Prenatal exposure to pollution, educational achievement and parental investments

Giovanni Scotti Bentivoglio*[1], Gerard van den Berg[1 2], and Melinda Mills[1 3]

[1]University of Groningen
[2]University Medical Center Groningen
[3]University of Oxford

January 2025

Abstract

Prenatal exposure to air pollution has been shown to have adverse short- and long-term effects on health and other several dimensions of human capital, including educational achievement. The overall effects on these outcomes, however, are not only the result of biological mechanisms, but also depend on other inputs, such as parental investments during the child's development. Using data from Germany, this study investigates the effect of prenatal exposure to air pollution on the transition rate of elementary school students to the academic track (*Gymnasium*) and on secondary school grades, and assesses directly how parental activities and involvement in early childhood moderate this effect. To address concerns related to the endogeneity of pollution exposure, we leverage variations in the height of the planetary boundary layer, an exogenous atmospheric variable strongly predicting pollution levels. We find that higher in-utero exposure to air pollution significantly decreases the likelihood of transition to the academic track. This effect is concentrated among low-SES children, since high parental education partially offsets the adverse effect of pollution exposure. Moreover, we find that parental activities can significantly moderate the effect of pollution exposure, but while educational activities, such as reading frequently with the child, strongly mitigate the adverse effects of pollution, recreational activities, such as watching tv with the child, have an aggravating effect.

JEL Codes: Q53, I24, D13, J13

Keywords: air pollution, fetal exposure, academic track, socioeconomic status, parental activities

*Department of Economics, Econometrics & Finance, Nettelbosje 2, 9747 AE, Groningen. Email: g.scotti.bentivoglio@rug.nl.

1 Introduction

According to the fetal origins hypothesis (Barker, 1990), conditions at birth can significantly shape a person's health trajectory and adverse environmental exposures in-utero can have substantial negative long-term impacts (Almond, Currie, and Duque, 2018). In this framework, a growing number of studies have examined the effect of prenatal exposure to air pollution on several later life socioeconomic outcomes (e.g. cognitive skills, educational achievement, income) in addition to health, finding significant negative effects. The extent to which exposure to pollution in-utero affects these outcomes, however, depends largely on how parents invest on the child's development during the first years of life.

Combining data from a longitudinal household survey and weather reanalysis datasets, this paper investigates the causal effect of prenatal exposure to pollution on educational achievement in Germany, and directly assesses how parental activities carried out with the child during early childhood moderate this effect. To measure educational achievement, we focus on school grades at age 12-14 and on the transition of elementary school students to the academic track in secondary school (*Gymnasium*). In light of the rigid early tracking of the German education system, which is in place since the beginning of secondary school, the secondary school track is a meaningful indicator of the performance in primary school but also a strong predictor of the educational trajectory of the child, being highly correlated with the enrolment and completion of tertiary education (Dustmann, 2004; Müller and Schneider, 2013).

To address potential endogeneity in the exposure to air pollution, related to sorting and avoidance behaviours, we leverage exogenous variations in the height of the planetary boundary layer (BLH), an atmospheric variable which strongly affects the concentration of pollutants on the ground.

Our results show that children exposed to a lower BLH on average in-utero (and thus worse pollution) are less likely to enroll to the academic track, while we do not find significant effects on school grades. This result is in line with previous studies finding negative effects of prenatal pollution exposure on educational achievement and cognitive skills in developed and developing countries (Sanders, 2012; Bharadwaj et al., 2017; Molina, 2021). We then investigate whether the size of the effect changes by demographic characteristics and socioeconomic status, finding that parental education strongly mitigates the adverse effects of exposure. A possible explanation is that higher educated parents, who are also likely to have better financial means, are better able to invest in the child's development.

After having established a negative socioeconomic gradient in the effect, we assess how a specific set of parental activities moderate the effect of pollution exposure on the likelihood of transition to the academic track. We find that frequently carrying out recreational activities with the child during early life, such as singing songs and watching tv and videos, reinforces the baseline effect of pollution exposure on educational achievement, while frequently carrying out an educational activity such as reading stories can mitigate the adverse effect of pollution. The effects of these activities are more relevant among lower educated households.

Our study aims to contribute to several strands of research. First, we contribute to the literature on the role of parental involvement in the child's human capital development. A few studies finding a socioeconomic gradient in the effect of pollution on educational achievement hint at the role of parental involvement as a possible explanatory mechanism, but cannot explicitly test it: we directly assess the moderating role of a set of parental activities carried out during early childhood, providing new insights on the effects of these activities on the child's educational achievement. We consider simple and common activities with a varying degree of parental involvement, that are carried out long before the outcomes of interest are measured.

Second, we measure educational achievement as the transition rate to the academic track, using an alternative measure to those commonly used in the related literature, such as low-stake school or cognitive test scores. This outcome is very specific to the German context and, even though it is observed during early childhood, it has relevant long-term implications on various future socioeconomic outcomes and on socioeconomic inequalities.

Third, this study adds to the growing number of recent studies instrumenting the variation in pollution levels with atmospheric variables or conditions. Specifically, while most of these studies instrument air pollution with thermal inversions (Arceo, Hanna, and Oliva, 2016; Jans, Johansson, and Nilsson, 2018; Molina, 2021; Colmer et al., 2021; Thies, 2024), we exploit plausibly exogenous variations in the boundary layer height (Schwartz, Bind, and Koutrakis, 2017; Godzinski and Castillo, 2021), a measure that is becoming increasingly available through atmospheric reanalysis models. While the former is usually measured as the occurrence of a thermal inversion, regardless of its intensity, boundary layer height is a continuous variable, and as such it can indicate the strength of the phenomenon.

The rest of the study is structured as follows: Section 2 provides a theoretical background on the effects of fetal exposure to air pollution on educational achievement and gives an overview of the different identification strategies used to estimate a causal effect in the related literature; Section 3 describes the data sources used to perform the analysis; Section 4 illustrates the empirical strategy and describes the atmospheric instrument; Section 5 illustrates the main results of the study, giving also an overview of the robustness checks performed; Section 6 concludes.

2 Theoretical background

2.1 Air pollution and educational achievement

There is substantial medical and epidemiological evidence documenting how pollutants can damage fetal development. The most cited harmful pollutants are particulate matter (PM2.5 or PM10) and carbon monoxide (CO). Particulate matter can circulate in fetal blood, inducing oxidative stress and inflammations in the fetus and increasing the risk of brain development and cognitive disorders after birth (Guxens et al., 2018; Chen et al., 2021). Carbon monoxide binds to hemoglobin and limits the body's ability to deliver oxygen to vital organs and tissues (Gorman et al., 2003; Smithline et al., 2003); this can affect the fetus both indirectly, as the mother transmits less oxygen, and directly, as CO can cross the placental barrier to circulate in fetal blood and gain access to the developing brain, with long-term consequences on neurocognitive functions (R. J. Levy, 2015).

While exposure to polluted air has serious, even short-term effects on health and cognitive functions throughout all the life course (e.g. Godzinski and Castillo, 2021; Bernardi and Keivabu, 2024), gestation is particularly critical, as brain formation proceeds very rapidly and the damage to the nervous system due to exposure in this period can be irreversible (Prado Bert et al., 2018).

Drawing from these premises, an expanding economic literature related to the fetal origins hypothesis has examined the effects of prenatal exposure to air pollution on human capital and socioeconomic outcomes. These include earnings and labour market performance (Isen, Rossin-Slater, and Walker, 2017; Black et al., 2019), educational achievement (Bharadwaj et al., 2017; Marcotte, 2017), crime (Grönqvist, Nilsson, and Robling, 2020), but also measures of cognitive skills, such as fluid intelligence (Molina, 2021) as well as measures of non-cognitive skills, such as neuroticism (Thies, 2024).

