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

Stephanie von Hinke **Emil Sorensen**

University of Bristol

EEA

August 2022

- Growing literature on contemporaneous effects of exposure to air pollution on individuals' human capital and health outcomes
  - Little empirical evidence on longer-term, cumulative effects of early-life pollution exposure
  - Despite the importance of the prenatal and early childhood period Almond and Currie (2011a); Almond and Currie (2011b)
- Lack of studies on very long-term effects of early life pollution exposure
  - Absence of high-quality historical pollution data
  - Implying most studies look at effects on child birth outcomes
  - A handful of studies explore effects in early adulthood or older age
- Ignoring such long-term effects may lead to underestimation of the total welfare effects caused by pollution

- We overcome the lack of historical pollution data by relying on reduced form analysis
- Exploit a severe pollution event: the London Smog
  - Unprecedented accumulation of pollutants in December 1952
  - Residential/industrial chimneys, vehicle exhausts, coal burning
  - Trapped under a layer of warm air due to a thermal inversion
  - Only 5 days: 5–9 December
- Similar to pollution levels currently reported in industrialising economies

- Examine the long-term effects of exposure to this smog on individuals in older age
- Identification exploits spatio-temporal variation in exposure using a difference-in-difference approach
- Identification similar to Bharadwaj et al. (2016) and Ball (2018), but:
  - Explore effects on human capital and health outcomes
  - Identify the gestational ages that are most sensitive to pollution
  - Examine heterogeneity of treatment effects wrt 3 sources of variation: genetic variation, gender, and SES

Two strands of literature

- Developmental Origins of Health and Disease (DOHaD) hypothesis
  - Early-life circumstances can have lifelong, irreversible, effects on individuals' health and well-being
  - Large and growing literature explores causal developmental origins
  - Most studies on pollution investigate effects on birth outcomes
  - Only few study effects in early adulthood and even fewer focus on outcomes in older age
- Our contribution:
  - Very long-term effects (age $\sim$60)
  - Identify gestational ages that are most sensitive

Two strands of literature

- 'Nature-nurture' ($G \times E$) interplay
  - Increasingly accepted that the two are inextricably linked
  - Large literature on importance of $G \times E$, with recent contributions from economics/social science (e.g., Schmitz and Conley, 2016; Bierut et al., 2018; Papageorge and Thom, 2020; Fletcher, 2012)
  - Most studies use endogenous environments
    - G and E are not independent (*rGE*, genetic nurture)
    - Unclear how to interpret parameters
- Our contribution:
  - Use London smog as a natural experiment
  - Ensure environment is orthogonal to observed and unobserved individual characteristics

Preview of the findings

1. Early-life pollution exposure matters for later-life human capital and health outcomes
2. Both prenatal and childhood exposure reduce human capital
3. Prenatal exposure increases probability of respiratory disease
4. Evidence of trimester-specific effects
5. Negative effects are stronger for those genetically predisposed
6. Lower social classes more adversely affected

