Prenatal Exposure to Air Pollution and the Development of Socio-Emotional Skills

Beate Thies*

September 2024

Abstract

Socio-emotional skills are important predictors for life outcomes like education, health and earnings. This paper provides causal evidence on the effect of in-utero exposure to air pollution on socio-emotional ability in childhood. Using thermal inversions to address endogeneity in pollution exposure and data from a representative household survey in Germany, I find that an increase in fine particulate matter concentration by 1 $\mu g/m^3$ during the prenatal period increases neuroticism and internalizing behavior at age 5-10 by 13% and 18% of a standard deviation, respectively. This implies that affected children are less emotionally stable and suggests adverse impacts on mental health. Back of the envelope computations indicate that a standard deviation increase in fine particulate matter reduces adult earnings by 0.23%-0.74% through its impact on socio-emotional ability.

JEL Codes: Q53, J24, J13

Keywords: air pollution, human capital, child development, non-cognitive skills, mental health

*Department of Economics, University of Vienna. Correspondence: beate.thies@univie.ac.at. I thank Omar Bamieh, Felix Holub, Vera Huwe, Dana Kassem, Alessandro Palma and Ulrich Wagner for helpful feedback and suggestions. I am also grateful to seminar participants at the University of Mannheim, the EAERE conference, the CRC TR 224 Summer School on the Development of Cognitive and Non-Cognitive Skills in Childhood and Adolescence, and the AURÖ Young Researchers' Workshop for their comments.

1. Introduction

Air pollution adversely affects several dimensions of human health and well-being, imposing high costs on society and hindering social mobility. Exposure to air pollution during the year of birth reduces educational attainment as well as adult earnings and employment (Isen et al., 2017; Voorheis, 2017). Colmer and Voorheis (2021) show that the adverse impacts of prenatal exposure even extend to the next generation. Since low-income families and minorities often live in more polluted neighbourhoods than more affluent groups (see e.g. Banzhaf et al., 2019; Currie et al., 2023; Rüttenauer, 2018), these long-run effects of air pollution not only impose a substantial economic cost on society, but also inhibit equality of opportunity.

Optimal policy responses to this issue might depend on the mechanisms driving the adverse long-run effects of early pollution exposure. As education and labor market outcomes are functions of human capital, its core components – cognitive and socio-emotional skills[1] – are potential mediators. The existing evidence regarding these channels is incomplete. While prenatal pollution exposure has been shown to reduce scores in school math and language tests (Bharadwaj et al., 2017; Sanders, 2012), as well as tests of fluid intelligence (Molina, 2021), i.e. measures of cognitive skills, evidence regarding its effect on socio-emotional abilities is missing. Therefore, the aim of this paper is to investigate whether in-utero exposure to air pollution affects socio-emotional skills, and to assess how important this potential channel is relative to the cognitive ability mechanism.

I employ data on socio-emotional abilities during childhood from the German Socio-Economic Panel (Goebel et al., 2019). The survey includes mother-reported Big Five personality traits and child behavior from the Strength and Difficulties Questionnaire (SDQ) for children aged 5-10. I combine these outcomes for a sample of 8,250 children born between 2000 and 2014 with data on satellite-derived particulate matter with a diameter of less than 2.5 µm (PM2.5), air pollution readings from outdoor monitors, and reanalysis data on meteorological conditions. To address endogeneity and measurement error in particulate matter exposure, I exploit exogenous variation in thermal inversions in an instrumental variable (IV) approach (following e.g. Arceo et al., 2016; Molina,

[1] While cognitive ability captures intelligence, socio-emotional skills - also known as non-cognitive skills or personality traits - comprise a variety of abilities that are weakly correlated with intelligence, such as social competencies, emotional stability and persistence.

2021). An inversion is a meteorological phenomenon during which air temperature increases with altitude. The warm upper air layer acts like a ceiling that prevents ground-level emissions from dispersing, causing an increase in surface-level pollution concentrations.

Results indicate that prenatal pollution exposure reduces emotional stability; a 1 unit increase in gestational PM2.5 exposure raises the Big Five trait neuroticism by 13% of a standard deviation. In line with this, it also increases internalizing behavior, which is based on the SDQ and relates to adverse emotional symptoms and problems in the interaction with peers, by 18% of a standard deviation. I do not find evidence for effects of prenatal air pollution exposure on other dimensions of the Big Five or externalizing behavior, which captures hyperactivity and conduct problems. The effects on neuroticism and internalizing behavior remain quantitatively similar when including sibling fixed effects. The effect sizes are of the same order of magnitude as the impact of air pollution on cognitive ability found in earlier work. Since previous research established negative effects of both neuroticism and internalizing behavior on labor market outcomes, they are plausible channels underlying the adverse long-run effects of early-life pollution exposure. Back-of-the envelope computations imply that an increase of PM2.5 by one standard deviation reduces earnings by roughly 0.23 – 0.74% via the deterioration in socio-emotional skills. The effects on both outcomes are mainly driven by increases in the upper tail of the distributions. This is especially concerning, as high levels of neuroticism and internalizing problems have been linked to mental health issues. Among children in high-income households, effects are slightly attenuated, suggesting potential for remediating investments into socio-emotional skills.

Understanding which mechanisms drive adverse long-term effects of early-life pollution exposure, and how important the respective channels are in quantitative terms is paramount when deciding about feasible and appropriate policy responses. While the predictive power of cognitive and socio-emotional skills for educational attainment and labour market performance is comparable, they differ crucially in how they respond to intervention programs and investments: There is growing evidence that socio-emotional skills are malleable up until adulthood and can be improved by way of low-cost interventions implemented in the classroom- or even work-environment (Adhvaryu et al., 2023; Alan et al., 2019; Sorrenti et al., 2024), whereas cognition is less malleable, especially after school start age (e.g. Almlund et al., 2011; Cunha et al., 2010). Hence, if long-term effects were driven purely by reduced cognitive skills, the only option to avoid them in future cohorts would be to reduce air pollution. If, on the other hand, socio-emotional abilities play a relevant role as well, as indicated by my results, long-term effects can also be alleviated *ex-post* via intervention programs or investments targeting these abilities. Given that it can be extremely costly or even impossible to reduce pollution levels in some circumstances (e.g. pollution arising from another jurisdiction or from natural sources like desert dust), and that it would only benefit future cohorts, investigating whether alternative options exist to compensate individuals ex-post for exposure to high levels of pollution while in utero, is highly policy-relevant.

This study contributes to the literature on the long-run consequences of early-life exposure to air pollution by providing causal evidence for a new relevant outcome, namely socio-emotional abilities. This complements work studying long-term impacts on cognitive skills (Sanders, 2012; Bharadwaj et al., 2017; Molina, 2021) and physical health (Klauber et al., 2024; Ferro et al., 2024). Moreover, using data on cohorts born in Germany after 2000, I study a setting with relatively low baseline pollution levels.[2] I thus shed light on a different part of the dose-response function than existing papers on the effects of gestational pollution exposure on human capital which are based on data from developing countries (Bharadwaj et al., 2017; Molina, 2021; Rosales-Rueda and Triyana, 2019) or cohorts born in the US during the 1970s-1980s (Isen et al., 2017; Sanders, 2012; Voorheis, 2017), i.e. settings with much higher pollution levels.

The paper also adds to the literature on the development of socio-emotional skills, a topic of major interest as the returns to these skills have increased over recent decades relative to the returns to cognitive ability (Deming, 2017; Edin et al., 2017). Related studies e.g. analyze the effects of family income (Akee et al., 2018), birth order (Black et al., 2018; Houmark, 2023), parents' labor market incentives (Hufe, 2024) or child care arrangements (Gupta and Simonsen, 2010; Houmark et al., 2022). Persson and Rossin-Slater (2018) and Adhvaryu et al. (2019) also investigate the impact of prenatal conditions, specifically maternal stress and malnutrition, on later life mental health and socio-emotional ability. Grönqvist et al. (2020) analyze how childhood exposure to lead affects adult outcomes, and identify negative effects on socio-emotional skills as an important mechanism. Relative to them, I investigate the role of less toxic air pollutants, which - as an externality of economic production and traffic - are omnipresent in both developing and developed countries.

[2] Mean gestational PM2.5 exposure in my sample is 14.7 µg/m^3.

Finally, my results on how in-utero exposure to air pollution affects neuroticism and internalizing behavior at ages 5-10 add new evidence on the *missing midlle* of human capital formation (Almond et al., 2018). This term refers to the relative lack of knowledge on how early life intervention programs or shocks transition through middle childhood.