The negative effects of pollution are particularly relevant for the more vulnerable groups: a few studies have assessed the presence of a socioeconomic gradient in the effect of prenatal pollution exposure on the outcomes mentioned above, finding low socioeconomic status (SES) children to be the most negatively affected group (Jans, Johansson, and Nilsson, 2018; Thies, 2024). This suggests that the effect of prenatal exposure to air pollution on these outcomes is not only the result of biological channels (i.e. the direct effect on health, cognitive and non-cognitive skills) but largely depends also on other inputs, including parental involvement during the child's early development.

While the role of parental involvement is acknowledged as a possible channel explaining differential effects of prenatal pollution exposure on human capital formation, it has been hardly examined directly, particularly in the context of a developed country. Conversely, a significant number of studies in the fetal origins literature have directly examined how parents adjust investments in response to the child's conditions at birth and in early life. In particular, these studies assess whether parents increase or decrease their investments in response to early insults to the child (i.e. compensating or reinforcing investments) and whether this depends on the socioeconomic status of the parents (Hsin, 2012; Restrepo, 2016; Fan and Porter, 2020). However, the results are mixed and depend to a large extent on the methodology and the research setting considered.

The recent study by Zhang et al. (2024) is, to our knowledge, the only study looking directly at the role of parental investments in a framework similar to ours. The authors estimate the effect of prenatal pollution exposure on cognitive test scores in China, finding worse effects for females than for males. To explain this gradient, they show that parents tend to be more involved with their male child (i.e. they help more with homework) when exposed to higher pollution in-utero. Differently from their study, we do not find that parents adjust their activities in response to pollution exposure; therefore, in our setting, we only assess parental activities as moderators. Moreover, the activities we examine are carried out in early childhood and before the start of formal education.

2.2 Identification issues and atmospheric instruments

The main concern for the identification of the effects of air pollution exposure on human capital and socioeconomic outcomes is that air pollution exposure is not random across locations and households. First, households from a higher socioeconomic status could have higher preferences for clean air, thus considering the level of pollution of an area when deciding where to live. Supporting this hypothesis, several studies show that air quality is capitalized in house prices (Chay and Greenstone, 2005; Sullivan, 2016). Residential sorting with respect to air quality may bias the estimates of the effect of pollution on educational achievement if high SES parents are more likely to invest in their children's schooling and development. Moreover, the actual exposure to air pollution, and its consequent effect on fetal development, depends to a large extent on the awareness and behavioural responses of the parents. These could consist in defensive investments, for example in pharmaceutical products (Deschenes, Greenstone, and Shapiro, 2017), or in avoidance behaviours, by reducing the time spent outdoors in response to smog alerts (Neidell, 2009; Moretti and Neidell, 2011).

In order to address the potential endogeneity of pollution exposure, past studies have employed a variety of empirical strategies. A number of studies have relied on within-family analyses with mother fixed effects, focusing on within-location temporal variations (e.g. Currie, Neidell, and Schmieder, 2009; Coneus and Spiess, 2012; Bharadwaj et al., 2017). Using Chilean data, Bharadwaj et al. (2017), find strong negative effects of fetal exposure to carbon monoxide and particulate matter on fourth-grade test scores.

Other studies have leveraged nation-wide policies tackling air pollution (Chay and Greenstone, 2003a; Isen, Rossin-Slater, and Walker, 2017), or the more recent introductions of Low Emission Zones at the local level (Brehm et al., 2022; Conte Keivabu and Rüttenauer, 2022). Isen, Rossin-Slater, and Walker (2017) exploit changes in air pollution driven by the Clean Air Act, and compare cohorts born before and after the enforcement of the reform, finding significant long-term effects of the reform on labour market performances.

In a similar way, economic shocks have been considered as exogenous sources of variation in pollution levels for limited periods of time. For example, the reduction in prenatal exposure to air pollution during the US industrial recession of the early 1980s has been shown to have short and long term beneficial effects, reducing infant mortality (Chay and Greenstone, 2003b) and increasing high school test scores (Sanders, 2012).

More recently, thanks to the increasing availability of atmospheric data, a few studies have used atmospheric events and characteristics, such as thermal inversion and planetary boundary layer height, to instrument air pollution (Jans, Johansson, and Nilsson, 2018; Godzinski and Castillo, 2021; Molina, 2021; Thies, 2024). The use of these measures comes from the atmospheric sciences literature, which shows that these atmospheric phenomena are important determinants of the dispersion and the concentration of pollutants on the ground (Jacobson, 2002; Levi et al., 2020). Being determined by vertical air layers, these atmospheric phenomena are plausibly not related to human activities and business cycles, but are purely meteorological in nature.

Using atmospheric mechanisms as exogenous sources of variation in pollution levels provides several advantages over other identification strategies: first, the theoretical effect of the instruments on air pollution is clear and widely acknowledged; second, differently from within-family studies, this strategy relies on cross-family comparisons, thus using larger and more representative samples; third, differently from studies that leverage specific policies or events, since these atmospheric phenomena are always present and regular in time, this strategy does not require narrow time windows around a cutoff for identification, thus allowing longer time periods for the analysis.

In terms of research setting and methodology, the closest related work is the recent study by Thies (2024), which examines the effect of prenatal pollution exposure on noncognitive skills during childhood using SOEP data. Thies instruments air pollution levels with thermal inversions and finds that higher exposure in utero increases neuroticism and negatively affects emotional stability at ages 5-10. Besides looking at a different outcome and using a different atmospheric measure, we investigate in depth the role of parental involvement during early childhood, assessing how it moderates the effects of pollution on educational achievement.

3 Data

This section describes the data sources employed to construct the final dataset, which matches socioeconomic and child information with pollution and atmospheric data based on the location and time of birth of the children. Our sample of interest includes children born between 2000 and 2010.

3.1 Socioeconomic data

Socioeconomic data comes from the German Socio-Economic Panel Study (SOEP, Goebel et al., 2019). SOEP is a large longitudinal household survey, started in 1984, covering approximately 30000 people from 20000 households every year. It contains detailed socioeconomic and demographic information at the household and individual level and closely follows parents and children over time.

Outcome variables For our study, the main outcomes of interest are measures of educational attainment in late childhood and early teenagehood, namely school grades in German and Maths at ages 12 and 14, and whether the student is enrolled in the academic track in secondary school, *Gymnasium*. Concerning the former outcomes, the national grading system from primary school through high school uses a six-point scale, ranging from 1 (excellent) to 6 (insufficient). For ease of interpretation, we reverse this scale. While school grades are available also during primary school (at age 8 and 10), these are not very indicative of school performance and display a very low variation, with values mainly ranging from 1 to 4.

Regarding the latter outcome, a brief overview of the German school system is needed in order to understand the implications of enrolling in the academic track. German education is characterized by a rigorous early tracking (generally at age 10[1]) system, where, based on the performance in the final years of primary school, students are allocated into one of the following secondary school tracks: the lower track, *Hauptschule*, lasting five years, the medium track, *Realschule*, lasting six years and the higher or academic track, *Gymnasium*, which lasts eight or nine years. While lower and medium tracks generally prepare students for apprenticeships and vocational trainings in blue-collar or white-collar occupations, *Gymnasium* offers a more academic program and is the only track granting access to university. Next to the classical three-tiered system, several states also have comprehensive schools (*Gesamtschule*), that offer multiple educational tracks in the same institute. Depending on the educational path chosen within the school, students can therefore obtain equivalent qualifications to those that are obtained in the lower, medium and higher tracks[2].

Students are allocated to one of the secondary school tracks based on the primary school teacher's recommendation, which takes into account the children's performance in the last years of primary school. While this recommendation is not binding in most states[3], parents tend to follow it. While in principle students can change track during the course of secondary school, mobility across tracks, and especially upward mobility, is very rare (Dustmann, Puhani, and Schönberg, 2017). Therefore, the rate of transition to *Gymnasium* is a meaningful indicator of early educational performance that can also have significant implications for later-life socioeconomic outcomes, such as career trajectories and earnings (Dustmann, 2004). This outcome has been studied in relation to the effects of pollution exposure: very relevant for our setting is a recent study by Brehm et al. (2022) which shows that the progressive implementation of Low Emission Zones in Germany significantly increased the rate of transition to the academic track for primary school children.

