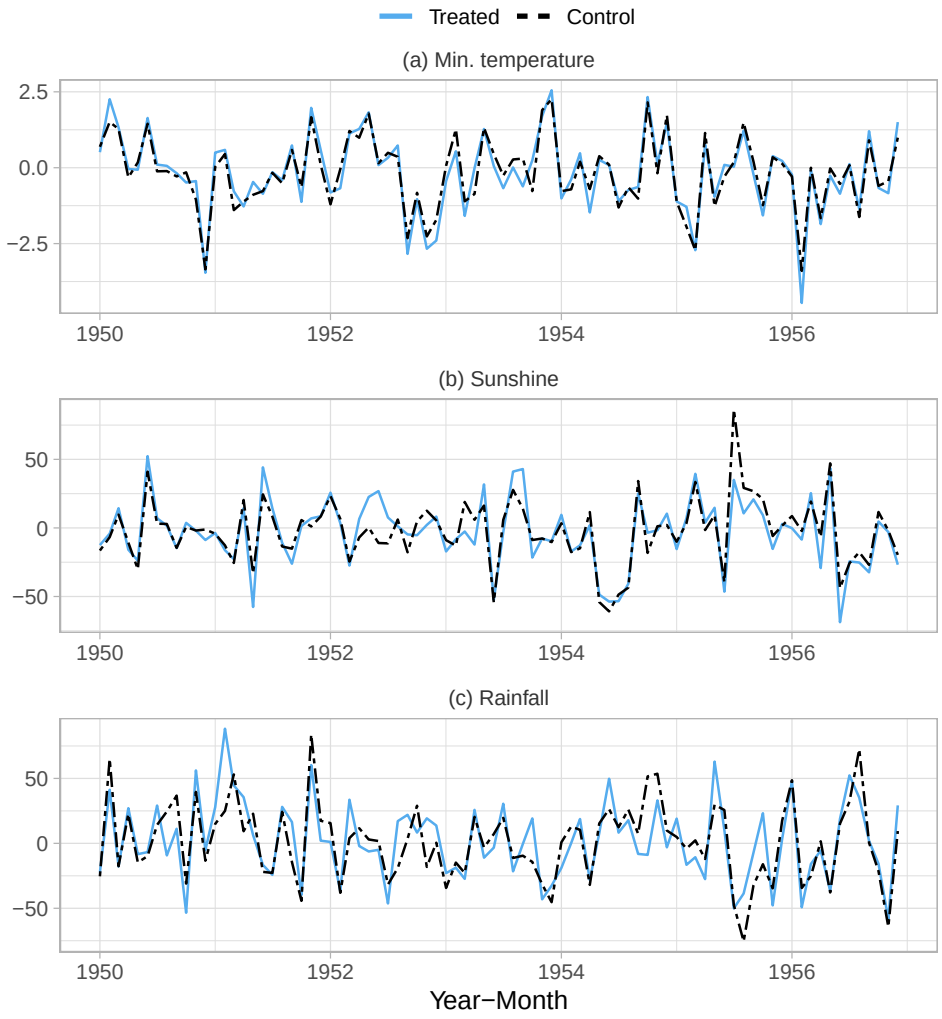
Columns: (1) the qualifications recorded in the UK Biobank, (2) the assigned years of education. A plus indicates that the individual must hold both of the specified qualifications simultaneously.

Table A.2: Heterogeneity across genetics – Qualifications. Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Exits education system with qualification:		
	(1) Upper secondary	(2) Lower secondary	(3) None
Panel (a) – High polygenic score			
Treated × In utero	-0.052^{***}	0.047^{***}	0.005
	(0.019)	(0.017)	(0.009)
Treated × Childhood	-0.043^{**}	0.023	0.020^{**}
	(0.020)	(0.019)	(0.010)
In utero	0.013	−0.015	0.002
	(0.015)	(0.012)	(0.008)
Childhood	−0.008	0.011	−0.002
	(0.032)	(0.024)	(0.019)
Panel (b) – Low polygenic score			
Treated × In utero	0.016	−0.028	0.012
	(0.031)	(0.022)	(0.021)
Treated × Childhood	−0.003	0.040	−0.036
	(0.033)	(0.027)	(0.022)
In utero	0.001	0.015	−0.016
	(0.015)	(0.014)	(0.014)
Childhood	0.046	−0.023	−0.023
	(0.031)	(0.026)	(0.028)

Columns: (1) exits at upper secondary level (university/college degree, A/AS-levels, professional/vocational training), (2) exits at lower secondary level (CSEs, GCSEs, O-levels), (3) exits with no qualifications. Panels: (a) high PGS subsample, (b) low PGS subsample. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

Figure A.2: Time series for minimum temperature, sunshine, and rainfall, in control and treated districts.