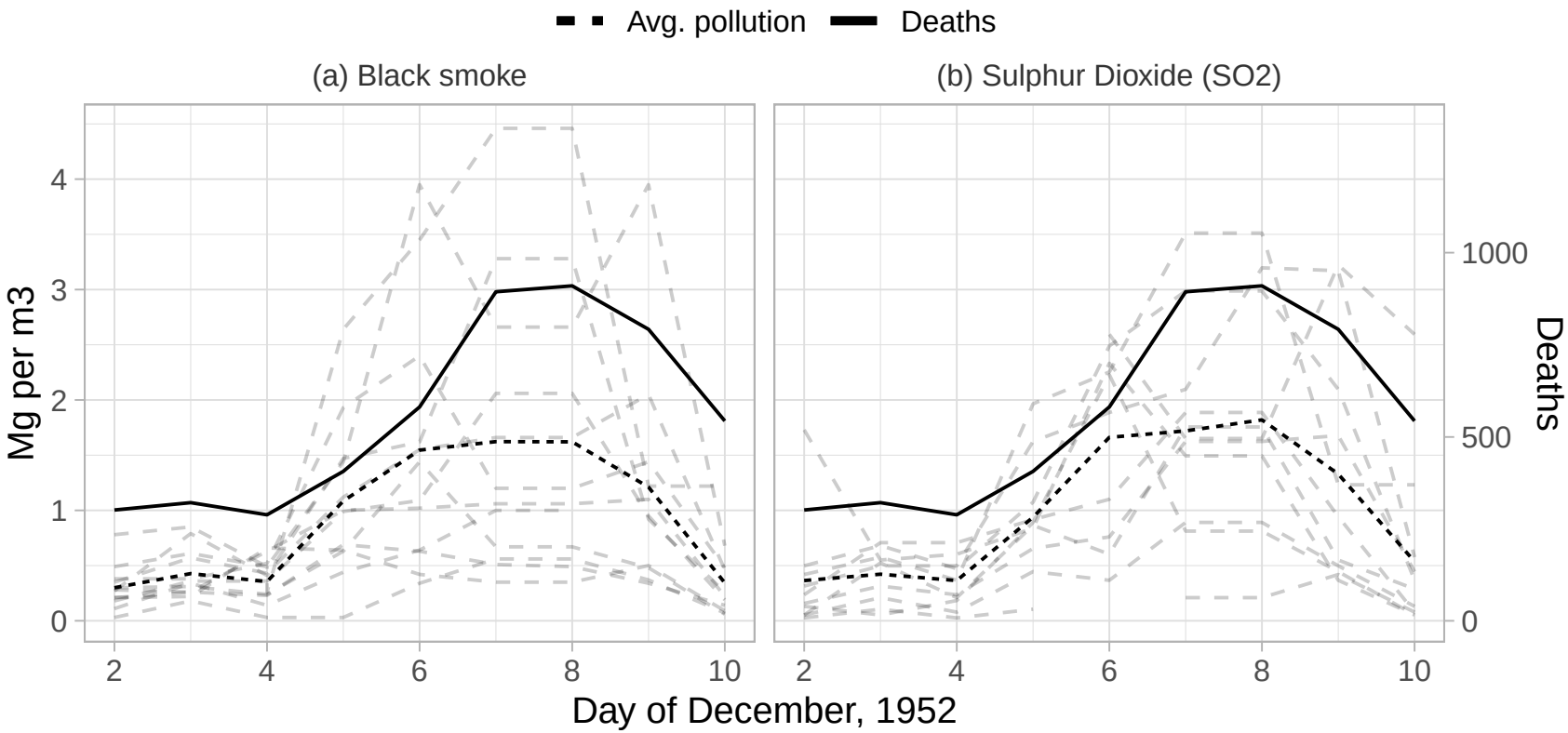
Background – The London Smog

The London Smog

London was covered by a thick smog between 5–9 December 1952

- A temperature inversion caused the cold air to be trapped under a layer of warm air. Very little wind.
- Led to an unprecedented accumulation of pollutants
  - Residential and industrial chimneys, vehicle exhausts and coal burning
  - Severe increases in the only two pollutants measured at that time: black smoke & SO_2
  - Likely also increases in carbon monoxide, carbon dioxide, sulphuric acid, and tar (Wilkins, 1954).
- Worse than Londoners ever experienced before
  - Reduced visibility and penetration into indoor areas
  - Public transport suspended, flights diverted, ambulances stopped, concerts, theatres cancelled

Background – The London Smog



Pollution and mortality during the London smog. Historical measurements of pollution (black smoke and SO_2) from stations in London in December 1952. 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 (1953).

The London Smog

- Medical statistics showed a substantial increase in mortality
  - Estimated 4,000 excess deaths, half of which attributed to bronchitis or pneumonia (Logan et al., 1953)
  - 60% among those 65+; 90% of these among those aged 45+
  - Small increase in mortality among newborns/infants
- Black smoke and SO_2 released into atmosphere via fuel combustion (e.g., coal burning)
- Black smoke is similar to PM2.5
  - Small particles that can penetrate into lung tissue
  - Interfere with transfer of oxygen to the blood
  - Passed through the placenta to developing fetus

Individuals

- UK Biobank (Fry et al., 2017; Bycroft et al., 2018):
  - Approx 500,000 individuals living in the United Kingdom
  - Assessed in 2006–2010 when participants were 40–69 years old
  - Demographics, physical and mental health, health behaviours, cognition, biomarkers, and economic outcomes
  - Obtained via questionnaires, interviews, nurse measurements
  - Links to GP, hospital records, and National Death Registry
  - All individuals have been genotyped
- Weather data from the MET Office
  - Link individuals' location of birth to nearest grid point
  - For period of the smog, merge in:
    - Minimum ambient temperature, average sunshine, average rainfall, wind speed