Parental activities We collect information on parental activities carried out with the child throughout childhood: the *Mother and child* questionnaires ask the mother in multiple waves how frequently she has carried out a series of activities with the child over the past couple of weeks prior to the survey, with values ranging from 'Every day' to 'Never'. Due to data limitations and for the purposes of this study, we focus on a subset of these activities, that are asked when the child is aged 1, 3 and 6:

- Singing songs to or with the child
- Painting or doing arts and crafts with the child
- Reading stories to or with the child
- Watching television or videos with the child

[1]Students start secondary school in the fifth grade, at age 10 in all federal states except for Berlin and Brandenburg, where secondary school starts in the seventh grade, at age 12.

[2]Depending on regional and organizational variations, *Gesamtschule* is referred to by a vast array of different terms, including *Mittelschule*, *Regelschule* and *Sekundarschule* (for schools that combine the lower and medium track) and *Integrierte Gesamtschule*, *Oberschule*, *Stadtteilschule* and *Gemeinschaftschule* (for schools that combine all three tracks).

[3]Teacher's recommendation was binding for the whole period of interest (from 2010 onwards) in Bavaria, Brandenburg, Saxony and Thuringia. Source: `www.kmk.org`

To make the values of parental activities comparable across ages, we standardize the values by age and take the average of the standardized values across ages; we then define a dummy for whether this value is above (below) the mean, which indicates a relatively high (low) frequency of the activity.

Pollution exposure is assigned to individuals based on the county (*Kreise*[4]) of residence at birth and on the date of birth. Information on the county of residence at birth is available for around 60% of the sample, while in the remaining cases, where the households entered the survey after the child was born, we proxy the county of birth with the county of residence in the year in which the household entered the survey. While incorrect assignments for this subsample can bias the estimates, it is unlikely that these are systematically correlated with pollution exposure. Moreover, for German-born children up to age 18, internal (cross-county) migration rates have been shown to be quite low for the period of interest (Sander, 2017; Stawarz and Sander, 2019).

From household and individual questionnaires, we also collect information on a wide range of demographic and socioeconomic characteristics. For the analysis sample, we consider only children from two-parent or single-parent households with nonmissing information on parental education, thus excluding children with unknown family education background. This results in N=5902 children for the *Gymnasium* sample and N=3167 or N=3164 for the German grade sample and the Maths grade sample, respectively. Descriptive statistics are reported in Table 1.

3.2 Atmospheric and air pollution data

The atmospheric data is collected from the ERA5 reanalysis dataset available in the Climate Data Store of the European Centre for Medium-Range Weather Forecasts (ECMWF), for the period 2000-2010. The data is available at an hourly level on a 0.25° latitude x 0.25° longitude grid, which roughly equals a 28 km x 17.5 km grid in Germany. To assign the value of boundary layer height to a county, we average the hourly values over the month and then compute the simple mean of the grids within the county's borders. As illustrated in Table 1, the average height of planetary boundary layer during gestation for the sample is around 575 meters with a standard deviation of 62 meters.

From the same data source, we then collect data at the monthly level on weather conditions, available at the 0.1° latitude x 0.1° longitude grid: total precipitations, surface temperature, and eastward and northward components of the 10-metre wind to compute wind speed. The monthly averages are assigned to the county in the same way as described for boundary layer height.

For the period of interest, we also collect data on monthly average on PM2.5 concentrations,

[4]Counties are the primary subdivision higher than a municipality in Germany and correspond to NUTS 3 administrative units. As of now, Germany is divided in 400 counties: 294 of these are rural (*Landkreis* or *Kreis*) and 106 are urban (*kreisfreie Städte* or *Stadtkreise*): these are usually large cities that do not belong to counties but take on county responsibilities themselves.

available at a 0.1° latitude x 0.1° longitude grid, from the Atmospheric Composition Analysis Group (ACAG). Estimates of PM2.5 concentration at this granular level are obtained thanks to a model that combines satellite-derived estimates with chemical transport models, and calibrates the ground-based observations using deep learning (Shen et al., 2024).

Table 1: Descriptive statistics

	Mean	Std.Dev.	N
Gymnasium	0.43	0.495	5902
German	4.36	0.85	3167
Maths	4.33	0.99	3164
BLH (m)	574.66	62.08	5902
PM2.5 ($\mu g/m^3$)	14.88	2.13	5902
Age at Birth	30.3	5.53	5902
Birth order	1.93	1.09	5902
Migration Background	0.29	0.45	5902
Female	0.49	0.5	5902
Single Parent HH	0.19	0.39	5883
Mother: No high school	0.15	0.36	5885
Mother: High school	0.58	0.49	5885
Mother: University	0.27	0.44	5885
Father: No high school	0.11	0.32	4786
Father: High school	0.5	0.5	4786
Father: University	0.39	0.49	4786

Descriptive statistics for the sample of children for whom there is information on secondary school track (the *Gymnasium* sample). German and Maths are reported for the school grades sample.

4 Empirical strategy

The baseline model relating pollution exposure to educational achievement is given by the following equations:

$$Y_{iymc} = \beta_1 Pol_{ymc} + \beta_2 W_{ycm} + \beta_3 X_i + \alpha_c + \alpha_m + \alpha_y + \epsilon_{iymc}. \tag{1a}$$

$$Y_{iymc} = \beta_1 Pol_{ymc} + \beta_2 W_{ycm} + \beta_3 X_i + \alpha_c + \alpha_m + \alpha_y + \alpha_s + \alpha_a + \epsilon_{iymc}. \tag{1b}$$

Y denotes a dummy for whether the student is enrolled in the academic track in equation 1a and school grades in 1b; Pol is the average air pollution exposure level during gestation (the nine months of pregnancy), where we consider the month of birth as the ninth month; W includes meteorological and weather conditions for the same period (second order polynomials of temperature, wind speed and precipitation levels); X includes child gender, migration background, birth order, maternal age at birth and parental education at birth, for which we set missing values to zero and add dummies for missing values; in equation 1b we also control for age in months and its squared; α captures county, month and year of birth fixed effects; in equation 1b we also include age fixed effects and school track fixed effects (lower, medium, higher track and *gesamtschule*) to (partially) control for the type of school attended by the student; standard errors are clustered at the county level. For school grades, the individuals that are surveyed both at age 12 and 14 are weighted by a factor of 0.5. The main effect is therefore identified by variations in the levels of prenatal pollution exposure between children born in the same county, netting out the average seasonal and yearly trends.

However, for the effect to be causal, exposure to pollution should be uncorrelated with unobserved determinants of educational achievement; this assumption is likely to be violated, since some households may take into account pollution levels when deciding where to live and may change their behaviour in response to pollution levels (residential sorting and avoidance behaviour, respectively); moreover, measurement errors in pollution exposure could bias the estimates. While county fixed effects partially alleviate residential sorting concerns, they do not control for location-specific trends in pollution that may actually be correlated with the outcome.

To address endogeneity concerns, we leverage variations in the height of planetary boundary layer (BLH). Boundary layer height measures the height of the lowest layer of the troposphere, which is in direct contact with the earth's surface. Pollutants disperse within this layer; therefore, the higher the BLH, the larger the air volume available for dispersion and the lower the concentrations of pollutants on the ground. Planetary boundary layer is closely related to thermal inversions, because often during this phenomenon, the temperature, which usually decreases with height, sharply increases at the top of the layer (Godzinski and Castillo, 2021).