2. Background: Socio-Emotional Skills and Prenatal Air Pollution

To capture the different dimensions of socio-emotional ability, I rely on the Big Five personality traits and behaviors assessed in the SDQ at ages 5-10. There is ample evidence in developmental psychology that children differ substantially in character traits by school start age and that these traits are predictive of their adult personality (Almlund et al., 2011; De Pauw, 2017). Moreover, childhood socio-emotional ability is a significant determinant of educational achievement (Carneiro et al., 2007; Johnston et al., 2014). The Big Five are a widely used taxonomy, comprising openness, conscientiousness, extraversion, agreeableness and neuroticism. Among these traits, conscientiousness (the tendency to be organized, responsible, and hard-working) and emotional stability (the opposite of neuroticism) show the most robust positive correlations with labor market success (e.g. Almlund et al., 2011; Cubel et al., 2016; Fletcher, 2013). The SDQ assesses children's prosocial behavior, externalizing behavior (which is directed outward, linked to aggression and hyperactivity, and tends to create conflict with one's environment), and internalizing behavior (which is directed inward, leads to distress and includes symptoms such as social withdrawal, shyness, fearfulness, and anxiety). Internalizing behavior has been linked to lower adult earnings, while externalizing behavior can increase earnings (Papageorge et al., 2019).

Findings from brain lesion studies and psychopharmacological research imply that the source of all socio-emotional abilities lies in the brain (e.g. Almlund et al., 2011). Personality neuroscience uses neuroimaging techniques to identify how personality traits depend on brain structure and function as well as levels of hormones and neurotransmitters (Allen and DeYoung, 2017). As skills governed by the brain, socio-emotional abilities might be affected by exposure to air pollution. Ultrafine particles can reach the brain tissue where they cause oxidative stress and neuroinflammation. Since the in-utero period is a phase of rapid brain growth, gestational air pollution exposure might cause irreversible damages to the nervous system by disrupting this process (de Prado Bert et al., 2018). The pollutants that are most likely to cause such damages are carbon monoxide (CO) and ultrafine particles, as both can cross the placenta, and thus pose most harm to the fetus. Recent evidence from brain imaging studies indeed suggests that gestational air pollution exposure is associated with a reduction in white matter volume and changes to brain structure in humans (de Prado Bert et al., 2018).

3. Data

3.1 Child Outcomes

Estimating the impact of prenatal pollution exposure on socio-emotional ability requires not only information on individuals' skills but also on their location and time of birth. Besides, the data must cover a sufficiently large number of individuals from multiple birth cohorts, since only temporal variation in air pollution is used in the estimation.

A data source that addresses these demands is the German Socio-Economic Panel (SOEP), a household panel survey started in 1984, which covers roughly 15,000 households. Children's socio-emotional skills are assessed in the *Mother-and-child* questionnaires which were introduced in 2008 for children aged 5-6, and in 2012 for children aged 9-10 years. Each Big Five personality trait is assessed with two questions on the child's behavior, which mothers answer on a scale from 0 to 10. The questions are presented in Appendix Table A.1. Notably, the items underlying openness are likely to capture at least in part the child's cognitive ability (e.g., '*My child is quick at learning new things vs. needs more time*'). Thus, the focus of my analysis lies on the other four traits which isolate non-cognitive skills. To construct the outcomes of interest, I collect the relevant information from the 2008-2020 mother-and-child questionnaires, recode items where necessary such that higher values reflect higher realizations of the respective trait, add up values for items within each domain, and standardize the resulting scores within age groups (5-6 and 9-10).

Mothers also answer the SDQ, which comprises five domains: emotional symptoms, peer relationship problems, conduct problems, hyperactivity/inattention, and prosocial behavior. They are assessed with two to five questions each, which mothers answer on a scale from one to seven (see Appendix Table A.1).

Emotional symptoms and peer problems can be aggregated to capture *internalizing behavior*, which is linked to the Big Five trait neuroticism (Griffith et al., 2010; Almlund et al., 2011). The hyperactivity and conduct problem scores can be aggregated to capture *externalizing behavior*, which is associated with high extraversion, low agreeableness, and low conscientiousness (Ehrler et al., 1999; Mezquita et al., 2015). Analogues to the Big Five, I standardize the resulting scores within age groups.[3]

Figure 1 shows the distributions of these outcomes (not standardized). All variables show substantial variation, indicating that they successfully pick up differences in child personality and behavior. However, the distributions of the Big Five traits have small spikes at the intermediate value of 10 (Figure 1a). This suggests that some mothers have difficulties in assessing their children's personality and thus opt for the middle values. While the vast majority of respondents does not follow that strategy, suggesting that their answers are informative, I still address this potential issue in robustness checks, including sibling fixed effect models and dropping uninformative answers.

To assign pollution exposure to individuals, I rely on information on year and month of birth as well as county of residence. For more than two thirds of the children, I can identify the county of residence during the year of birth. In the remaining cases, the households entered the panel after the child was born such that the place of birth is unknown. As a proxy for county of birth, I assign the county of residence during the first wave the household was interviewed. Incorrect assignments can induce measurement error in prenatal pollution exposure which is not addressed by the IV strategy. Since measurement error causes attenuation bias, any results would reflect a lower bound.

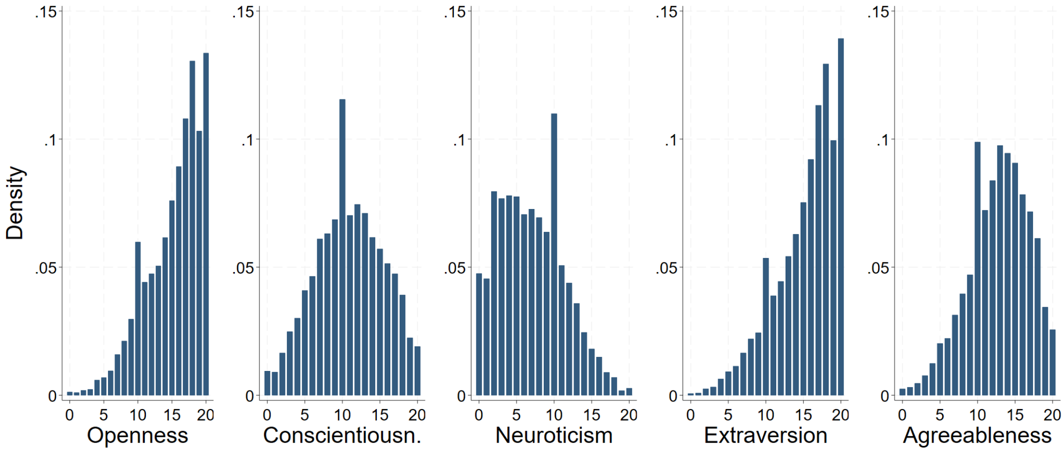
The SOEP also provides relevant demographic and socioeconomic background variables, e.g. the child's gender, age in month and migration history, whether it lives in a single-parent household, parental education[4], and birth order.

I restrict the analysis sample to children born in counties with at least 15 unique children in the data. Since effects are identified from variation in pre-

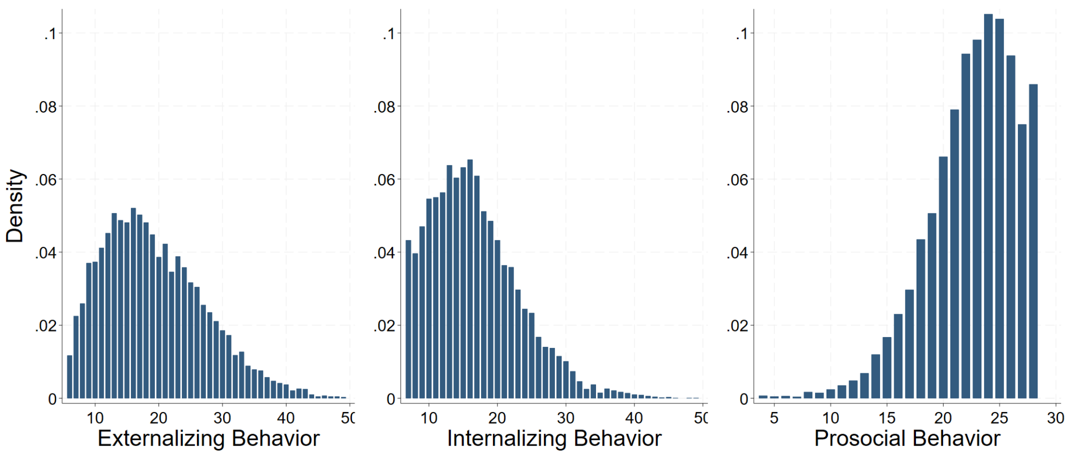
[3]These outcomes are available for fewer observations than the Big Five because the "Families in Germany" Study, which was integrated into the SOEP in 2014, included the Big Five, but not the SDQ.

[4]I construct dummy variables reflecting low, medium and high levels of education for mothers and fathers. Low education: no high school degree (Abitur) and no vocational training. Medium education: completed high school or vocational training, but no tertiary degree. High education: tertiary degree. I include children in the sample with missing information on parental education, and define separate dummies for these cases.