Outcomes

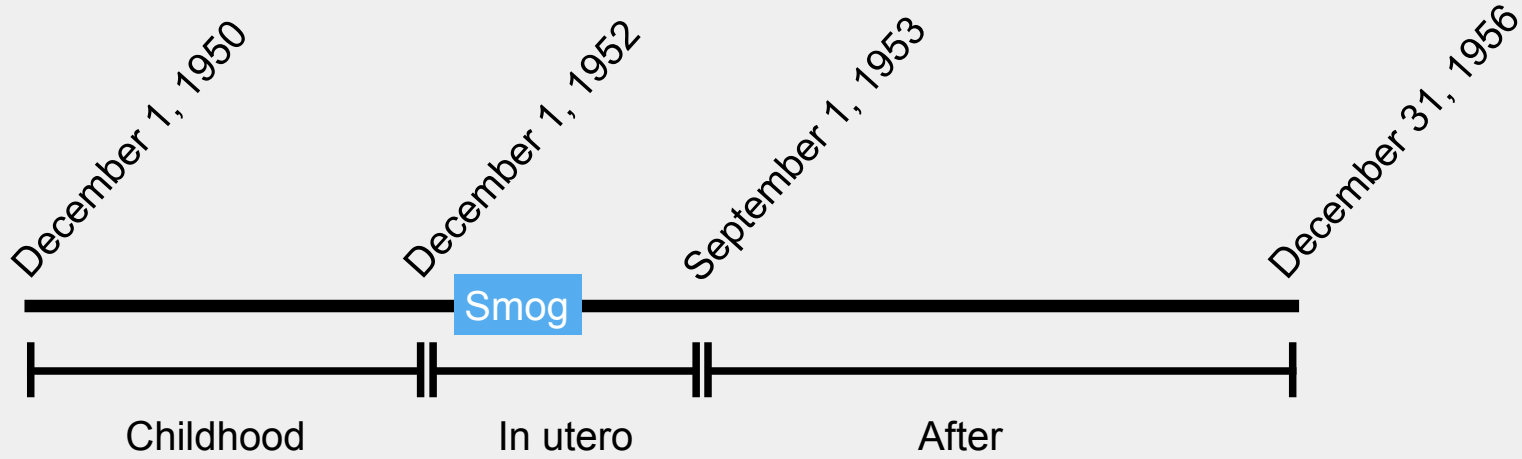
Motivated by the existing literature on

- Effects on human capital outcomes (e.g., Almond, 2006)
  - **Educational attainment**, defined based on qualifications (as in, e.g., Lee, 2018)
  - **Fluid intelligence**, score based on problem solving questions that require logic and reasoning ability
- Effects on health outcomes (e.g., Currie and Walker, 2011)
  - **Respiratory disease**, measured using hospitalisations
  - **COVID-19**, measured using hospitalisations/deaths

Treatment

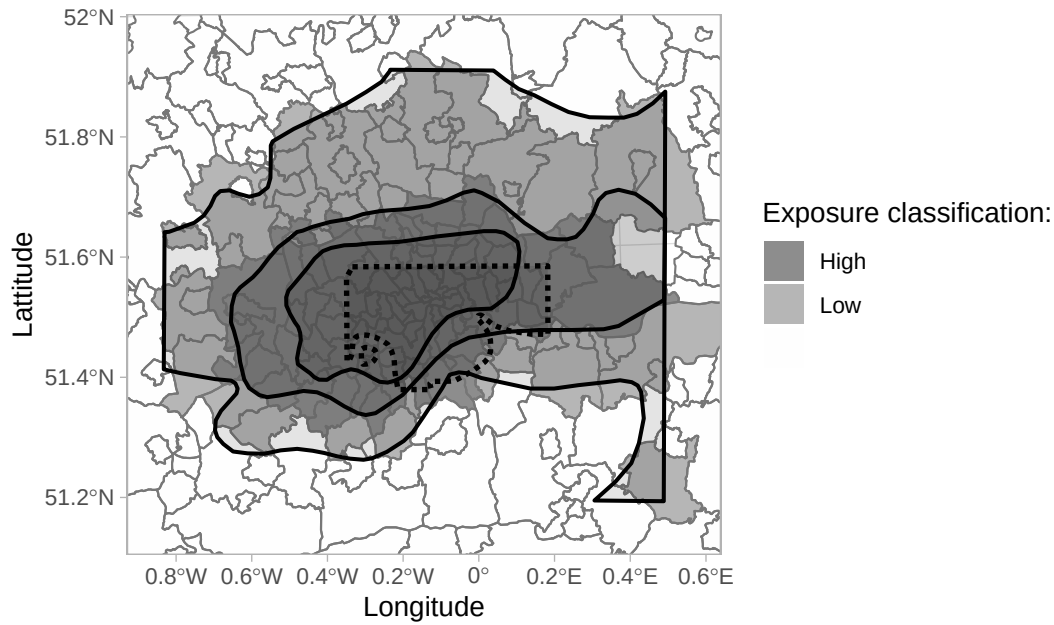
Use tempo-spatial information to assign 'treatment' status