The level of BLH is characterized by high seasonal fluctuations, as illustrated in Figure 2, and is influenced by wind characteristics and weather conditions. However, controlling for these conditions and for geographic characteristics, variations in BLH are as good as random (in other words, the random combination of all these conditions), and generate exogenous variations in air pollution. Therefore, variations in BLH should affect educational achievement only through variations in the level of air pollution. Table 5 in the Appendix shows the effect BLH on PM2.5 for the samples of interest (the *gymnasium* sample and the school grades sample). Even after controlling for weather characteristics and fixed effects, BLH is negatively and significantly correlated with PM2.5: a one s.d. decrease in BLH increases PM2.5 by 0.55-1 $\mu g/m^3$. The large difference between the coefficients in the two samples is due to the fact that the school grades sample consists only of students who are 12 or 14 at the time of the survey, and since the data is available until 2021, it comprises only children born between 2002 and 2009. This difference persists even replicating this analysis in the atmospheric data, thus it cannot be attributed to the different spatial composition of the two samples. While this large difference in the coefficients can be seen as a limitation of the atmospheric data, the main results for gymnasium are confirmed also in this alternative sample.

In our framework, air pollution should be thought as the endogenous variable. However, the level of air pollution is the result of the concentration and the interaction of several pollutants, in addition to PM2.5, and variations in BLH are likely to affect those pollutants as well. In other words, since we do not have data on other pollutants, the condition of exclusion restriction of the instrument is not fully satisfied in our setting. Therefore, following the approach of Molina (2021), in what follows we will show the reduced form effects of BLH on educational achievement.

To empirically strengthen the credibility of the exogeneity assumption, we check for the presence of a systematic correlation of BLH with the background household and child characteristics, that are included in the baseline model. Table 6 in the Appendix shows this for the two samples of interest; the instrument is exogenous to all background controls except for paternal tertiary education but only in the *gymnasium* sample, for which the instrument is significantly correlated at the 10% level. This could possibly indicate still some mild form of sample sorting with respect to parental education, but the inclusion of parental education among the set of controls in the baseline model should alleviate this concern.

5 Results

5.1 Main Results

Table 2 shows the reduced form effects of in-utero exposure to pollution on our measures of educational attainment, namely the rate of transition to *Gymnasium* and school grades in secondary school. We find that a 1 standard deviation decrease in BLH decreases the likelihood of enrolment in the academic track by 3.7 percentage points, a sizeable effect considering the baseline share of students in *Gymnasium.*

The size of the coefficient should be carefully interpreted, as this outcome is unique to the German context and hardly comparable to the most common measures of educational attainment in the related literature, such as standardized cognitive tests (Molina, 2021) and national school exams (Bharadwaj et al., 2017; Sanders, 2012). The most relevant study to benchmark our estimate is the study by Brehm et al. (2022), who show that the staggered implementation of Low Emission Zones (LEZ) in North-Rhine Westphalia, the most populated federal state in Germany, increased the rates of transition to the academic track by 0.9-1.6 percentage points.

The larger effect found in our study is plausible for two reasons. First, while their study focuses on the short to medium-term effects of the LEZ introduction (one to six years after implementation) for primary school students, we look at the long-term effects of pollution exposure during gestation, a period when the effects of pollution exposure are particularly harmful and potentially irreversible on key dimensions of fetal development. Second, while the implementation of LEZ reduced concentrations of NO_2 by 1.6-2.1 $\mu g/m^3$ and of PM_{10} by 0.8-1.3 $\mu g/m^3$, we find that a 1 standard deviation decrease in BLH in-utero (around 62 m) increases concentrations of $PM_{2.5}$ by 0.55-1 $\mu g/m^3$. Compared to $PM_{2.5}$, PM_{10} concentrations are larger and more volatile by definition, as these include larger and less harmful particles; therefore, for the same decrease in boundary layer height, it is plausible to expect larger variations in PM_{10} concentrations than those caused by the implementation of Low Emission Zones.

Columns 2 and 3 show a similar effect on school grades, but the effects are not significant (marginally insignificant for German). This is probably due to the fact that, differently from standardized test scores or school exams, school grades are not a very accurate measure of school performance and are highly influenced by several contextual factors, including the type of school and the type of track where the student is enrolled, which are not fully absorbed by the fixed effects in our specification.

Table 2: In-utero exposure to air pollution and educational achievement

	(1)	(2)	(3)
	School track	School grades: age 12-14	
Dep. Var.	Gymnasium	German	Maths
BLH in-utero	0.0371**	0.0703	0.0630
	(.0179)	(.05)	(.0594)
Observations	5902	3167	3164
Mean Dep. Var.	0.43	4.36	4.33
Counties	373	346	346

The table shows the effects of BLH in-utero (in standard deviations) on the likelihood of transition to *gymnasium*, the academic track, and on school grades at age 12-14. All regressions include county, month and year of birth fixed effects, parental and child characteristics and period-specific weather controls as described in equation 1a-1b. SE are clustered at the county level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

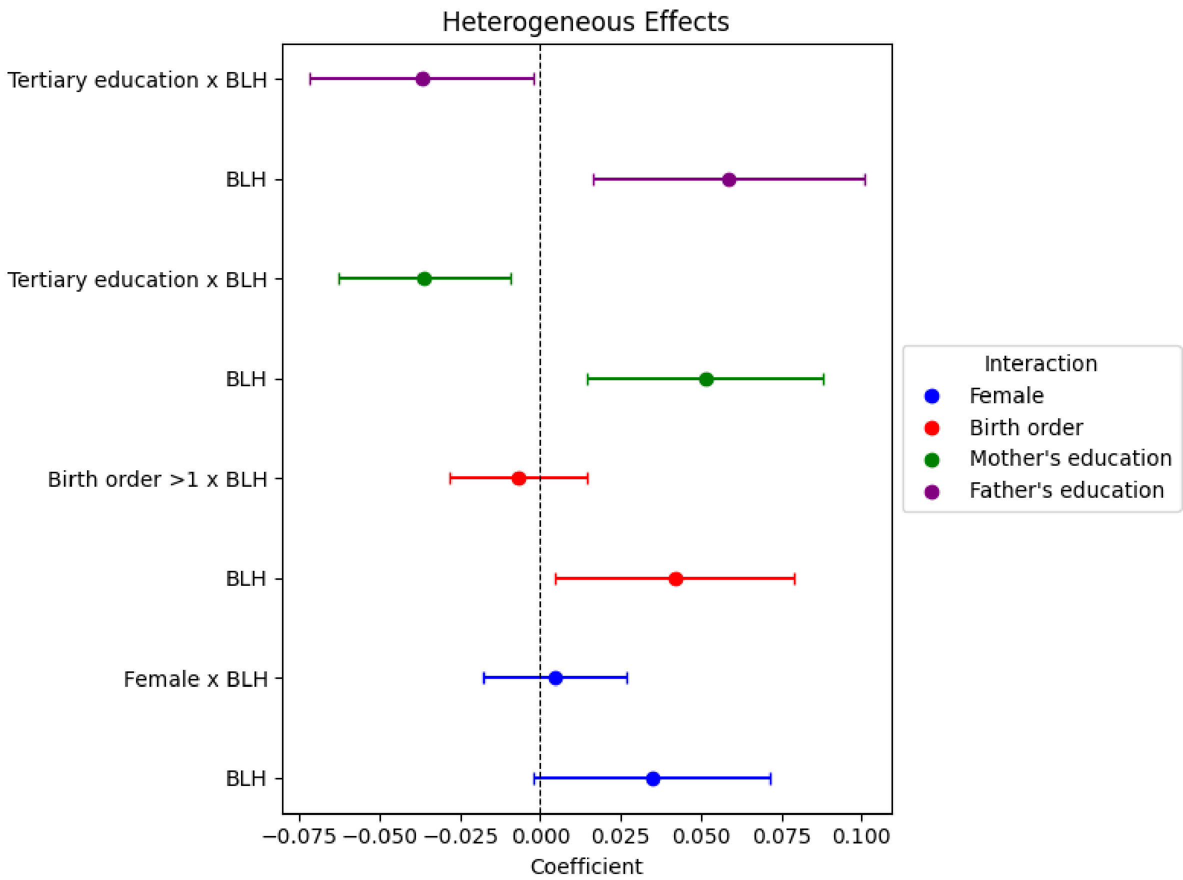


Figure 1: Effect heterogeneity by demographics and SES

The figure shows estimated coefficients of reduced form regressions of educational achievement (*gymnasium*) on BLH during the in-utero period and interactions between BLH and sociodemographic characteristics of child and parents. Each color represents a different regression. The coefficients of BLH are the baseline coefficients. All regressions include county, month and year of birth fixed effects, parental and child characteristics and weather controls as described in equation 1a.