Figure 1: Distribution of mother-reported Socio-Emotional Skills



(a) Big Five



(b) Child Behaviors (SDQ)

Note: Based on data from the SOEP, version 38. The distributions are based on all available observations. Panel (a) is based on 13,300 observations, and panel (b) on 9,900 observations.

natal PM2.5 concentration between children born in the same county, this restriction serves to reduce the influence of outliers. This yields 11,242 observations of 8,253 unique individuals for the Big Five, and 7,119 observations of 5,464 unique individuals for the SDQ-based outcomes. Sample characteristics are displayed in Table A.2.

Validating the Outcomes. In Appendix B, I demonstrate that the mother-assessed outcome variables contain substantial information regarding the children's socio-emotional skills. I show that the mother-reported variables are significantly and positively correlated with father-reports at the same age, and with self-reports during adolescence . Moreover, I show that the mother-reported Big Five are significant predictors of the mother-assessed probability that the child will graduate from the academic track of the German school system and of child-reported life satisfaction, self esteem, and risk aversion, with plausible signs.

3.2 Environmental conditions

Air Quality. PM2.5 is the air pollutant of main interest because very small particles can cross the placenta, but it has only been measured comprehensively in Germany since 2008. I thus collect monthly PM2.5 concentrations from van Donkelaar et al. (2021), which are generated by combining satellite measurements with a chemical transport model, and calibrating to ground-based monitor readings. Concentrations are reported on a global 0.1° longitude-by-latitude grid (roughly 10km x 10km) and are available from 1998 onward. To aggregate data to counties, I average concentrations across all grid points falling into a county. In case of small counties without any grid point on their territory, I assign the inverse-distance weighted average from the 10 points closest to their centroid. Finally, I compute average PM2.5 during children's in-utero period, defined as the nine month period ending with the month of birth. Sample average gestational PM2.5 concentration is 14.7 $\mu g/m^3$, i.e. falling between the annual standards in the European Union (25 $\mu g/m^3$) and the US (9 $\mu g/m^3$).

In addition, I use daily readings of PM10 (particulate matter with a diameter of at most 10 µm), CO, and NO2 (nitrogen dioxide) from outdoor monitors operated by Germany's federal environmental agency (Umweltbundesamt, 2024). PM10 contains both PM2.5 and larger, less harmful particles, and serves to verify the results obtained with the satellite-based data. The latter two are common co-pollutants of particulate matter, and especially CO might also cause damage to the unborn child. To assign data to counties, I compute inverse distance weighted averages, based on up to three closest stations within a radius of 60km around the county centroid. I aggregate daily concentrations to the in-utero period (= 270 days), keeping only observations with less than 115 missing daily values.[5] Summary statistics on air quality are depicted in Table A.2.

[5] NO2 [CO] concentrations are measured consistently at 252 [64] stations over the sample period. Thus, CO is available for only a subset of all counties.

Thermal Inversions. During normal times, air temperature decreases with altitude. A thermal inversion occurs when temperature instead *increases* with altitude. To measure this, I use reanalysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF) on surface level and upper air temperature for the years 1999-2019. The data are reported at hourly frequency on a 0.25° latitude-longitude grid. I measure upper air temperature at a pressure level 50 hPa below the surface pressure in a county, which corresponds to approximately 400-500m higher altitude. Counties are assigned the inverse distance weighted average temperatures across all grid points within 30km of their centroid. For each day, I average both surface level and upper air temperature between 2 am and 6 am. If nightly upper air temperature exceeds nightly surface temperature, the county experiences a night-time inversion on that day. The instrumental variable is the share of days with a night-time inversion during the in-utero period. Its sample mean is 0.39, with a standard deviation (SD) of 0.07.

Weather conditions. I collect reanalysis data on average monthly meteorological conditions between 1999 and 2019 from the ECMWF, to include as control variables. Air temperature, precipitation, dewpoint temperature, and wind speed are reported on a 0.1° longitude-latitude grid. To aggregate data from the gridpoint × month- to the county × in-utero period-level, I proceed in the same way as with the PM2.5 data. I compute relative humidity from air temperature and dewpoint temperature using the R package `weathermetrics`.

4. Empirical Strategy

The baseline model relating children's socio-emotional skills to air pollution is given by:

$$SocEm_{icyma} = \beta PM_{cym} + \gamma' \mathbf{X}_i + \delta' \mathbf{W}_{cym} + \theta_c + \theta_y + \theta_m + \theta_a + u_{icyma}, \qquad (1)$$

where $SocEm_{icyma}$ denotes a measure of socio-emotional ability of individual i, born in county c in month m of year y, and assessed at age a. PM_{cym} is local PM2.5 concentration during the gestational period. The model includes county fixed effects, θ_c, to account for persistent difference in pollution and skill levels across locations, month-of-birth fixed effects, θ_m, to control for seasonality in air quality, and year-of-birth fixed effects, θ_y, to capture changes in socio-emotional skills and air quality over time which affect individuals in all counties equally. θ_a controls for age-specific trends in outcomes. Individual background characteristics in $\mathbf{X}_i$ comprise child gender, age in month and its square, migration background and parental education. Meteorological conditions during the gestational period, $\mathbf{W}_{cym}$, include third order polynomials in mean temperature and precipitation, as well as wind speed and relative humidity in linear form. Standard errors are clustered at the county level. Observations are weighted by the inverse of the number of times a child is observed in the data, i.e. children observed at both age 5-6 and 9-10 are assigned a weight of 0.5 in both cases.

In model 1, omitted variable bias might arise, e.g. due to region-specific economic shocks affecting both prenatal air pollution as well as parental income, which might be spent on investments into the child's skill development. Moreover, individual pollution exposure is measured with error, causing attenuation bias.

To address these issues, I exploit thermal inversions to extract exogenous variation in air quality, following e.g. Colmer et al. (2021) and Molina (2021). Under normal conditions, air temperature decreases with altitude. Emissions released at the ground level rise and disperse in the air. During an inversion, upper air layers are warmer than ground level air. The warm upper air acts like a ceiling that traps emissions close to the ground, causing an increase in surface-level pollution concentrations. Inversions are a meteorological phenomenon. They exhibit a seasonal pattern, with inversions occurring more frequently during winter than summer. However, conditional on month-of-birth fixed effects and weather controls, it is as good as random how often the specific combination of meteorological conditions occurs that gives rise to an inversion. Importantly, the frequency of inversions should be plausibly uncorrelated with local business cycles. At pollution levels common in Germany during the sample period, inversions usually do not lead to visible smog events or extremely poor air quality, and are thus unlikely to trigger avoidance behavior. Following e.g. Jans et al. (2018) and Molina (2021), I exclusively consider nighttime inversions, which should be even less likely to induce any behavioral responses.

I use this instrument in a Two-Stage Least-Squares (2SLS) approach with the following first stage model:

$$PM_{icym} = \alpha Inv_{cym} + \rho' \mathbf{X}_i + \eta' \mathbf{W}_{cym} + \phi_c + \phi_y + \phi_m + \phi_a + \epsilon_{icyma} \quad (2)$$

where Inv_{cym} is the share of days on which a nighttime inversion occurred during the in-utero period.

The first stage results in Appendix Table A.3 show that a standard deviation increase in prenatal inversion frequency (SD = 0.07), induces a rise in prenatal PM2.5 concentrations by 0.36 to 0.45 $\mu g/m^3$ across analysis samples. This corresponds to 17-21% of a standard deviation of PM2.5. A more extreme increase in inversion frequency by 0.233 – a shift from the 5th to the 95th percentile of the sample distribution – implies a change in PM2.5 by 1.2-1.5 $\mu g/m^3$ or 56-71% of a standard deviation. The associated F-Statistics are large, exceeding or just slightly below the threshold of 104 for valid t-ratio inference in a single-instrument model (Lee et al., 2022). Inversions also significantly increase readings of PM10, NO2 and CO at ground monitors, but for CO, the F-Statistic is far below 10. With only one instrument, I cannot disentangle the distinct effects of the different pollutants on socio-emotional development. However, the estimates suggest that PM2.5 is likely to drive results in the second stage, because (i) unlike CO and fine particles, NO2 and coarse particles are not commonly considered as major risk factors in the medical literature, and (ii) the first stage effect for CO is weak.

To test the credibility of the exogeneity assumption, I check whether inversion frequency is systematically correlated with observed predetermined family charcateritsics (Appendix Table A.4). None of the tested variables are significantly correlated with the instrument and the point estimates are close to zero.

5. Results

5.1 Main results

Panel A of Table 1 displays 2SLS estimates of the effect of in-utero exposure to PM2.5 on the Big Five personality traits. I find a significant impact on neuroticism which increases by 13.5% of a standard deviation for an increase in PM2.5 concentration by 1 $\mu g/m^3$.[6] This implies that individuals exposed to higher levels of air pollution during the prenatal period are less emotionally stable in childhood, i.e. more fearful and less self-confident. For the remaining four traits, I cannot reject the null that prenatal particulate matter exposure

[6]The coefficient remains significant at the 5%-level after applying the tF-correction following Lee et al. (2022).