We take the measurements at the birth locations of all individuals in our sample and average these by year-month and treatment status. Before averaging we remove seasonality using a set of month dummies.

Table A.3: Difference-in-Difference estimates of a male birth, comparing treated to control districts defined as urban England and Wales.

	Dependent variable:
	(1)
	Male
Treated × In utero	−0.007
	(0.021)
Treated × Childhood	0.011
	(0.020)
In utero	0.001
	(0.011)
Childhood	−0.044**
	(0.022)
Observations	65,060
R^2	0.014

Columns: (1) being born as male. Includes fixed-effects for district, month of birth, and year of birth. Also controls for year-month linear time trends by administrative county. Urban England and Wales are defined as districts that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

B. Choice of time trends

The main analysis controls for administrative-county-specific (year-month) trends to allow the outcome of interest to trend differently in each administrative county. We here explore the sensitivity of the trend-specifications. Panel (a) of Table B.1, includes administrative county-specific *annual* (as opposed to year-month) trends. Panel (b) specifies *year-month* trends for each of the over 1400 districts observed in our data, and Panel (c) allows for district-specific *annual* trends. Finally, Panel (d) does not include any trends and only accounts for year of birth dummies.

Table B.1: Specification of trends. Difference-in-Difference estimates comparing treated to control districts defined as urban England and Wales.

	Dependent variable:		
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any
Panel (a) – Year by administrative county			
Treated × In utero	−0.087	−0.101**	0.016
	(0.089)	(0.050)	(0.010)
Treated × Childhood	−0.119	−0.142**	−0.014
	(0.086)	(0.066)	(0.012)
Panel (b) – Year-month by district			
Treated × In utero	−0.120	−0.162***	0.022*
	(0.091)	(0.048)	(0.012)
Treated × Childhood	−0.174*	−0.245***	−0.003
	(0.097)	(0.064)	(0.016)
Panel (c) – Year by district			
Treated × In utero	−0.112	−0.148***	0.017
	(0.090)	(0.047)	(0.012)
Treated × Childhood	−0.167*	−0.228***	−0.013
	(0.093)	(0.062)	(0.015)
Panel (d) – No trend			
Treated × In utero	−0.050	−0.072*	0.024***
	(0.075)	(0.043)	(0.008)
Treated × Childhood	−0.047	−0.086**	−0.001
	(0.047)	(0.034)	(0.008)

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation. Panels: (a) Year trend at administrative county ($n = 174$) level, (b) Year-month trend at district ($n = 785$) level, (c) Year trend at district level. (d) No trends. We always include district FE, year-of-birth FE, and month-of-birth FE. Standard errors are clustered by district.

The estimates are generally consistent across the different specifications. For educational attainment, we find negative estimates throughout for both in utero and childhood exposure. For fluid intelligence, we find clear evidence of a negative effect of smog exposure that is slightly larger for exposure in childhood compared to prenatally. Finally, for respiratory disease, the estimates are always positive, but they are slightly smaller when accounting for annual trends at either the administrative county or district level. Nevertheless, the magnitude of the estimates remains in the same ballpark, suggesting that smog exposure increases the probability of being diagnosed with respiratory disease.