5.2 Effect heterogeneity

After having determined that higher prenatal pollution exposure significantly decreases the rate of transition to the academic track, we assess the heterogeneity of the effects with respect to various socio-demographic dimensions. We run the reduced form model described in equation 1a, including interaction terms with the dimension of interest, namely gender, order of birth (in this case a binary variable for whether the child is the first born), maternal education and paternal education.

Figure 1 plots the coefficients of interest from these regressions (Table 7 shows the point estimates of these regressions). Considering the demographic characteristics of the child, we do not find heterogeneous effects with respect to child gender or to birth order, with the latter implying that first born children are not differentially affected from children born in higher parities. However, considering parental education, we find that the negative effect of a decrease in BLH on the likelihood of enrolling to the academic track is much smaller in absolute terms for children with highly educated parents (around 50% of the baseline effect is absorbed by parental education) as compared to children from lower educated families, with the reduction being particularly significant for high maternal education. The presence of a negative SES gradient is in line with previous literature and suggests that higher parental education strongly mitigates the adverse effects of prenatal pollution exposure (Voorheis et al., 2017; Bernardi and Keivabu, 2024; Thies, 2024).

Since the level of boundary layer height is not systematically correlated with socioeconomic and demographic characteristics of the household, including parental education, this effect can be mainly attributed to the possibility of high SES households to remediate the effect of pollution exposure during the child's development rather than to the possibility to enjoy a better air quality in the first place. In other words, a higher parental education likely entails better financial means invested in the child's development and a higher awareness of the effects and the implications of exposure to air pollution.

In our setting, the large mitigating effect of high parental education could also be partially attributed to a higher relative importance of the final parental decision regarding the secondary school track to which the child is allocated, as compared to households from lower education backgrounds, relative to other factors such as teacher recommendation and child's performance in primary school. While this hypothesis cannot be tested directly in this setting, several studies in the German context suggest that the choice of secondary school track and the subsequent educational achievements of the child are strongly correlated with parental education (Waldinger, 2007; Dodin et al., 2024).

5.3 The role of parental activities

The mitigating effect of parental education suggests that higher educated parents can remediate the adverse effects of prenatal pollution exposure more effectively than lower educated parents. One of the mechanisms explaining this socioeconomic gradient could be the differential parental response to the child's pollution exposure in terms of activities during early childhood. In other words, higher educated parents could increase the frequency of "beneficial" activities during early childhood in response to higher prenatal pollution exposure, thus mitigating its adverse effect.

For the purpose of this study, we focus on the activities described in Section 3. Table 8 illustrates the baseline effect of the frequency of each activity on the likelihood of transition to Gymnasium: singing and reading frequently significantly increase the likelihood of transition to the academic track; conversely, frequently watching tv has negative effects; finally, frequent painting and crafting activities with the child do not significantly affect the educational trajectory of the child.

Having determined how each activity is correlated with the educational achievement of the child, we assess whether these causally mediate the differential effect of pollution by parental SES. Table 9 shows that exposure to air pollution does not affect the frequency of any of the activities of interest, regardless of the mother's education level, thus ruling out the mediation hypothesis. Therefore, parents do not seem to adjust the frequency of these activities in response to variations in pollution exposure.

The absence of evidence on parental response to prenatal pollution is consistent with a setting like the German one, where pollution concentrations are relatively low and rarely give rise to major public health concerns. However, it may be the case that parents with a higher education, carry out beneficial activities more frequently than lower educated parents, and harmful or numbing activities less frequently, independent of their children's exposure to pollution. Table 10 in the Appendix shows how parental education is correlated with the frequency of the activities: compared to lower educated parents, parents with tertiary education tend to read and to sing songs with the child more frequently on one hand, and to paint and do crafts and watch tv less frequently on the other. To understand how these parental activities moderate the effect of prenatal exposure to pollution, we assess effect heterogeneity with respect to the frequency (high vs low) of each activity of interest.

Table 3 shows that three out of four parental activities significantly moderate the effects of pollution exposure, although in different directions, as shown by the interaction coefficients. To better interpret these coefficients, we need to take into account the baseline effects of the activities; the positive coefficient on the interaction for singing (column 1) indicates that the benefits of singing increase with better air quality and decrease with higher pollution; more precisely, the effect of singing at the mean level of BLH is positive, in line with the baseline effect, but for sufficiently low levels of BLH below the mean (hence for sufficiently high levels of pollution), the effect of singing becomes negative. The interpretation changes when considering watching tv, even though the coefficients in column 3 show the same dynamics of those in column 1: in this case, the positive sign of the interaction coefficient means that the detrimental effects of watching tv decrease as air quality increases and increase with higher pollution; however, differently from the case of singing, the effect of watching tv is negative at almost all levels of BLH and reaches zero only at the highest levels of BLH, meaning that only very low levels of pollution completely counterbalance the detrimental effects of watching tv. Column 4 shows the opposite pattern for reading: the beneficial effects of reading are particularly relevant at high levels of pollution, becoming less relevant as air quality increases; in this case, the positive effects of reading converge to zero for the highest values of BLH. Finally, for painting (column 2), the coefficient on the interaction is not significant, in line with the small and non-significant baseline effect.

Overall, this suggests that watching tv reinforces and reading compensates the negative effect of exposure to pollution on educational achievement. Conversely, for singing, not only the size, but also the direction of the effect changes depending on the level of air pollution: at high (low) levels of air pollution, singing further decreases (increases) the likelihood of enrolling to the academic track; that is, singing amplifies the effect of pollution exposure.

Following the literature on parental activities and in particular the distinction defined in Del Bono et al. (2016), we can characterize singing, painting and doing crafts and watching tv and videos as recreational activities, whereas reading with the child can be ascribed to the category of educational or cognitively stimulating activities. According to this categorization, the results show that recreational activities tend to reinforce the effect of exposure to air pollution, whereas educational activities tend to mitigate the negative effect of exposure. Importantly, there is no evidence in our data that parents consider some of these activities as substitutes, therefore this effect does not seem to be driven by selective parental behaviours, but by the activities themselves.

It is now relevant to assess whether the heterogeneous effects with respect to parental activities exhibit a differential pattern by parental education. Table 4 reports the results of the effect heterogeneity with respect to parental activities decomposing by maternal education (tertiary degree vs less). Overall, the effects seem to be concentrated among children from lower educated households: in other words, the reinforcing effect of recreational activities and the mitigating effect of educational activities are particularly relevant for lower SES children, whereas these activities seem to matter less for children in highly educated households. This is in line with the presence of a negative SES gradient illustrated in Figure 1, which remains even after controlling for parental activities. The coefficients in the subsample of children with university-educated mothers, instead, are not significant, likely due to data and sample size limitations, and should just be considered as suggestive: for singing and reading (columns 2 and 8), the interaction coefficients are very similar to those in the subsample for mothers without a tertiary degree (columns 1 and 7) while for watching tv the sign and the size of the interaction coefficient differs across mother's education status, suggesting that watching tv does not really play a relevant role in high SES households.

5.4 Robustness checks

Table 2 shows that a decrease in boundary layer height and a consequential increase in pollution concentration in-utero decreases the likelihood of transitioning to the academic track in secondary school. In the Appendix, we test the robustness of this result to alternative samples (Table 11) and to alternative specifications (Table 12).