Table 1: Effects of prenatal PM2.5 on Childhood Socio-emotional Skills

Panel A: Big Five	Openness	Conscientious	Extraversion	Agreeable	Neuroticism
PM2.5	−.078	−.043	−.001	.020	.135**
in-utero	(.069)	(.072)	(.071)	(.069)	(.066)
Observations	11,243	11,243	11,239	11,193	11,242
Counties	219	219	219	218	219
1st Stage F-Stat.	100	101	99	99	101
Panel B: SDQ	Externalizing Behavior	Internalizing Behavior	Prosocial Behavior		
PM2.5	.032	.185**	−.021		
in-utero	(.072)	(.074)	(.071)		
Observations	7,183	7,119	7,188		
Counties	163	160	162		
1st Stage F-Stat.	135	131	134		

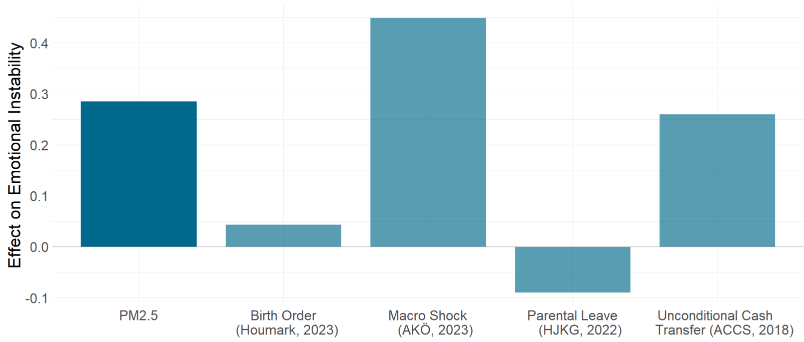
Note: The Table displays 2SLS-estimates of the parameter β in Equation (2), reflecting the effects of a 1 $\mu g/m^3$ increase in PM2.5 concentration during the in-utero period on the Big Five (Panel A) and on externalizing, internalizing, and prosocial behavior based on the Strength and Difficulties Questionnaire (Panel B). PM2.5 is instrumented by inversion frequency. Outcomes are standardized within age groups. Regressions are weighted by the inverse of the number of times a child is observed in the data. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

has no effect on these outcomes. The largest point estimate emerges for Openness (−0.078 SD), which, as mentioned above, likely partially captures cognitive skills. Thus, the negative point estimate is in line with existing results on the negative impact of in-utero exposure to air pollution on cognitive ability. The fact that it is not statistically significant is unsurprising given that this measure of cognitive ability is relatively crude. The remaining point estimates are substantially smaller in absolute terms.

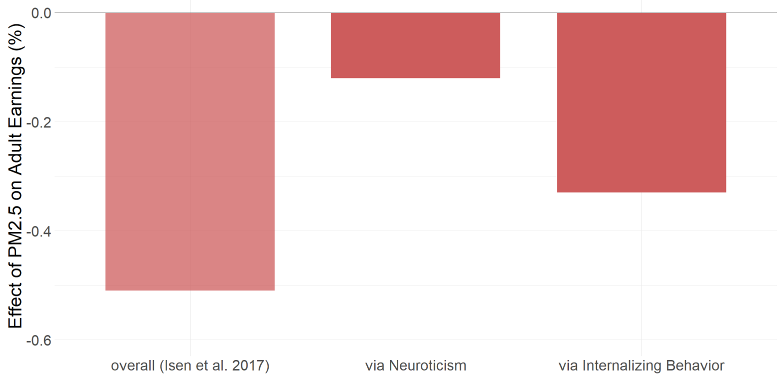
Panel B of Table 1 displays 2SLS estimates of the effect of PM2.5 exposure on child behaviors assessed in the SDQ. Internalizing behavior – the outcome most related to neuroticism – increases by 18.5% of a standard deviation in response to a unit increase in prenatal PM2.5, implying that affected children are shyer, more fearful and have more problems in interacting with peers. Point estimates for externalizing problems and prosocial behavior are small and insignificant.

Is the effect of fetal PM2.5 exposure on emotional stability economically relevant? To assess its magnitude, in Figure 2a I compare the impact of a standard deviation increase in PM2.5 (2.1 $\mu g/m^3$) on Neuroticism to the effects of four other factors that have been found to cause significant changes in this outcome during childhood or adolescence. The comparison strongly suggests that the

Figure 2: Effect Magnitude



(a) Change in Neuroticism for different Treatments



(b) Implied Earnings Effects of PM2.5

Note: Panel (a) illustrates the magnitude of the effects of different treatments on emotional instability, measured in standard deviations of the outcome. Effect of PM2.5 is based on Table (1) and refers to a standard deviation increase in prenatal PM2.5. The effect of birth order on Neuroticism is from Houmark (2023), the effect of a macro shock on self-confidence is from Azmat et al. (2023), the effect of a parental leave extension (by 3.2 months) on Neuroticism is from Houmark et al. (2022), and the effect of an unconditional cash transfer (of $3,500 on average) on Neuroticism is from Akee et al. (2018). Panel (b) shows estimated percentage changes in adult earnings due to a 1 µg/m^3 increase in prenatal PM2.5 via its effects on neuroticism and internalizing behavior, respectively. Estimates are derived by combining results from Table (1) with effects given in Fletcher (2013, Table 4) and Papageorge et al. (2019, Table 1). The left bar depicts the overall earnings effect of pollution in the year of birth from Isen et al. (2017, Table 4). Their estimates for TSP exposure are transformed to a 1 unit increase in PM2.5 using a TSP-PM2.5 ratio of 4.38 (Voorheis, 2017).

effect of prenatal PM2.5 exposure is of relevant magnitude, exceeding e.g. the size of effects found for birth order, an extension of maternity leave, or an unconditional cash transfer.

To assess how important socio-emotional skills are as a mechanism for the adverse long-run impacts of prenatal pollution exposure, I compare the effect of PM2.5 on neuroticism and internalizing behavior to estimates for the impact on cognitive ability and physical health. Results by Molina (2021) imply that a one µg/m^3 increase in PM10 concentration during the second trimester of pregnancy reduces cognitive ability in adulthood by 2.4% of a standard deviation.[7] Sanders (2012) finds that a 10 µg/m^3 increase in total suspended particles (TSP, which comprises PM10 and larger, less harmful particles) exposure during the year of birth reduces high school math test scores by 6% of a standard deviation. For context, in the US the average TSP-to-PM2.5 ratio is estimated to be 4.38 (Voorheis, 2017). While these effects are smaller then my results, the outcomes were assessed at older ages. Regarding physical health impacts, Ferro et al. (2024) find that a one µg/m^3 increase in prenatal PM10 exposure increases annual hospitalisations between ages 5 and 10 by 1-3%. Overall, the comparisons suggest that relative to cognitive skills and physical health, also socio-emotional ability – specifically emotional stability – is a relevant channel through which early life pollution exposure generates adverse long-run effects.

To approximate the earnings impact of poor air quality via the reduction in emotional stability, I conduct a back-of-the-envelope calculation. Using a sibling fixed effects approach, Fletcher (2013) finds that a standard deviation increase in neuroticism in young adulthood reduces annual earnings by 5-6%. Combining this with my 2SLS results implies that a standard deviation increase in prenatal PM2.5 concentration reduces annual earnings by .23%-.27% through its effect on neuroticism.[8] Similarly, combining the estimated effect of PM2.5 on internalizing behavior with estimates on how this outcome affects earnings (Papageorge et al., 2019) implies a slightly larger reduction in adult earnings by .63%-.74%.[9] For context, Isen et al. (2017) find that a reduction in TSP during the year of birth by 10 µg/m^3 increases adult earnings by 1%-1.4%. While a comparison of results derived from different settings and time periods has limitations, it at least suggests that the socio-emotional ability channel plays a relevant role (see Figure 2b).

[7] This number is derived by combining quasi first stage- and reduced form results in Molina (2021).

[8] A SD increase in gestational PM2.5 exposure is estimated to raise Neuroticism during childhood by 2.1 x .135 = 0.28 SD. To approximate how gestational air pollution exposure affects neuroticism in early adulthood, I multiply this with the correlation between neuroticism in childhood and at ages 16-17 (Figure B.1): 0.28 x 0.16 = 0.045 SD. This implies an earnings reduction of 0.045 x 5% to 0.045 x 6%.

[9] Papageorge et al. (2019) find that a SD increase in internalizing behavior at age 11 reduces adult earnings by 5 to 5.8%. I find that a SD increase in PM2.5 increases internalizing behavior by 2.1 x 0.185 = 0.389 SD at ages 5-10. Based on the correlation in Figure B.1 this implies an increase in internalizing behavior at age 11-12 by 0.128 SD. Hence, a SD increase in prenatal PM2.5 would lead to a 0.128 x 5% to 0.128 x 5.8% decrease in adult earnings.