- Split sample along the time (year-month of birth) dimension:
  - Define those exposed **in childhood**, **in utero**, and **not**, based on whether prenatal period precedes, overlaps, or follows the smog
  - For this, assume each pregnancy took 9 months
  - Cut off in December 1956 (RoSLA, SCA)



Treatment

- Split sample along the spatial dimension:
  - Identify one of $>$ 1400 Local Government Districts of birth in E&W
    - Overlay visibility and SO_2 from Wilkins (1954) onto district shapefile
    - Exposed districts are those with reduced visibility and high SO_2



- **Treated**: Exposed (high and low) districts
- **Control**: Other urban districts with population density $>$ 400 per km^2

Sample

- Restrict our sample to:
  - Individuals born in England or Wales with valid birth coordinates
  - Individuals born between 1950–1956 in treated/control districts
  - Those with white European ancestry, ensuring a genetically homogeneous sample
- Final sample includes between 27,000–65,000 individuals

Table: 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).

Specification

- Difference-in-difference using spatial and time variation:
  - Define three groups across time: Childhood, *in utero*, and after
  - Define two groups across space: Treated and control districts

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

  - Individual i, born in district j (admin county k) at year-month t
  - L_i: Relevant London (treated) districts
  - $\mathbf{X}_i$: Controls for sex, weather, month of birth
  - Year and district fixed effects, and district-specific linear trends
  - Standard errors are clustered by district

Compare Y_{ijt} for those exposed at different ages to those conceived after the smog in treated districts, relative to others born at the same time, but in control districts

Potential concerns

- Common trend
  - Parallel trends in treated and control after event ⏩ Trends
  - New difference-in-difference estimators ⏩ Robust estimates
    - Borusyak et al. (2021), Callaway & Sant-Anna (2021), Goodman-Bacon (2021), de Chaisemartin & d'Haultfæuille (2020).
- Anticipation, avoidance behaviour
  - Unlikely prior to smog: no awareness of effects of pollution
  - 'Reverse' difference-in-difference: Control cohorts must have same outcomes as they would have had in the absence of the smog.
- Fetal selection
  - Evidence of increased infant mortality and foetal loss
  - If stronger children survived, suggests estimates are lower bound
- Measurement error in timing, exposure, location

Results - Human capital outcomes

Table: Human capital outcomes — 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. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

Results - Health outcomes

Table: Health outcomes — 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. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

Results - Trimester-specific effects

Table: 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. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

Overview

- Genetic heterogeneity: ⏩ Qualifications
  - Explore treatment effect heterogeneity with respect to individuals' genetic variation
    - Can one's genetic variation 'protect' against adverse circumstances?
    - Does it 'exacerbate' the effect?
  - We use polygenic scores (PGS) to measure one's 'genetic predisposition' towards the outcome.
- Heterogeneity by SES: ⏩ SES differences
  - Low SES groups experience larger adverse impact of pollution exposure
- Heterogeneity by sex: ⏩ Sex differences
  - Stronger impact of pollution on educational attainment for females, but otherwise little differences across sex.