In our main specification, we do not impose a minimum number of observations per county. However, in this fixed effects model, counties with very few observations and potentially high

Table 3: Heterogeneous Effects by Parental Activities

Gymnasium	(1)	(2)	(3)	(4)
BLH in-utero	.044	.0542*	.0363	0.0889***
	(.0326)	(.0328)	(.0364)	(.0349)
Singing (high freq.)	-0.376**			
	(.167)			
Singing (high freq.) x BLH	0.0478***			
	(.0177)			
Painting (high freq.)		-.111		
		(.171)		
Painting (high freq.) x BLH		.0142		
		(.0180)		
Watching tv (high freq.)			-0.397**	
			(.183)	
Watching tv (high freq.) x BLH			0.039**	
			(.0196)	
Reading (high freq.)				0.463***
				(.175)
Reading (high freq.) x BLH				-0.043**
				(.0184)
Observations	2534	2534	2535	2537

The table shows the effects of BLH in-utero (in standard deviations) on the likelihood of transition to *gymnasium* interacted with each parental activity of interest. All regressions include county, month and year of birth fixed effects, parental and child characteristics and period-specific weather controls as described in equation 1a. SE clustered at county level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 4: Heterogeneous Effects by Parental Activities and Mother's Education

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mother's tertiary degree	No	Yes	No	Yes	No	Yes	No	Yes
Dep.Var.				Gymnasium				
BLH in utero	.0563	-.0095	.0683	.0015	.0359	-.0032	0.095**	.0123
	(.0429)	(.0730)	(.0454)	(.0621)	(.0493)	(.0587)	(.0448)	(.0724)
Singing (high freq.)	-.301	-.341						
	(.203)	(.393)						
Singing (high freq.) x BLH	.0407*	.0408						
	(.0214)	(.0424)						
Painting (high freq.)			-.119	.103				
			(.217)	(.359)				
Painting (high freq.) x BLH			.014	-.0135				
			(.0227)	(.0378)				
Watching tv (high freq.)					-.553**	.085		
					(.267)	(.337)		
Watching tv (high freq.) x BLH					.0545*	-.0095		
					(.0284)	(.0366)		
Reading (high freq.)							0.411*	.273
							(.21)	(.424)
Reading (high freq.) x BLH							-0.037*	-.0218
							(.0222)	(.0449)
Observations	1740	703	1737	706	1739	705	1738	707

The table shows the effects of BLH in-utero (in standard deviations) on the likelihood of transition to *gymnasium* interacted with each parental activity of interest, decomposing the sample by maternal education (tertiary vs less). All regressions include county, month and year of birth fixed effects, parental and child characteristics and period-specific weather controls as described in equation 1a. SE clustered at county level. *** p<0.01, ** p<0.05, * p<0.1.

variance in the dependent and independent variables could drive our result. To reduce the influence of outliers, in line with previous studies, in columns 1 and 2 of Table 11 we restrict the sample to counties with at least 10 or 20 observations, respectively. The coefficients are larger and more significant than in the baseline model, as expected. In column 3 we exclude the survey year 2020, as the outburst of the pandemics may have heavily affected trends in the outcome: the coefficient is slightly lower but remains significant.

Column 1 of Table 12 replicates the reduced form model in equation 1a excluding background household and child characteristics, to check whether the main result is driven by the inclusion of the set of socioeconomic and demographic controls. In column 2, the coefficient is virtually unchanged when excluding the quadratic terms from the set of weather controls. In column 3 we add one lead and one lag (average value of the nine months before pregnancy and of the nine months after birth) of the instrument and of the weather variables to the main specification. Given the temporal auto-correlation of the weather variables and of the instrument, omitting atmospheric variables from before pregnancy and after birth could result in their effect loading onto the coefficient of interest. The coefficient on the lead and of the lag is not significant, but remains significant for the in-utero period. This is further confirmed in column 4, where we regress the dependent variable only on the lag and its correspondent weather controls. The absence of any effect, both in terms of significance and size, in the period before pregnancy confirms that the effect of pollution exposure on educational achievement operates only through the effects of pollutants on the fetus' neurocognitive and physical development. Finally, in column 5 we add state-by-year fixed effects in the main model to address the possibility that the result is driven by regional trends in the outcome, which could be related to reforms or changes of the school system at the state level during the period of interest. Also in this case, the coefficient is significant and similar in magnitude to the coefficient in the main specification.

6 Conclusion

This paper examines the effect of prenatal exposure to air pollution on educational achievement in Germany, assessing whether this effect changes depending on the socioeconomic background of the household and testing how specific parental activities carried out during early childhood moderate the impact of pollution. Educational achievement is measured by school grades at ages 12-14 and by the rate of transition to the academic track (*Gymnasium*) in secondary school. Due to the German early and rigid tracking system, characterized by a low mobility across tracks, the chosen track for secondary school has relevant implications for the academic and professional trajectory of the child (Dustmann, 2004).

To address potential endogeneity in the exposure to air pollution, we leverage exogenous variations in the height of the planetary boundary layer (BLH), an atmospheric variable strongly affecting pollution concentrations on the ground.

We find that an increase in pollution exposure during gestation significantly decreases the likelihood of transitioning to the academic track, with the effect being concentrated among children from less educated families. This socioeconomic gradient cannot be explained by a differential parental response to pollution exposure in terms of parental activities. However, parental activities moderate the impact of pollution in opposite directions, depending on the nature of the activity. While an educational activity like reading can effectively mitigate the adverse impact of pollution, recreational activities mostly reinforce the baseline effects of pollution exposure: in particular, a passive and not engaging activity such as watching tv, tends to aggravate the adverse effect of pollution. Considering that parents with tertiary education tend to read stories more frequently and watch tv less frequently than lower educated parents, these findings suggest that parental activities are one of the reasons underlying the socioeconomic gradient in the effect of prenatal pollution exposure on educational achievement. Finally, when we further decompose the analysis by household SES, the moderating effect of parental activities seems less relevant in high-SES households.

These results offer new insights for the evaluation of the long-term effects of prenatal pollution exposure. On the one hand, a high socioeconomic status, as measured by parental education, significantly remediates the adverse effects of pollution on educational achievement. This leaves the opportunity for socially equitable policies, such as targeted subsidies or early-life programs for children of lower educated households, to limit the vulnerability of disadvantaged children and to reduce the socioeconomic gap with higher educated households.

On the other hand, parental activities that educate and cognitively stimulate the child, can also effectively remediate the impacts of pollution exposure, particularly for lower educated households. This highlights the importance of the quality of parental involvement during the child's early development, which is not unequivocally determined by parental education and socioeconomic status.

In this regard, disentangling the relative roles of parental involvement and parental SES is an important avenue for future research on the long-term impacts of adverse prenatal exposures.

References

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.

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.

Barker, David J (1990). “The fetal and infant origins of adult disease.” In: *BMJ: British Medical Journal* 301.6761, p. 1111.

Bernardi, Fabrizio and Risto Conte Keivabu (2024). “Poor air quality at school and educational inequality by family socioeconomic status in Italy”. In: *Research in Social Stratification and Mobility* 91, p. 100932.

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.

Black, Sandra E et al. (2019). “This is only a test? Long-run and intergenerational impacts of prenatal exposure to radioactive fallout”. In: *Review of Economics and Statistics* 101.3, pp. 531–546.

Brehm, Johannes et al. (2022). *From low emission zone to academic track: Environmental policy effects on educational achievement in elementary school.* 980. Ruhr Economic Papers.

Chay, Kenneth Y and Michael Greenstone (Oct. 2003a). *Air Quality, Infant Mortality, and the Clean Air Act of 1970.* Working Paper 10053. National Bureau of Economic Research. DOI: 10.3386/w10053.

— (2005). “Does air quality matter? Evidence from the housing market”. In: *Journal of political Economy* 113.2, pp. 376–424.

— (Aug. 2003b). “The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession”. In: *The Quarterly Journal of Economics* 118.3, pp. 1121–1167. ISSN: 0033-5533. DOI: 10.1162/00335530360698513.