5.2 Additional Results

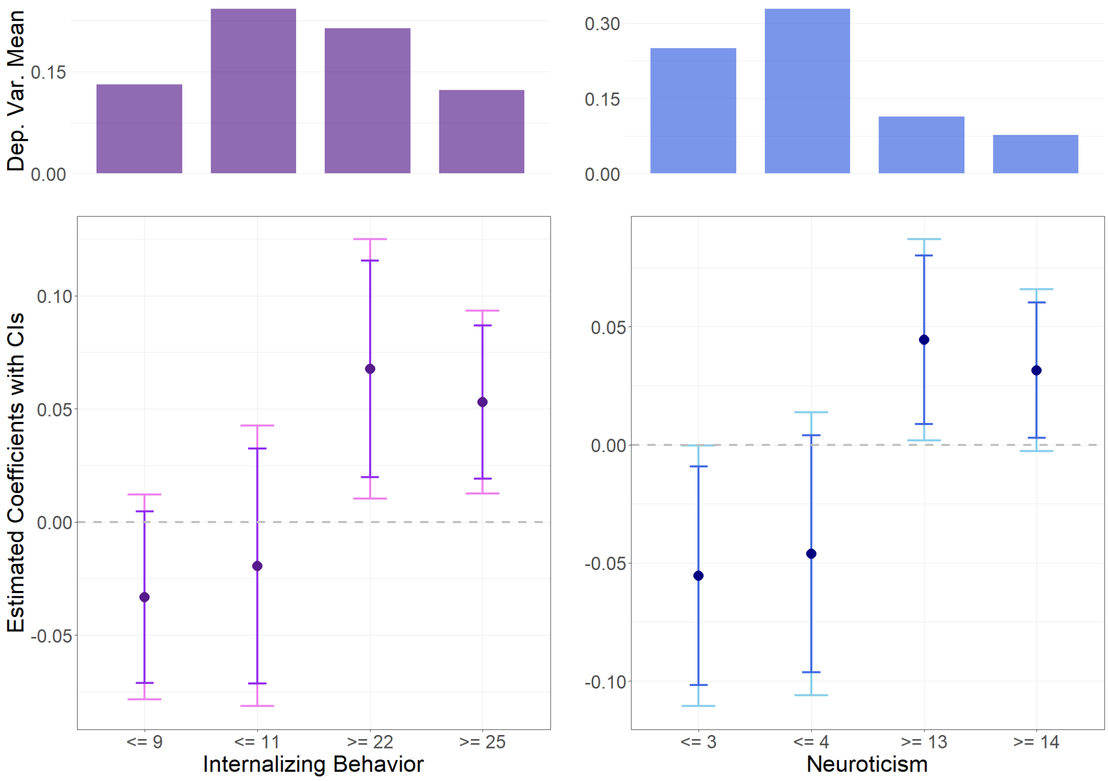
Having established that in-utero exposure to particulate matter decreases emotional stability, I next examine this effect in terms of heterogeneity and critical windows of exposure.

Effect Heterogeneity. I analyze effect heterogeneity with respect to child and family characteristics, namely child gender, age at which the outcomes are is assessed (5-6 vs. 9-10 years), maternal education, and current household income (above vs. below median). To do so, I run the reduced form version of the IV model and include an interaction term between inversion frequency and the characteristic of interest. Results from the baseline reduced form model and the models including interaction terms are illustrated in Appendix Figure A.1. I find that the effect of a standard deviation increase in inversion frequency on neuroticism and internalizing behavior is about 16-28% smaller for children in richer households as compared to poorer households. This suggests that parental resources can be used for remediating investments into children's emotional stability. I also find strong heterogeneity by gender for internalizing behavior, with the effect of inversions being more than twice as large for boys as compared to girls, but no comparable pattern for neuroticism. There is no significant heterogeneity by age or maternal education.

I also examine heterogeneity in effect magnitude along the distribution of the outcomes. I construct indicator variables for the value of neuroticism or internalizing behavior, respectively, falling below or above certain thresholds (before standardization) and use them as regression outcomes. Air pollution exposure affects neuroticism at both ends of the distribution. It reduces the probability to have a very low score and increases the probability to have a very high score (Figure 3). In terms of magnitude, the probability to have a score of at most three falls by 5.5 pp., whereas the probability to have a score of at least 13 increases by 4.5 pp. The effects in the upper tail are, however, larger relative to the lower baseline probability of high scores. For internalizing behavior, the overall effects are mostly driven by increases in the upper part of the distribution.

The increase in the probability of very high scores for both outcomes is noteworthy as it implies potential effects on mental health. In modern personality psychology, mental disorders are conceptualized as extreme realizations of the Big Five traits. High values of neuroticism show particularly robust and con-

Figure 3: PM2.5 Effects along the Distribution of the Outcomes



Note: Bottom right plot: 2SLS point estimates of the effect of an increase in prenatal PM2.5 concentration by 1 µg/m^3 on indicator variables for a child's neuroticism score falling into different ranges, as depicted on the x-axis, along with 90%- and 95%-confidence intervals. The score can vary between 0 and 20. Top right plot: sample averages of the dependent variables. Bottom left plot: 2SLS point estimates of the effect of prenatal PM2.5 concentration on indicator variables for the internalizing behavior score from the SDQ falling into different ranges, as depicted on the x-axis, along with 90%- and 95%-confidence intervals. The score can vary between 7 and 49. Top left plot: sample averages of the dependent variables.

sistent correlations with a range of mental health issues, e.g. depression and anxiety disorders (Almlund et al., 2011). Similarly, internalizing problems are often linked to these issues. Hence, prenatal particulate matter exposure might not just reduce an important socio-emotional skill within the range of 'normal' variations in personality, but could even give rise to mental health problems.

Timing of Exposure. To assess critical windows of exposure, I split the in-utero period into three trimesters, and run the reduced form regression with variables measuring inversion frequency separately for each trimester. Regressions include the same individual and family characteristics as before and trimester-specific weather controls. I find that the overall effects are driven by the second and third trimesters (Table A.5), in line with findings on the effect of air pollution on cognitive ability (Molina, 2021, significant effect only in the second trimester) and on performance in school tests (Bharadwaj et al., 2017, effect mainly in the third trimester).

Additionally, I investigate whether postnatal air pollution exposure also affects emotional stability. Table A.6 shows results from 2SLS regressions, where I find insignificant coefficients on PM2.5 concentration during the first nine months *after* birth for both outcomes of interest. As an alternative, I estimate a reduced form model including both inversion frequency in the postnatal period as well as inversion frequency during the nine months *before* conception. Coefficients are insignificant for both of these periods, but remain significant for the in-utero period. The absence of any effect during the placebo period before conception suggests that the effect during the in-utero period operates solely through physiological channels.

Robustness Tests. The main finding implies that in-utero exposure to particulate matter increases neuroticism and internalizing behavior. In Appendix C, I test the robustness of this result to (i) changes in sample construction and weighting, (ii) inclusion of additional background controls, (iii) use of more stringent fixed effects, (iv) variation in the included weather controls, (v) adding a lead and lag of the instrument to capture potential autocorrelation, (vi) using monitor-measured PM10 concentration instead of satellite-assessed PM2.5 concentration, and (vii) dropping 'uninformative' answers, where mothers pick intermediate values very often, or have a very low variance of answers across all relevant survey items.[10]

Finally, I incorporate the IV approach into a model with family fixed effects, exploiting only variation in inversion-induced air pollution within sibling groups. This model accounts for potential differences in mothers' subjective assessment of child behavior or their reference groups, which might confound the results, because only variation in socio-emotional outcomes between children of the same mother is used. Information on neuroticism (internalizing behavior) is available for 5,811 children across 2,503 families (3,917 children across 1,749 families). The results are depicted in Table 2. I find positive coefficients that are larger than in the baseline model for both neuroticism and internalizing behavior, but the effect on neuroticism is not statistically significant. This is not surprising given the reduced sample sizes and identifying variation. How-

[10]Respondents with high capability to memorize and recall relevant information about their child's behavior are likely to deviate more in their answers from the mean across different items (Falk et al., 2021)

Table 2: Effects of PM2.5: Models with Family Fixed Effects

	Neuroticism		Internalizing Behavior	
PM2.5	.145		.317***	
in-utero	(.096)		(.121)	
Inversion Frequency		.053		.125***
in-utero		(.043)		(.057)
Observations	8,045	8,045	5,119	5,119
Unique Children	5,811	5,811	3,917	3,917
Sibling Groups	2,503	2,503	1,749	1,749
F-Statistic	62		43	

Notes: Columns 1 and 3 shows 2SLS results from a regression of Neuroticism and internalizing behavior, respectively, on prenatal PM2.5 concentration, instrumented with inversion frequency during the same period. Regression controls for family-, year-, month-, and age group fixed effects, weather conditions during the in-utero period, child's gender, age in months and its square,. Columns 2 and 4 show reduced form results. All regressions are weighted by the inverse of the number of times a child is observed in the data. tF-adjusted standard errors clustered at the county level in parentheses. *p<0.1; **p<0.05; ***p<0.01

ever, the fact that I continue to find positive and sizable estimates alleviates concerns about systematic differences in reporting standards across mothers.[11]

6. Conclusion

Exploiting quasi-random variation in air quality induced by thermal inversions, I study how in-utero exposure to air pollution affects children's socio-emotional development. I find that a 1 $\mu g/m^3$ increase in PM2.5 increases neuroticism and internalizing behavior by 13% and 18% of a standard deviation, respectively, implying that affected children have lower emotional stability. These effects are of the same order of magnitude as effects of gestational pollution exposure on cognitive ability reported in the literature, and are a plausible channel contributing to adverse long-run earnings effects.