Results – Genetic heterogeneity

Table: 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)

Panels: (a) subsample with above-median polygenic score. (b) subsample with below-median polygenic score. Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

Overview

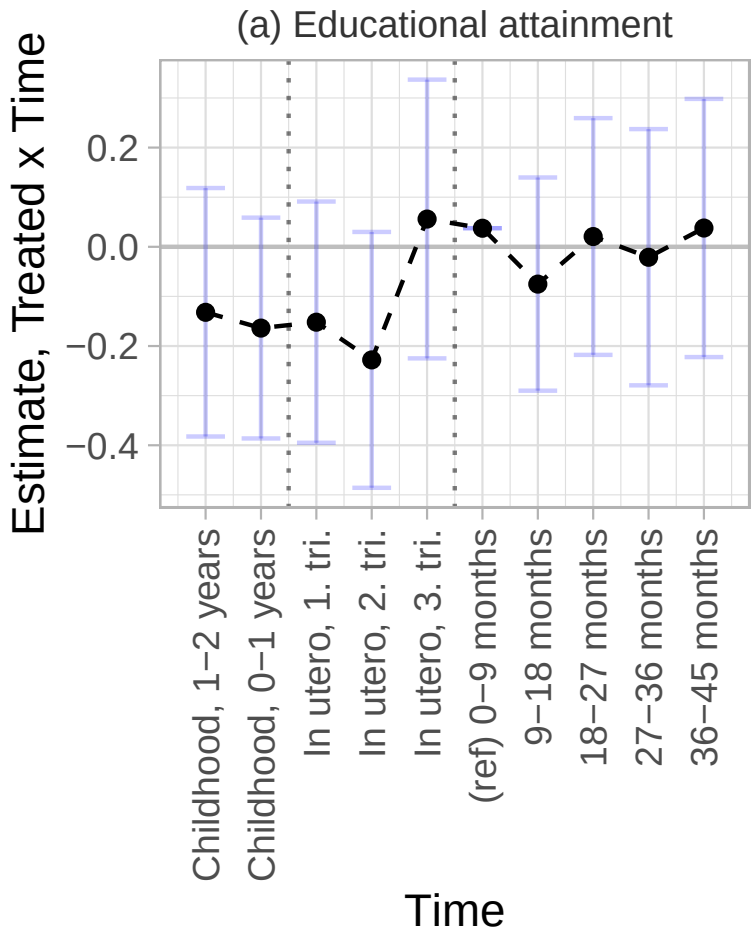
- Definition of treated and control districts
  - Dropping 'low exposure' districts ⏩ Drop low exposure districts
  - Define exposure based on birth coordinates ⏩ Birth coordinates
  - Control cities instead of urban districts ⏩ Control cities
  - Definition of urban control districts ⏩ Urban cutoff
- Definition of 'after' birth cohorts ⏩ After cutoff
- Robustness to specification of trends ⏩ Trend specifications
- Sensitive age in childhood? ⏩ Childhood ages 0-1 and 1-2
- Sex and SES differences ⏩ Sex differences ⏩ SES differences
- Robust estimates using Borusyak et al. (2021) ⏩ Robust estimates

- New evidence on the effects of pollution exposure on human capital and health
  - Rely on reduced form analysis and natural experiment
  - Among the first to estimate the very long-term pollution effects
- Those exposed to the smog have worse respiratory health, lower fluid intelligence, with suggestive evidence of fewer years of education
  - Differential effects by gestational age of exposure
  - Treatment effect heterogeneity: genetic predisposition, gender, SES
  - Highlights the joint role that nature and nurture play in shaping individuals' outcomes

- Interpretation of estimates:
  - Severe pollution event: Current London pollution not comparable, but it compares to current smogs in industrialising economies
  - Estimates capture the effect of being exposed to a large pollution shock, relative to already high levels throughout early childhood
  - Estimates likely to be a lower bound of the 'true' pollution effect
    - Selection (in utero / infant)
    - Selection of survival until age 40+
    - Measurement error (in timing, exposure, location)