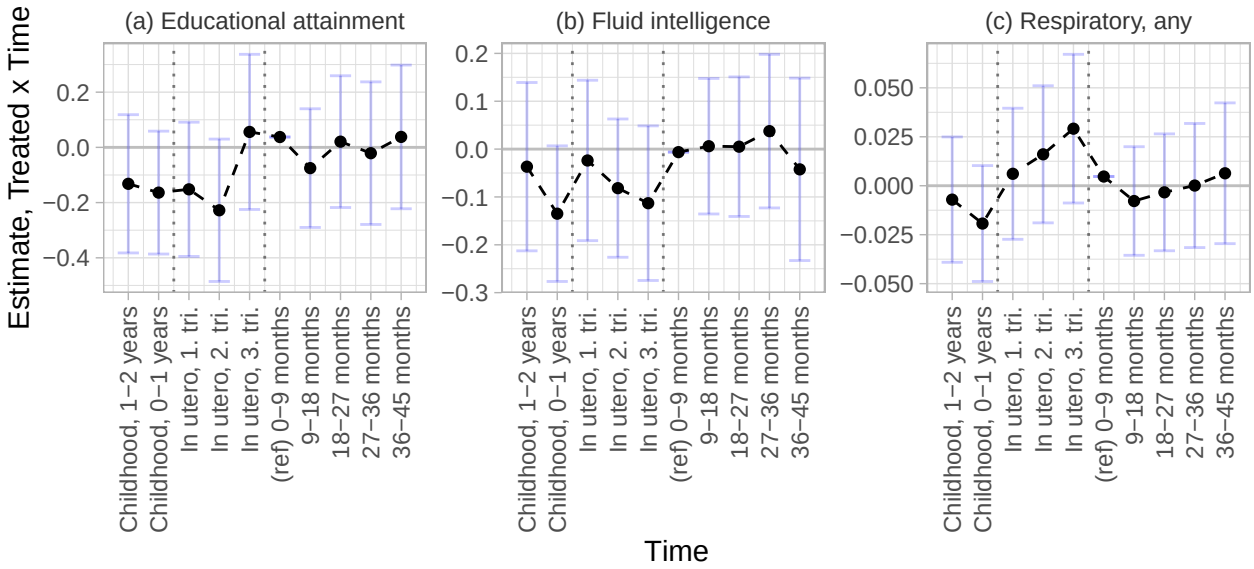
C. Common time trends

Our specification implicitly assumes that individuals who were born in districts that were affected by the London smog would have had similar trends in their outcomes of interest in the absence of the smog compared to those born in districts that were not affected. To explore this common trend assumption empirically, we compare the trends in the relevant outcomes of interest among those conceived at different points in time throughout our observation period in treated and control districts. Note here, that we are mainly interested in comparing individuals in treated and control districts who are conceived *after* the smog, since those who are conceived before or during the smog were potentially exposed either *in utero* or in childhood. We here focus on our main outcomes of interest: education, fluid intelligence and respiratory disease.

Figure C.1 shows the conditional difference in the mean of the relevant outcome for those born in treated versus control districts across childhood, trimesters in utero, and 9 month intervals throughout our observation period. We condition on the same controls and fixed-effects as in the main analysis. The two vertical dotted lines indicate the threshold for potential exposure in childhood and *in utero*.

Figure C.1a and Figure C.1b show that those exposed in utero or in early childhood have lower education and fluid intelligence compared to those born at the same time, but in control districts. Similarly, Figure C.1c shows that those who are exposed to the smog whilst *in utero* have a higher probability of being diagnosed with respiratory disease compared to those born at the same time, but in control districts. For all outcomes, we see no suggestion of differential trends for those conceived after the smog, i.e., those to the right of the second vertical dotted line. In other words, we find no evidence to suggest that those born in treated districts have differential trends in our outcomes of interest compared to those born control districts.

Figure C.1: The conditional differences in the means of the relevant outcome for those born in treated versus control districts.



Shows the conditional differences across childhood, trimesters in utero, and 9 month intervals throughout our observation period. We condition on the same controls as in the main analysis, and we include fixed-effects for district. We also control for year-month linear time trends by administrative county. The two vertical dotted lines indicate the threshold for potential exposure in childhood and *in utero*. The vertical bars show 90% confidence intervals around the point estimates. Control districts are defined as districts in England and Wales that had a population density above 400 individuals per km^2 in 1951. Standard errors are clustered by district.

D. A brief background to genetics

Humans have 46 chromosomes stored in every cell apart from sex-cells. The chromosomes exist in pairs such that each pair has a maternal and paternal copy. A single chromosome consists of a double-strand of deoxyribonucleic acid (DNA) containing a large number of 'base pairs': pairs of nucleotide molecules (referred to as the 'letters' A (adenine) that binds with T (thymine), and G (guanine) that binds with C (cytosine)) that together make up the human genome. In a population there will be variation in the base pairs at some locations. Such variation is known as a single nucleotide polymorphism (SNP, pronounced 'snip') – a change in the base pair at one particular locus (location) – and is the most commonly studied genetic variation. When there are two possible base pairs at a given location (i.e., two alleles), the most frequent base pair is called the major allele, while the less frequent is called the minor allele. As humans have two copies of each chromosome, any given individual can have either zero, one, or two copies of the minor allele.

To identify specific SNPs that are robustly associated with a particular outcome of interest, so-called Genome-Wide Association Studies (GWAS) relate each SNP to the outcome in a hypothesis-free approach. As there are more SNPs than individuals, the SNP effects cannot be identified in a multivariate regression model. Instead, a GWAS runs a large number of univariate regressions of the outcome on each SNP. These analyses have shown that most outcomes of interest in the social sciences are 'polygenic': they are affected by a large number of SNPs, each with a very small effect. To increase the predictive power of the SNPs, it is therefore custom to aggregate the individual SNPs into so-called polygenic scores, as:

$$G_i = \sum_{j=1}^{J} \beta_j X_{ij},$$

where X_{ij} is a count of the number of minor alleles (i.e., 0, 1 or 2) at SNP j for individual i, and β_j is its effect size obtained from an independent GWAS. Hence, the polygenic scores are weighted linear combinations of SNPs, where the weights are estimated in an independent

GWAS. This is motivated by an additive genetic model where all SNPs contribute additively to the overall genetic predisposition of an individual (see e.g., Purcell et al., 2009).