These results have important policy implications since socio-emotional skills are malleable during childhood and adolescence. Investments targeted at emotional stability, e.g. mentoring programs, might be a feasible strategy to alleviate the negative impacts of prenatal air pollution exposure on education and earnings. Individuals born in places and during periods of poor air quality, e.g. from natural sources such as wild fires, could at least in part be compensated

[11] The IV estimate for internalizing behavior remains significant at the 5% level after tF-adjustment (Lee et al., 2022).

for these bad starting conditions.

The finding that the effects on neuroticism and internalizing behavior are driven by increases at the upper end of the distribution suggests that in-utero exposure to air pollution might not only reduce emotional stability within the range of 'normal' variations in socio-emotional ability, but could even induce mental health issues. A more comprehensive analysis of the relationship between gestational pollution exposure and mental health is an important avenue for further research.

References

Adhvaryu, A., J. Fenske, and A. Nyshadham (2019). Early life circumstance and adult mental health. *Journal of Political Economy 127*(4), 1516–1549.

Adhvaryu, A., N. Kala, and A. Nyshadham (2023). Returns to on-the-job soft skills training. *Journal of Political Economy 131*(8), 2165–2208.

Akee, R., W. Copeland, E. J. Costello, and E. Simeonova (2018). How does household income affect child personality traits and behaviors? *American Economic Review 108*(3), 775–827.

Alan, S., T. Boneva, and S. Ertac (2019, 02). Ever Failed, Try Again, Succeed Better: Results from a Randomized Educational Intervention on Grit*. *The Quarterly Journal of Economics 134*(3), 1121–1162.

Allen, T. A. and C. G. DeYoung (2017). Personality neuroscience and the five factor model. *Oxford handbook of the five factor model 319*, 52.

Almlund, M., A. L. Duckworth, J. Heckman, and T. Kautz (2011). Personality psychology and economics. In *Handbook of the Economics of Education*, Volume 4, pp. 1–181. Elsevier.

Almond, D., J. Currie, and V. Duque (2018, December). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature 56*(4), 1360–1446.

Arceo, E., R. Hanna, and P. Oliva (2016). Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from mexico city. *The Economic Journal 126*(591), 257–280.

Azmat, G., K. M. Kaufmann, and Y. Özdemir (2023). Socioemotional development during adolescence: Evidence from a large macro shock.

Banzhaf, H. S., L. Ma, and C. Timmins (2019). Environmental justice: Establishing causal relationships. *Annual Review of Resource Economics 11*, 377–398.

Bharadwaj, P., M. Gibson, J. G. Zivin, and C. Neilson (2017). Gray matters: Fetal pollution exposure and human capital formation. *Journal of the Association of Environmental and Resource Economists 4*(2), 505–542.

Black, S. E., E. Grönqvist, and B. Öckert (2018). Born to lead? the effect of birth order on noncognitive abilities. *Review of Economics and Statistics 100*(2), 274–286.

Carneiro, P., C. Crawford, and A. Goodman (2007). The impact of early cognitive and non-cognitive skills on later outcomes. cee dp 92. *Centre for the Economics of Education (NJ1)*.

Colmer, J., D. Lin, S. Liu, and J. Shimshack (2021). Why are pollution damages lower in developed countries? insights from high-income, high-particulate matter hong kong. *Journal of Health Economics 79*, 102511.

Colmer, J. and J. Voorheis (2021). The grandkids aren't alright: The intergenerational effects of prenatal pollution exposure.

Cubel, M., A. Nuevo-Chiquero, S. Sanchez-Pages, and M. Vidal-Fernandez (2016). Do personality traits affect productivity? evidence from the laboratory. *The Economic Journal 126*(592), 654–681.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Currie, J., J. Voorheis, and R. Walker (2023, January). What caused racial disparities in particulate exposure to fall? new evidence from the clean air act and satellite-based measures of air quality. *American Economic Review 113*(1), 71–97.

De Pauw, S. S. (2017). Childhood personality and temperament. *The Oxford handbook of the five factor model 1*.

de Prado Bert, P., E. M. H. Mercader, J. Pujol, J. Sunyer, and M. Mortamais (2018). The effects of air pollution on the brain: a review of studies interfacing environmental epidemiology and neuroimaging. *Current environmental health reports 5*, 351–364.

Deming, D. J. (2017). The growing importance of social skills in the labor market. *The quarterly journal of economics 132*(4), 1593–1640.

Edin, P.-A., P. Fredriksson, M. Nybom, and B. Öckert (2017). The Rising Return to Non-Cognitive Skill. Technical report, IZA Discussion Papers.

Ehrler, D. J., J. G. Evans, and R. L. McGhee (1999). Extending big-five theory into childhood: A preliminary investigation into the relationship between big-five personality traits and behavior problems in children. *Psychology in the Schools 36*(6), 451–458.

Falk, A., T. Neuber, and P. Strack (2021, July). Limited Self-Knowledge and Survey Response Behavior. CRC TR 224 Discussion Paper Series 2021-307, University of Bonn and University of Mannheim, Germany.

Ferro, S., A. Palma, C. Serra, and M. Stafoggia (2024). Beyond birth: The medium-term health impact of prenatal exposure to air pollution. *Journal of Environmental Economics and Management 127*, 103009.

Fletcher, J. M. (2013). The effects of personality traits on adult labor market outcomes: Evidence from siblings. *Journal of Economic Behavior & Organization 89*, 122–135.

Goebel, J., M. M. Grabka, S. Liebig, M. Kroh, D. Richter, C. Schröder, and J. Schupp (2019). The german socio-economic panel (soep). *Jahrbücher für Nationalökonomie und Statistik 239*(2), 345–360.

Griffith, J. W., R. E. Zinbarg, M. G. Craske, S. Mineka, R. D. Rose, A. M. Waters, and J. M. Sutton (2010). Neuroticism as a common dimension in the internalizing disorders. *Psychological Medicine 40*(7), 1125–1136.

Grönqvist, H., J. P. Nilsson, and P.-O. Robling (2020). Understanding how low levels of early lead exposure affect children's life trajectories. *Journal of Political Economy 128*(9), 3376–3433.

Gupta, N. D. and M. Simonsen (2010). Non-cognitive child outcomes and universal high quality child care. *Journal of public Economics 94*(1-2), 30–43.

Houmark, M. A. (2023). First among equals? how birth order shapes child development. *How Birth Order Shapes Child Development (November 8, 2023)*.

Houmark, M. A., C. M. Løchte Jørgensen, I. L. Kristiansen, and M. Gensowski

(2022). Effects of extending paid parental leave on children's socio-emotional skills and well-being in adolescence. *Univ. of Copenhagen Dept. of Economics Discussion Paper, CEBI Working Paper 14*, 22.

Hufe, P. (2024). The parental wage gap and the development of socio-emotional skills in children.

Isen, A., M. Rossin-Slater, and W. R. Walker (2017). Every breath you take—every dollar you'll make: The long-term consequences of the clean air act of 1970. *Journal of Political Economy 125*(3), 848–902.

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

Johnston, D., C. Propper, S. Pudney, and M. Shields (2014). Child mental health and educational attainment: Multiple observers and the measurement error problem. *Journal of Applied Econometrics 29*(6), 880–900.

Klauber, H., F. Holub, N. Koch, N. Pestel, N. Ritter, and A. Rohlf (2024). Killing prescriptions softly: Low emission zones and child health from birth to school. *American Economic Journal: Economic Policy 16*(2), 220–248.

Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter (2022, October). Valid t-ratio inference for iv. *American Economic Review 112*(10), 3260–90.

Mezquita, L., M. I. Ibáñez, H. Villa, L. Fañanás, J. Moya-Higueras, and G. Ortet (2015). Five-factor model and internalizing and externalizing syndromes: A 5-year prospective study. *Personality and Individual Differences 79*, 98–103.

Molina, T. (2021). Pollution, ability, and gender-specific investment responses to shocks. *Journal of the European Economic Association 19*(1), 580–619.

Papageorge, N. W., V. Ronda, and Y. Zheng (2019). The economic value of breaking bad: Misbehavior, schooling and the labor market. Technical report, National Bureau of Economic Research.