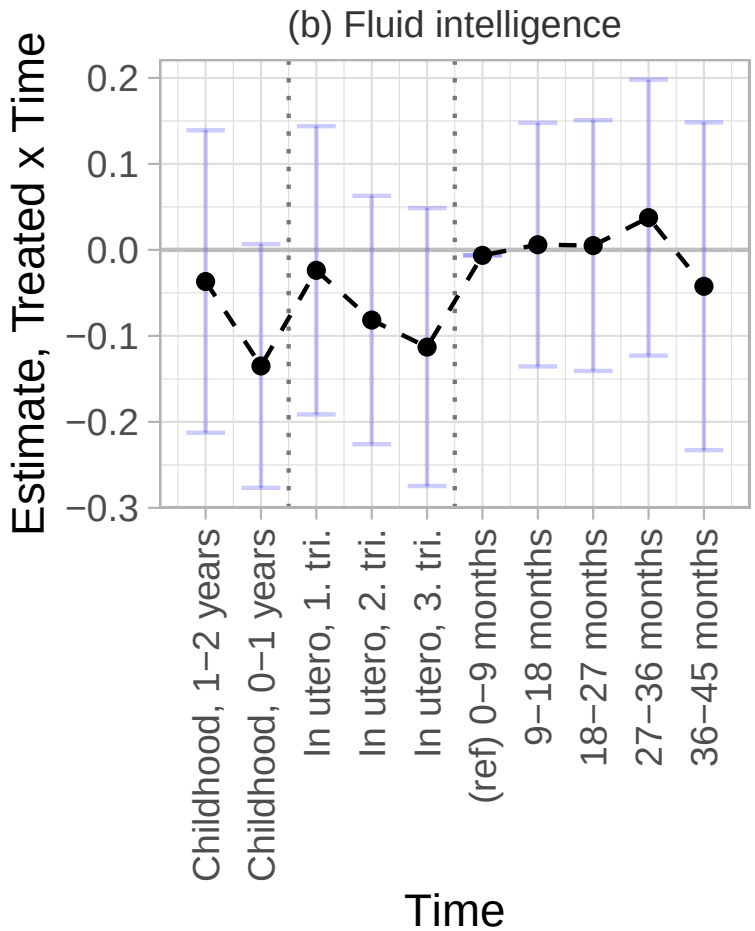
- Limitations
  - Reduced form: Cannot identify which pollutants matter more
  - Generalisability of UK Biobank sample
- Policy implications
  - Large, long-term, benefits for the population
  - Population: Forecasting allows us to avoid heavy pollution
  - Policy: Encouraging evidence to implement incentives/regulation that aim to reduce pollution

Robustness — Parallel trends



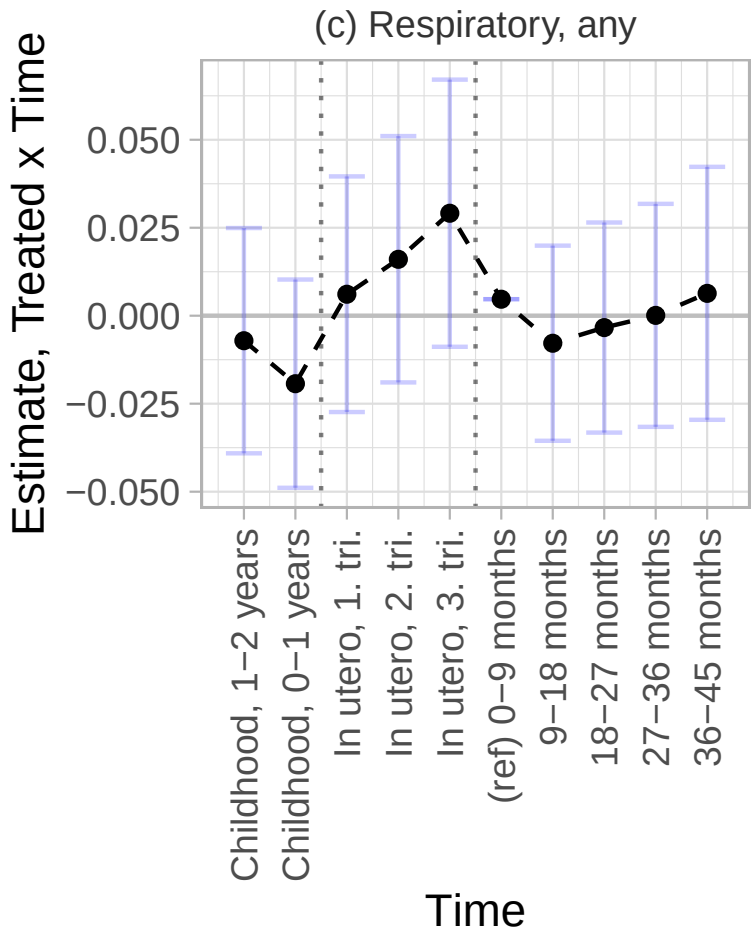
⏩ Back to 'Empirical Strategy' ⏩ Back to 'Robustness Overview'