We conduct our own tailor-made GWAS for each of the main outcomes in our analysis: educational attainment, fluid intelligence score, respiratory disease (acute, chronic, and combined), and COVID-19. To avoid overfitting, we partition the UK Biobank into three non-overlapping samples: (1) a GWAS discovery sample, (2) a reference/tuning sample, and (3) the analysis sample. We use samples (1) and (2) to construct polygenic scores for the individuals in the analysis sample (i.e., sample (3)). We then use these polygenic scores to explore the extent to which one's genetic predisposition can 'protect' or 'exacerbate' the effect of early-life pollution exposure. The analysis sample is outlined in Section 3 and contains individuals born in treated or control districts in 1950–1956. The GWAS discovery sample contains individuals born outside the study period, i.e., 1934-1949 and 1957-1970, as well as individuals born outside treated and control districts in the years 1950-1956. From the GWAS discovery sample, we randomly sample 20,000 unrelated individuals of white British ancestry that we exclusively use for the reference sample.[23]

The GWAS discovery sample sizes and descriptive statistics are reported in Table D.1. The sample size varies depending on the outcome of interest. For our GWAS, we follow the quality control (QC) procedure described by Elsworth et al. (2019) to remove genetic outliers and ensure the genotypes are well-measured. We follow the literature and include minimal covariates in the GWAS, controlling for gender, genotyping array, birth year, and the first 20 genetic principal components.[24] To maximise the size of the discovery sample, we use BOLT-LMM (Loh et al., 2015) to run the GWAS. Since BOLT-LMM uses a linear mixed model, it allow us to include related individuals and to relax the restrictions on ancestry (i.e., European ancestry instead of white British individuals only).

Using the GWAS estimates, we use LDPred2 (Privé et al., 2020) to construct polygenic

[23]We use the reference sample to estimate genetic correlations (LD structure) and to select tuning parameters for the LDpred2 method.

[24]Principal components are commonly used to control for population stratification, see Price et al. (2006) and Novembre and Stephens (2008).

scores. We construct the polygenic scores under both an infinitesimal model (all SNPs are causal) and a model where the proportion of causal SNPs is estimated as an additional parameter using a grid search (for more information, see Privé et al., 2020). To avoid overfitting, this grid search is done in the reference sample. We reduce the computational burden by including only SNPs that are in HapMap3 (Altshuler et al., 2010), resulting in approximately 1.6 million SNPs.[25] We also filter the SNPs using a minor allele frequency threshold of 0.01 and an info score threshold of 0.97.

All polygenic scores are standardised to have zero mean and unit standard deviation in the analysis sample. To validate their predictive power, we use a linear regression model to test them against their target outcome in the analysis sample. We control for sex and the first 20 genetic principal components, and we include fixed effects for year-month of birth. We report the incremental R^2 defined as the increase in R^2 when the polygenic score is included as a covariate. Table D.2 reports the results, showing that each polygenic score is highly predictive of its outcome, with the incremental R^2 ranging between 0.1% and 10%.

[25] See also https://www.sanger.ac.uk/resources/downloads/human/hapmap3.html (accessed 18 October 2021).

Table D.1: Descriptive statistics – GWAS sample.

	Obs. (1)	Mean (2)	Std. dev. (3)
Educational attainment	378,503	14.736	5.175
Fluid intelligence	138,933	0.065	0.981
Respiratory disease	377,577	0.111	0.314
– Acute	377,573	0.089	0.285
– Chronic	377,547	0.020	0.139

Columns: (1) Number of observations in the GWAS sample. (2) Sample mean in the GWAS sample. (3) Sample standard deviation in the GWAS sample. Rows: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation, (4)-(5) splits (3) into acute and chronic causes of respiratory hospitalisation. Fluid intelligence is standardised but during the GWAS routine a small number of individuals are discarded from the sample and this causes the mean and variance to differ slightly from zero and unity above.

Table D.2: Predictive power of polygenic scores. Linear regression estimates of the main outcomes regressed on their corresponding polygenic score.

	Dependent variable:				
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any	(4) Respiratory, acute	(5) Respiratory, chronic
Polygenic score	0.723***	0.631***	0.014***	0.009***	0.005***
	(0.008)	(0.012)	(0.001)	(0.001)	(0.000)
Observations	64,681	26,877	64,923	64,920	64,920
R^2	0.112	0.097	0.005	0.004	0.002
Incremental R^2	0.100	0.087	0.002	0.001	0.001

Columns:(1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation, (4)-(5) splits (3) into acute and chronic causes of respiratory hospitalisation. We only include sex and the first 20 genetic principal components as covariates. All specifications contain year-of-birth and month-of-birth fixed effects. Standard errors are heteroskedasticity robust. The incremental R^2 is the increase in R^2 relative to a null model excluding the polygenic score. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.