Persson, P. and M. Rossin-Slater (2018). Family ruptures, stress, and the mental health of the next generation. *American economic review 108*(4-5), 1214–1252.

Rosales-Rueda, M. and M. Triyana (2019). The persistent effects of early-life exposure to air pollution evidence from the indonesian forest fires. *Journal of Human Resources 54*(4), 1037–1080.

Rüttenauer, T. (2018). Neighbours matter: A nation-wide small-area assessment of environmental inequality in germany. *Social Science Research 70*, 198–211.

Sanders, N. J. (2012). What doesn't kill you makes you weaker: Prenatal pollution exposure and educational outcomes. *Journal of Human Resources 47*(3), 826–850.

Sorrenti, G., U. Zölitz, D. Ribeaud, and M. Eisner (2024, 02). The Causal Impact of Socio-Emotional Skills Training on Educational Success. *The Review of Economic Studies*, rdae018.

Umweltbundesamt (2024). Luftdaten. Technical report, Umweltbundesamt, Fachgebiet II 4.2, Beurteilung der Luftqualität.

van Donkelaar, A., M. S. Hammer, L. Bindle, M. Brauer, J. R. Brook, M. J. Garay, N. C. Hsu, O. V. Kalashnikova, R. A. Kahn, C. Lee, R. C. Levy, A. Lyapustin, A. M. Sayer, and R. V. Martin (2021). Monthly global estimates of fine particulate matter and their uncertainty. *Environmental Science & Technology 55*(22), 15287–15300. PMID: 34724610.

Voorheis, J. (2017). Air quality, human capital formation and the long-term effects of environmental inequality at birth. Technical report, Center for Economic Studies, US Census Bureau.

Online Appendix

A Additional Tables and Figures

Table A.1: Socio-emotional Skills in the SOEP Mother-and-child-questionnaires

Panel A: Big Five	
Dimension	*How would you rank your child in comparison to other children of the same age? My child is ...*
Openness	understands quickly — needs more time not that interested -– hungry for knowledge
Conscientiousness	focused – easily distracted tidy -– untidy
Extraversion	talkative -– quiet withdrawn -– sociable
Agreeableness	obstinate – compliant good-natured — irritable
Neuroticism	self-confident -– insecure fearful — fearless

Panel B: Strength and Difficulties Questionnaire		
Aggregate Scale	*Subscale*	*To what extent do the following statements apply to your child?*
Internalizing Behavior	Emotional Problems	is often unhappy or dejected is nervous/clingy in new situations, loses self-confidence easily has many fears, becomes frightened easily
	Peer Problems	is a loner, usually plays by him/herself is popular with other children is often made fun of or picked on by other children gets along better with adults than with other children
Externalizing Behavior	Hyperactivity	is agitated, hyperactive, cannot sit still is fidgety is easily distracted and lacks concentration finishes tasks, is able to concentrate thinks before acting
	Conduct Problems	often has tantrums, has a temper quarrels a lot with other children, picks on them
Prosocial Behavior	Prosocial Behavior	is considerate likes to share with others (sweets, toys, crayons) is helpful if others are hurt, sick, or sad helps others of his/her own accord

Note: Each questions on the Big Five is answered on an 11-point Likert scale ranging from 0 (= does not apply at all) to 10 (= fully applies). Each questions from the SDQ is answered on an 7-point Likert scale ranging from 1 (= does not apply at all) to 7 (= fully applies).

Table A.2: Summary Statistics

	Big Five				SDQ			
Observations	11,242				7,119			
Unique children	8,253				5,464			
Counties	219				160			
Birth cohorts	2000-2014				2002-2014			
Age [years]	7.8				7.8			
Migration background [%]	29.6				32.9			
College-educated mothers [%]	26.5				28.5			
Single-parent households [%]	16.4				14.0			
County of birth imputed [%]	31.1				23.1			
	Mean	SD	5th Pc.	95th Pc.	Mean	SD	5th Pc.	95th Pc.
PM2.5 in-utero [μg/m^3]	14.7	2.1	11.6	18.6	14.7	2.1	11.6	18.5
NO2 in-utero [μg/m^3]	28.8	10.9	12.5	47.5	29.0	11.0	12.6	47.6
CO in-utero [μg/m^3]	467.8	191.1	210	844	439.5	171.8	204	781
Inversions in-utero [%]	39.2	7.0	27.3	50.9	38.8	6.7	27.1	50.2

Note: Based on data from the SOEP, version 38. Samples include all available observations per child for the outcomes given at the top of the table, but only for children born in counties with at least 15 unique individuals in the sample.

Table A.3: First Stage Results

	PM2.5	PM10	NO2	CO
Panel A: Sample for Big Five				
Inversions	0.364***	0.901***	0.702***	11.647**
in utero	(0.036)	(0.097)	(0.079)	(4.994)
Observations	11,242	7,810	10,590	6,721
Counties	219	162	207	139
1st Stage F-Stat.	100.8	86.9	79.2	5.4
Outcome Mean	14.72	23.80	28.91	467.42
Panel C: Sample for SDQ				
Inversions	0.449***	0.780***	0.814***	11.65**
in utero	(0.039)	(0.094)	(0.097)	(4.679)
Observations	7,119	5,795	6,865	4,219
Counties	160	139	155	96
1st Stage F-Stat.	130.9	68.6	69.9	6.2
Outcome Mean	14.71	23.49	29.00	440.83

Note: The table depicts OLS estimates of the parameter α in Equation (2), reflecting the effects of inversion frequency during the nine-month in-utero period on air pollution concentration during the same period. All outcomes are measured in μg/m^3. Coefficients reflect the effect of a standard deviation increase in inversion frequency (SD = 0.07). Regressions include weather controls and individual background characteristics as described in the text. Regressions are weighted by the inverse of the number of times a child is observed in the data. Standard errors clustered at the county level are reported in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table A.4: Instrument Validity: Falsification Tests

	Maternal age	Migration history	Mother: Tertiary degree	Mother: Less than high school	Father: Tertiary degree	Father: Less than high school
Panel A: Sample for Big Five						
Inversions	-.099	-.015	-.013	.007	-.008	.007
in utero	(.145)	(.012)	(.011)	(.011)	(.015)	(.013)
Observations	11,033	11,242	10,938	10,938	7,420	7,420
Panel C: Sample for SDQ						
Inversions	-.133	-.006	-.013	-.005	-.013	.006
in utero	(.195)	(.014)	(.015)	(.014)	(.020)	(.018)
Observations	6,931	7,119	6,860	6,860	4,523	4,523

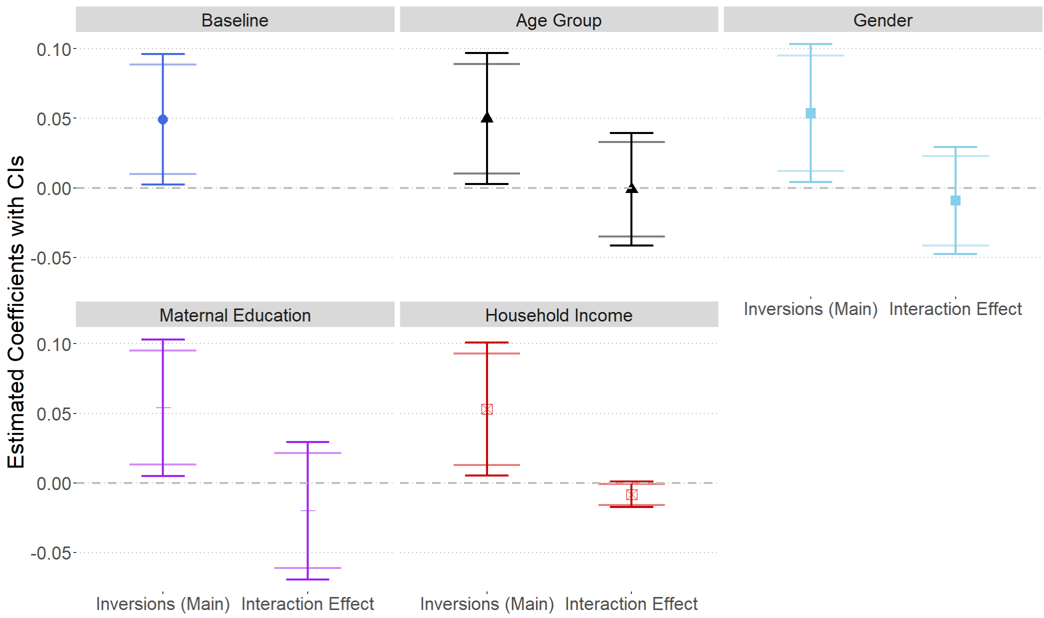
Note: The table depicts coefficients from OLS regressions of family characteristics on inversion frequency during the child's in-utero period. Estimates reflect the effect of a standard deviation increase in inversion frequency. Maternal age is measured in years. The other outcomes are dummy variables. Regressions control for county, year, month and age group fixed effects, child's gender and age in month, and weather controls (temperature and precipitation in cubic form, and wind speed and relative humidity in linear form). Regressions are weighted by the inverse of the number of times a child is observed in the data. Standard errors clustered at the county level are reported in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01