Robustness — Parallel trends



Back to 'Empirical Strategy'
Back to 'Robustness Overview'

Robustness — Parallel trends



⏩ Back to 'Empirical Strategy'
⏩ Back to 'Robustness Overview'

Robustness — Dropping 'low exposure' districts

Table: Definition of exposure. 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) – 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)

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. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

⏩ Back to 'Robustness Overview'

Robustness — Control cities instead of urban districts

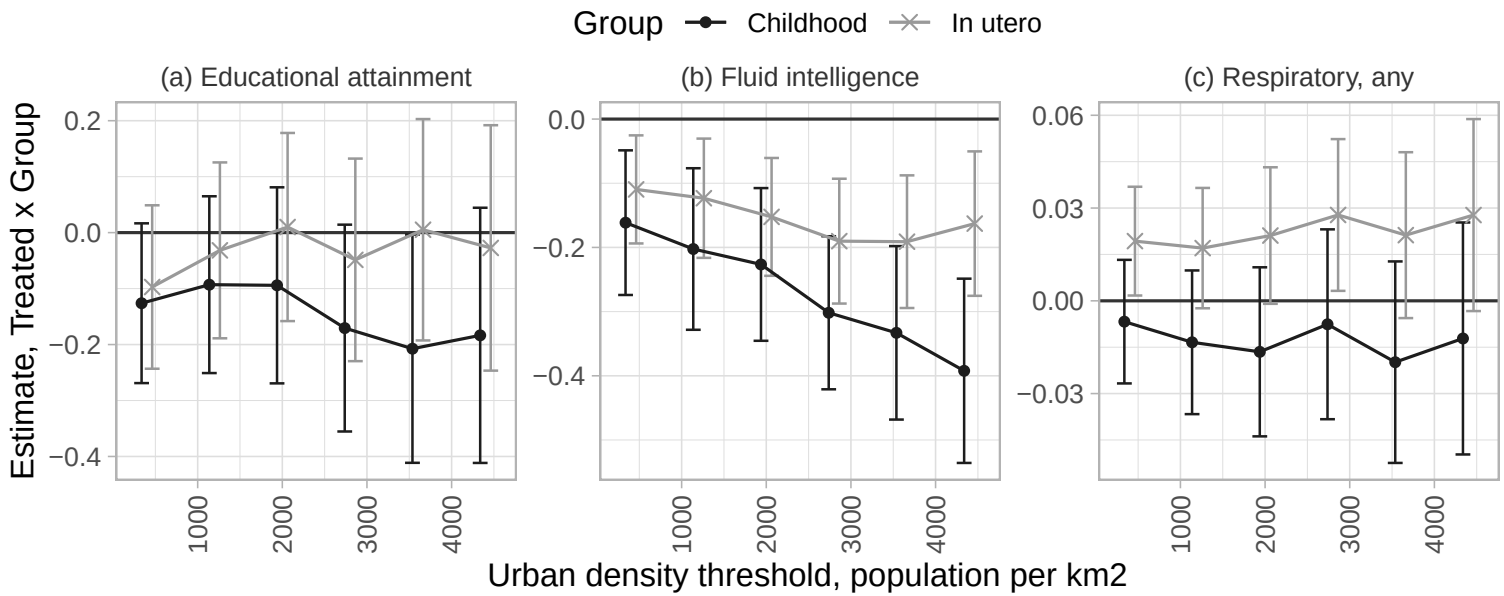
Table: Difference-in-Difference estimates comparing treated to control cities.

	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

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. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