Chen, Hui et al. (2021). “Particulate matter, an intrauterine toxin affecting foetal development and beyond”. In: *Antioxidants* 10.5, p. 732.

Colmer, Jonathan et al. (2021). “Why are pollution damages lower in developed countries? Insights from high-Income, high-particulate matter Hong Kong”. In: *Journal of Health Economics* 79, p. 102511.

Coneus, Katja and C Katharina Spiess (2012). “Pollution exposure and child health: Evidence for infants and toddlers in Germany”. In: *Journal of Health Economics* 31.1, pp. 180–196.

Conte Keivabu, Risto and Tobias Rüttenauer (2022). “London congestion charge: the impact on air pollution and school attendance by socioeconomic status”. In: *Population and Environment* 43.4, pp. 576–596.

Currie, Janet, Matthew Neidell, and Johannes F Schmieder (2009). “Air pollution and infant health: Lessons from New Jersey”. In: *Journal of health economics* 28.3, pp. 688–703.

Del Bono, Emilia et al. (2016). “Early maternal time investment and early child outcomes”. In: *The Economic Journal* 126.596, F96–F135.

Deschenes, Olivier, Michael Greenstone, and Joseph S Shapiro (2017). “Defensive investments and the demand for air quality: Evidence from the NOx budget program”. In: *American Economic Review* 107.10, pp. 2958–2989.

Dodin, Majed et al. (2024). “Social mobility in Germany”. In: *Journal of Public Economics* 232.105074.

Dustmann, Christian (2004). “Parental background, secondary school track choice, and wages”. In: *Oxford Economic Papers* 56.2, pp. 209–230.

Dustmann, Christian, Patrick A Puhani, and Uta Schönberg (2017). “The long-term effects of early track choice”. In: *The Economic Journal* 127.603, pp. 1348–1380.

Fan, Wei and Catherine Porter (2020). “Reinforcement or compensation? Parental responses to children’s revealed human capital levels”. In: *Journal of Population Economics* 33.1, pp. 233–270.

Godzinski, Alexandre and Milena Suarez Castillo (2021). “Disentangling the effects of air pollutants with many instruments”. In: *Journal of Environmental Economics and Management* 109, p. 102489.

Goebel, Jan et al. (2019). “The German socio-economic panel (SOEP)”. In: *Jahrbücher für Nationalökonomie und Statistik* 239.2, pp. 345–360.

Gorman, Des et al. (2003). “The clinical toxicology of carbon monoxide”. In: *Toxicology* 187.1, pp. 25–38.

Grönqvist, Hans, J Peter Nilsson, and Per-Olof Robling (2020). “Understanding how low levels of early lead exposure affect children’s life trajectories”. In: *Journal of Political Economy* 128.9, pp. 3376–3433.

Guxens, Mònica et al. (2018). “Air pollution exposure during fetal life, brain morphology, and cognitive function in school-age children”. In: *Biological psychiatry* 84.4, pp. 295–303.

Hsin, Amy (2012). “Is biology destiny? Birth weight and differential parental treatment”. In: *Demography* 49.4, pp. 1385–1405.

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.

Jacobson, Mark Z (2002). *Atmospheric pollution: history, science, and regulation.* Cambridge University Press.

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.

Levi, Yoav et al. (2020). "On the association between characteristics of the atmospheric boundary layer and air pollution concentrations". In: *Atmospheric Research* 231, p. 104675.

Levy, Richard J (2015). "Carbon monoxide pollution and neurodevelopment: a public health concern". In: *Neurotoxicology and teratology* 49, pp. 31–40.

Marcotte, Dave E (2017). "Something in the air? Air quality and children's educational outcomes". In: *Economics of Education Review* 56, pp. 141–151.

Molina, Teresa (2021). "Pollution, ability, and gender-specific investment responses to shocks". In: *Journal of the European Economic Association* 19.1, pp. 580–619.

Moretti, Enrico and Matthew Neidell (2011). "Pollution, health, and avoidance behavior: evidence from the ports of Los Angeles". In: *Journal of human Resources* 46.1, pp. 154–175.

Müller, Sophie and Thorsten Schneider (2013). "Educational pathways and dropout from higher education in Germany". In: *Longitudinal and Life Course Studies* 4.3, pp. 218–241.

Neidell, Matthew (2009). "Information, avoidance behavior, and health: the effect of ozone on asthma hospitalizations". In: *Journal of Human resources* 44.2, pp. 450–478.

Prado Bert, Paula de et al. (2018). "The effects of air pollution on the brain: a review of studies interfacing environmental epidemiology and neuroimaging". In: *Current environmental health reports* 5, pp. 351–364.

Restrepo, Brandon J (2016). "Parental investment responses to a low birth weight outcome: who compensates and who reinforces?" In: *Journal of Population Economics* 29.4, pp. 969–989.

Sander, Nikola (2017). "Germany: Internal migration within a changing nation". In: *Internal Migration in the Developed World*. Routledge, pp. 226–241.

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.

Schwartz, Joel, Marie-Abele Bind, and Petros Koutrakis (2017). "Estimating causal effects of local air pollution on daily deaths: effect of low levels". In: *Environmental health perspectives* 125.1, pp. 23–29.

Shen, Siyuan et al. (2024). "Enhancing Global Estimation of Fine Particulate Matter Concentrations by Including Geophysical a Priori Information in Deep Learning". In: *ACS ES&T Air* 1.5, pp. 332–345.

Smithline, Howard A et al. (2003). "Whole body oxygen consumption and critical oxygen delivery in response to prolonged and severe carbon monoxide poisoning". In: *Resuscitation* 56.1, pp. 97–104.

Stawarz, Nico and Nikola Sander (2019). "The impact of internal migration on the spatial distribution of population in Germany over the period 1991-2017". In: *Comparative Population Studies* 44.

Sullivan, Daniel M (2016). "The true cost of air pollution: Evidence from house prices and migration". In: *Harvard University*.

Thies, Beate (2024). "Prenatal Exposure to Air Pollution and the Development of Socio-Emotional Skills". Working Paper.

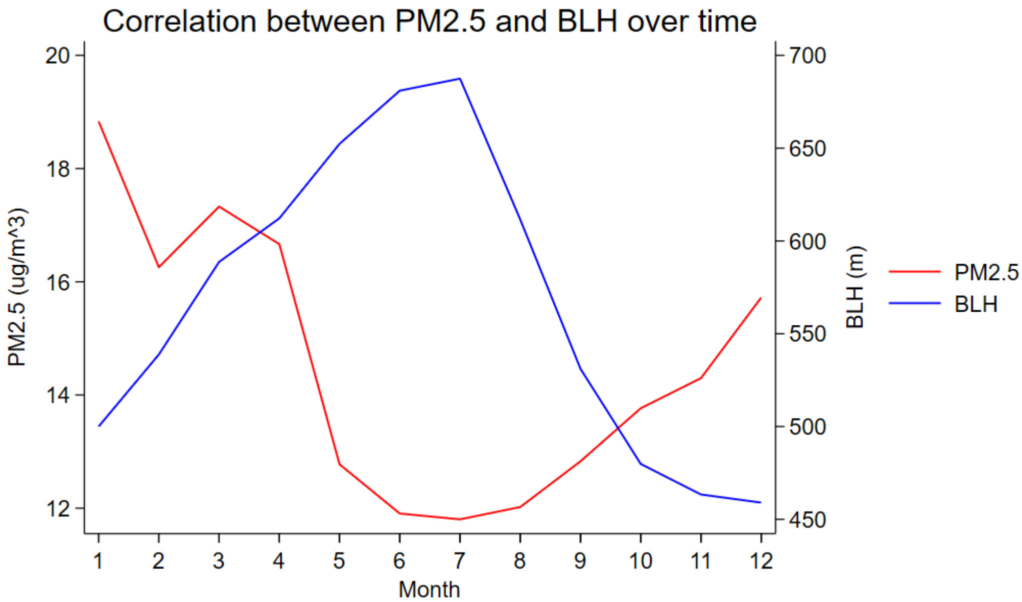
Voorheis, John et al. (2017). "Air quality, human capital formation and the long-term effects of environmental inequality at birth". In: *US Census Center for Economic Studies, CARRA Working Paper Series* 5.

Waldinger, Fabian (2007). "Does Ability Tracking Exacerbate the Role of Family Background for Student's Test Scores?" Working Paper.

Zhang, Xin et al. (2024). *Fetal Pollution Exposure, Cognitive Ability, and Gender-Specific Parental Investment*. eng. IZA Discussion Papers 17288. Bonn. URL: `https://hdl.handle.net/10419/305730`.

Appendix

Figure 2: Boundary Layer Height and PM2.5



The figure shows the correlation between the monthly values of BLH and PM2.5 averaged over the years 2000-2010 for our sample.

Table 5: Effect of BLH on PM2.5

PM2.5	(1)	(2)
Sample	Gymnasium	German
BLH in-utero	-0.552***	-0.995***
	(.0466)	(.0796)
Mean Dep. Var.	14.88	14.96
Counties	373	346
Observations	5902	3167

The table shows the effects of BLH in-utero (in standard deviations) on the concentrations of PM2.5 for the *gymnasium* sample and for the school grades sample. Both regressions include county, month and year of birth fixed effects and period-specific weather controls as described in equation 1a-1b. SE are clustered at the county level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 6: Falsification checks

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep.Var	MAB	Migration Background	Mother: No HS	Mother: HS	Mother: Tertiary	Father: No HS	Father: HS	Father: Tertiary
				Panel A: Gymnasium Sample				
BLH in-utero	-0.287	-0.011	-0.014	0.021	-0.008	-0.015	-0.027	0.042*
	(.224)	(.018)	(.015)	(.02)	(.018)	(.014)	(.026)	(.025)
Observations	5828	5828	5812	5812	5812	4709	4709	4709
				Panel B: School Grades Sample				
BLH in-utero	-0.293	0.015	-0.016	0.044	-0.028	-0.014	-0.023	0.037
	(.319)	(.028)	(.022)	(.029)	(.027)	(.022)	(.039)	(.038)
Observations	5574	5574	5410	5410	5410	4258	4258	4258

The table shows the effects of BLH in-utero (in standard deviations) on a set of household and child characteristics for the two samples of analysis. MAB (Mother's age at birth) is measured in years. All regressions include county, month and year of birth fixed effects, child gender and period-specific weather controls as described in equation 1a-1b. SE are clustered at the county level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 7: Effect heterogeneity by SES

Gymnasium	(1)	(2)	(3)	(4)
BLH in-utero	0.0348*	0.0418**	0.0513***	0.0584***
	(.0187)	(.0189)	(.0187)	(.0215)
Female x BLH	0.0044			
	(.0114)			
Birth order >1 x BLH		-0.007		
		(.0108)		
Mother: Tertiary x BLH			-0.0361***	
			(.0137)	
Father: Tertiary x BLH				-0.0369**
				(.0177)
Mean Dep. Var.	0.43	0.43	0.43	0.46
Counties	373	373	373	357
Observations	5902	5902	5885	4776

The table shows the effects of BLH in-utero (in standard deviations) on the likelihood of transition to *gymnasium* interacted with gender, birth order (higher parity vs first) maternal and paternal education (tertiary vs less). All regressions include county, month and year of birth fixed effects, parental and child characteristics and period-specific weather controls as described in equation 1a. SE are clustered at the county level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 8: Parental activities and Gymnasium

	Gymnasium
Singing (high freq.)	0.0531**
	(.0226)
Painting (high freq.)	-0.0002
	(.0208)
Watching tv (high freq.)	-0.0412**
	(.0207)
Reading (high freq.)	0.0559***
	(.0216)
Mean Dep. Var.	0.43
Observations	2524

The table shows the baseline effects of the binary parental activities (high vs low frequency) on the likelihood of transition to *gymnasium* from a regression including county, month and year of birth fixed effects, parental and child characteristics as described in equation 1a, and excluding BLH and weather controls. SE are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Parental activities and BLH

	(1)	(2)	(3)	(4)
High Frequency of:	Singing	Painting & Crafts	Watching TV	Reading
BLH in-utero	-0.0204	0.0588	-0.0392	0.0114
	(.0383)	(.0357)	(.0401)	(.038)
Mother: Tertiary	-0.0986	0.027	0.2108	0.0767
	(.2612)	(.2227)	(.2376)	(.2181)
Tertiary x BLH	0.0152	-0.0085	-0.0293	0.0061
	(.0276)	(.0245)	(.0255)	(.0234)
Mean Dep. Var.	0.58	0.53	0.56	0.61
Counties	299	300	299	300
Observations	2526	2526	2527	2529

The table shows the effects of BLH in-utero (in standard deviations) on the frequency of each parental activity (1 for high frequency and 0 for low frequency), by maternal education status (tertiary vs less). All regressions include county, month and year of birth fixed effects, parental and child characteristics and period-specific weather controls as described in equation 1a. SE are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table 10: Parental education and activity frequency

	(1)	(2)	(3)	(4)
High Frequency of:	Singing	Painting & Crafts	Watching TV	Reading
Mother: Tertiary	0.03	-0.074***	-0.048	0.131***
	(.027)	(.028)	(.03)	(.026)
Father: Tertiary	0.061**	-0.009	-0.131***	0.063**
	(.027)	(.027)	(.033)	(.028)
Observations	2524	2524	2524	2524

The table shows the correlation of paternal education (in binary form, tertiary education vs less) with the frequency of each parental activiy. All regressions include county, month and year of birth fixed effects, parental and child characteristics as described in equation 1a, and the parental activities other than the outcome variable. SE are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Robustness: Alternative Samples

Gymnasium	(1)	(2)	(3)
BLH in-utero	0.0469**	0.0497**	0.0301*
	(.0184)	(.0199)	(.0181)
Minimum Obs. per county	10	20	All
Years	All	All	No 2020
Mean Dep. Var.	0.43	0.44	0.42
Counties	245	140	373
Observations	5305	4143	5453

The table shows the effects of BLH in-utero (in standard deviations) on the likelihood of transition to *gymnasium*, using different samples from the baseline sample of eqiation 1a. Columns 1 and 2 restrict the sample to counties with at least 10 and 20 observations, respectively. Column 3 excludes survey year 2020 from the analysis. All regressions include county, month and year of birth fixed effects, parental and child characteristics and period-specific weather controls as described in equation 1a. SE are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.

Table 12: Robustness: Alternative Specifications

Gymnasium	(1)	(2)	(3)	(4)	(5)
BLH in-utero	0.0444**	0.0373**	0.0471**		0.0389*
	(.0206)	(.0175)	(.0236)		(.0213)
BLH prenatal			-0.0027	-0.0079	
			(.0209)	(.0182)	
BLH postnatal			-0.0198		
			(.023)		
Mean Dep. Var.	0.43	0.43	0.43	0.43	0.43
Counties	373	373	368	369	373
Observations	5902	5902	5542	5642	5895
Specification	No SES controls	Linear weather	Ext. Weather controls	1 Lag BLH	Year*region FE

The table shows the effects of BLH in-utero (in standard deviations) on the likelihood of transition to *gymnasium*, using different specifications from the baseline regression in equation 1a. In column 1 we do not include the set of child and household background characteristics X. In column 2 we do not include weather controls in quadratic form. In column 3 we add one lead and one lag (nine months average) of the instrument and of second order weather controls. In column 4 we run the same regression as column 3 but only keeping the lag of the instrument and the correspondent weather controls. In column 5 we expand the set of fixed effects by including year of birth by state of birth fixed effects. SE are clustered at the county level. *** p<0.01, ** p<0.05, * p<0.1.