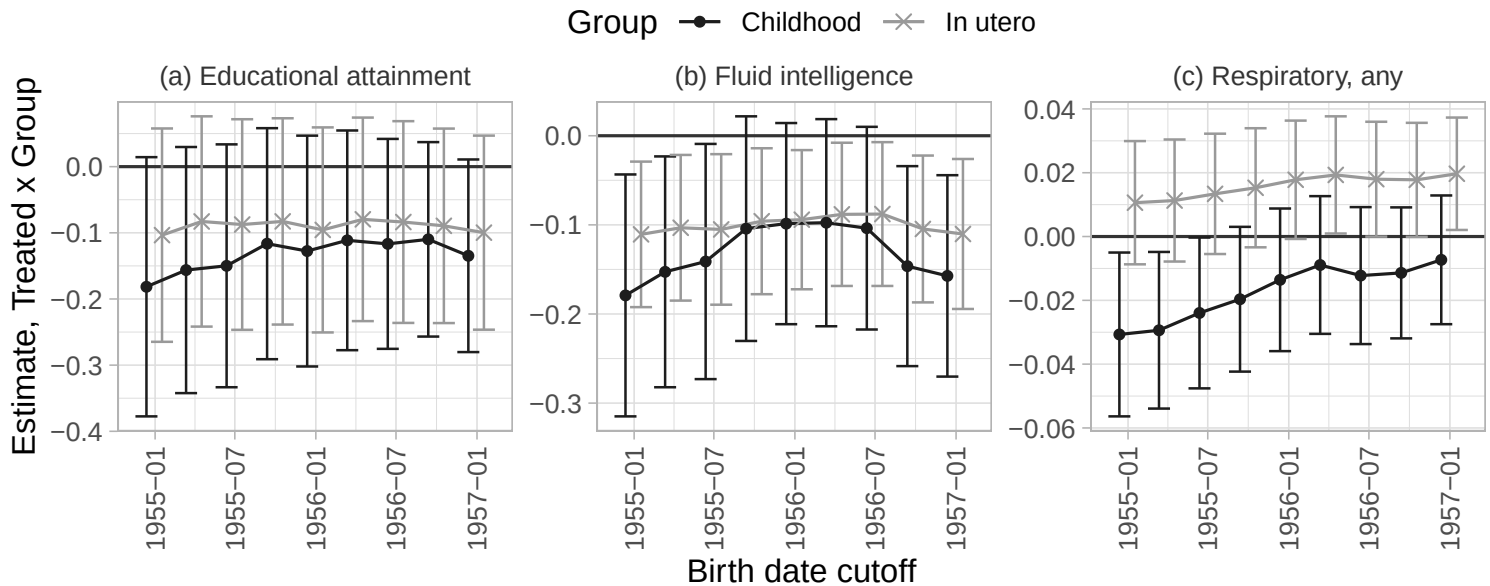
⏩ Back to 'Robustness Overview'

Robustness — Definition of urban control districts



Back to 'Robustness Overview'

Robustness — Definition of ‘after’ birth cohorts



Back to ‘Robustness Overview’

Robustness — Choice of trends

Table: 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)

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.

⏩ Back to 'Robustness Overview'

Robustness — Childhood effects

Table: 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

Standard errors are clustered by district. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

⏩ Back to 'Robustness Overview'

Heterogeneity — Sex differences

Table: 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)

Panels: (a) female subsample, (b) male subsample. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

⏩ Back to 'Heterogeneity' ⏩ Back to 'Robustness Overview'

Heterogeneity — SES differences, high SES groups

Table: 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)

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. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

⏩ Back to 'Heterogeneity' ⏩ Back to 'Robustness Overview'

Heterogeneity — SES differences, low SES groups

Table: 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 (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)

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. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

⏩ Back to 'Heterogeneity' ⏩ Back to 'Robustness Overview'

Heterogeneity — Qualifications

Table: 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)	(2)	(3)
	Upper secondary	Lower secondary	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)

Panels: (a) high PGS subsample, (b) low PGS subsample. Includes fixed-effects for district, month of birth, and year of birth. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

⏩ Back to 'Heterogeneity' ⏩ Back to 'Robustness Overview'

Table: Robust estimates of average treatment effects (ATTs) using estimator from Borusyak et al. (2021). Compares treated to control districts defined as urban England and Wales.

	Dependent variable:		
	(1) Educational attainment	(2) Fluid intelligence	(3) Respiratory, any
ATT, In utero	−0.049	−0.137	0.025***
	(0.075)	(0.089)	(0.009)
ATT, Childhood	−0.043	−0.161**	−0.001
	(0.047)	(0.069)	(0.008)

Columns: (1) educational attainment in years, (2) standardised fluid intelligence score, (3) ever experienced a (primary) respiratory hospitalisation. Reports the robust estimates of the ATTs for the in utero and childhood cohorts when we assign equal weights across units in these cohorts. (*): $p < 0.1$, (**): $p < 0.05$, (***): $p < 0.01$.

⏩ Back to 'Empirical Strategy' ⏩ Back to 'Robustness Overview'