Table A.5: Effects of Air Pollution by Trimester

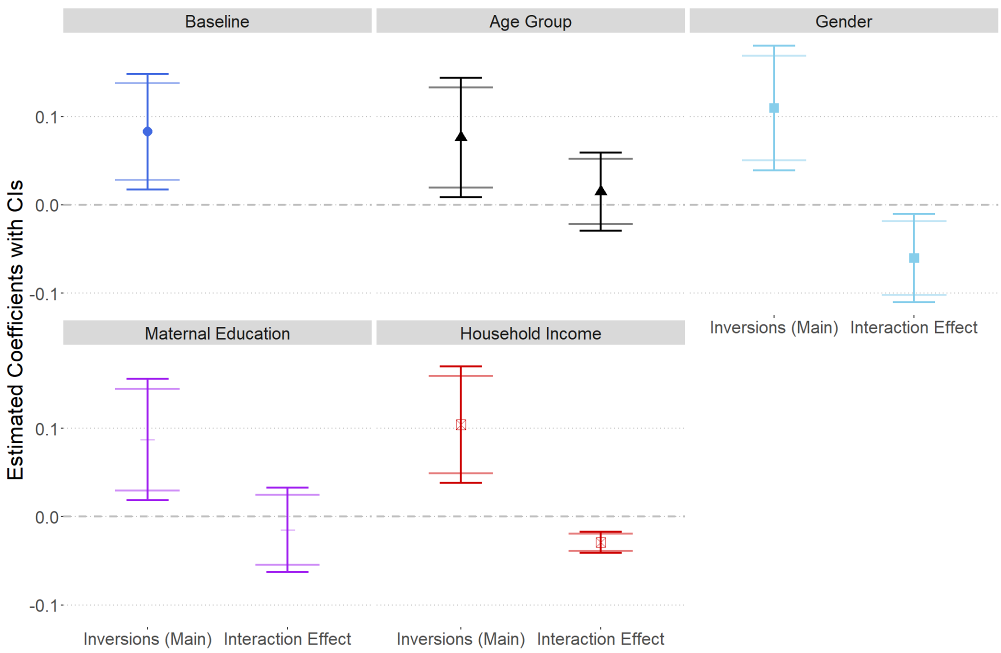
	Neuroticism	Internalizing Behavior
Inversion Frequency	.0128	.0181
1st trimester	(.0217)	(.0256)
Inversion Frequency	.0373**	.0504*
2nd trimester	(.0177)	(.0258)
Inversion Frequency	.0300	.0784***
3rd trimester	(.0198)	(.0296)
Observations	11,244	7,119

Note: Table depicts results from a regressions of age group-standardized neuroticism and internalizing behavior on variables measuring inversion frequency during the three trimesters of the in-utero period. Estimated coefficients reflect the effect of a standard deviation increase in inversion frequency (SD = 0.115). Controls include , year-, month-, county- and age group fixed effects, child gender, age in month and its square, migration background, parental education, and weather controls (temperature and precipitation in cubic form, and wind speed and relative humidity in linear form) for each trimester. Standard errors clustered at the county level are in parentheses.

Figure A.1: Effect Heterogeneity



(a) Neuroticism



(b) Internalizing Behavior

Note: Plots show estimated coefficients from OLS regressions of age group-standardized Neuroticism (panel a) and internalizing behavior (panel b) on a variable measuring inversion frequency, and its interaction with the factor given in the plot titles. Age Group: indicator for child being aged 9-10 years. Gender: indicator for child being female. Maternal Education: indicator for mother having a university degree. Household income: indicator for above median income. Fixed effects and controls as in Equation (2). Estimates reflect the effect of a standard dev. increase in inversion frequency. Error bars reflect 90%- and 95% confidence intervals.

Table A.6: Effects of PM2.5: Postnatal and Placebo Periods

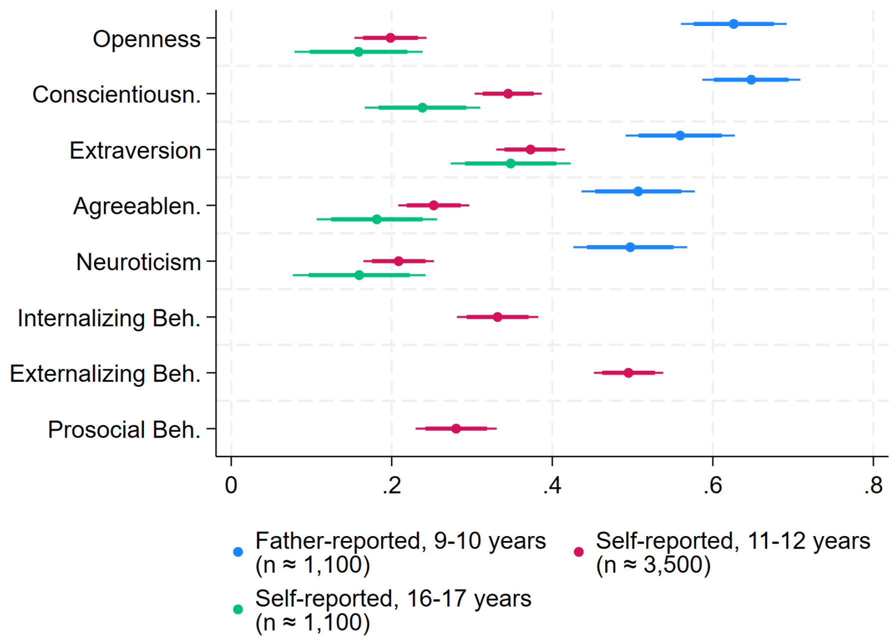
	Neuroticism		Internalizing Behavior	
PM2.5	−.029		−.110	
postnatal period	(.075)		(.070)	
Inversion Frequency		−.010		−.056
postnatal period		(.029)		(.034)
Inversion Frequency		.047*		.069*
in-utero		(.025)		(.037)
Inversion Frequency		.005		−.026
before conception		(.029)		(.035)
Observations	11,242	11,242	7,119	7,119
F-Statistic	109		127	

Notes: Columns 1 and 3 shows 2SLS results from a regression of Neuroticism and internalizing behavior, respectively, on average PM2.5 concentration during the first nine months *after* the month of birth, instrumented with inversion frequency during the same period. Regression controls for county, year, month, and age group fixed effects, weather conditions during the exposure period, child's gender, age in months and its square, parental education, and migration background, as well as inversion frequency and weather conditions during the in-utero period. Columns 2 and 4 show reduced form results where inversion frequency is included for the in-utero period as well as the nine months periods after birth and before conception. Weather controls are also included for all three periods separately. All regressions are weighted by the inverse of the number of times a child is observed in the data. Standard errors clustered at the county level in parentheses. *p<0.1; **p<0.05; ***p<0.01

B Validating the Outcome Variables

Figure B.1 shows correlations of the mother-reported socio-emotional skills with father- and self-reports. Correlations are measured between the mother-reported Big Five assessed at ages 9-10 and (i) father-reported Big Five assessed at the same age range, (ii) self-reported Big Five at ages 11-12 and (iii) self-reported Big Five at ages 16-17, as well as between mother-reported behaviors from the SDQ and self-reported behaviors at age 11-12. All variables are standardized within the respective age-by-respondent cell. All correlations are positive and statistically significant, implying that mother-reports are in line with reports by others and predictive for socio-emotional ability during adolescence. In the main analysis, I use only mother-reported skills as they are available for the largest number of observations.

Figure B.1: Mother-, Father- and Self-reported Socio-emotional Skills



Note: Correlations between mother-assessed and either father- or self-reported socio-emotional skills. Father-reported Big Five are measured with the same scales as mother-reported variables. Self-reported Big Five and Behaviors are based on more items per domain and shorter Likert scales. All variables are standardized. Sample sizes refer to individuals with mother-reported outcomes at age 9-10 for whom data from other reporters are also available. 95%- and 99%-confidence intervals are based on heteroscedasticity robust standard errors.

Figure B.2 shows partial correlations between the mother-reported Big Five and measures of the child's school performance, well-being and preferences, controling for parental education, single-parent households, child gender and migration background. Mother-reported Big Five at age 5-6 are significant predictors of the mother-assessed probability that the child will graduate from the academic track of the German school system when the child is 7 to 8 years old (i.e. before track choice is made). Openness is a strong positive predictor, which is unsurprising, given that it likely partially captures cognitive skills. More importantly, conscientiousness and neuroticism are also significant predictors, with positive and negative sign, respectively. Furthermore, child-reported life satisfaction, self esteem, and risk aversion at ages 11-12 are all correlated with the mother-assessed Big Five at age 9-10, with plausible signs.

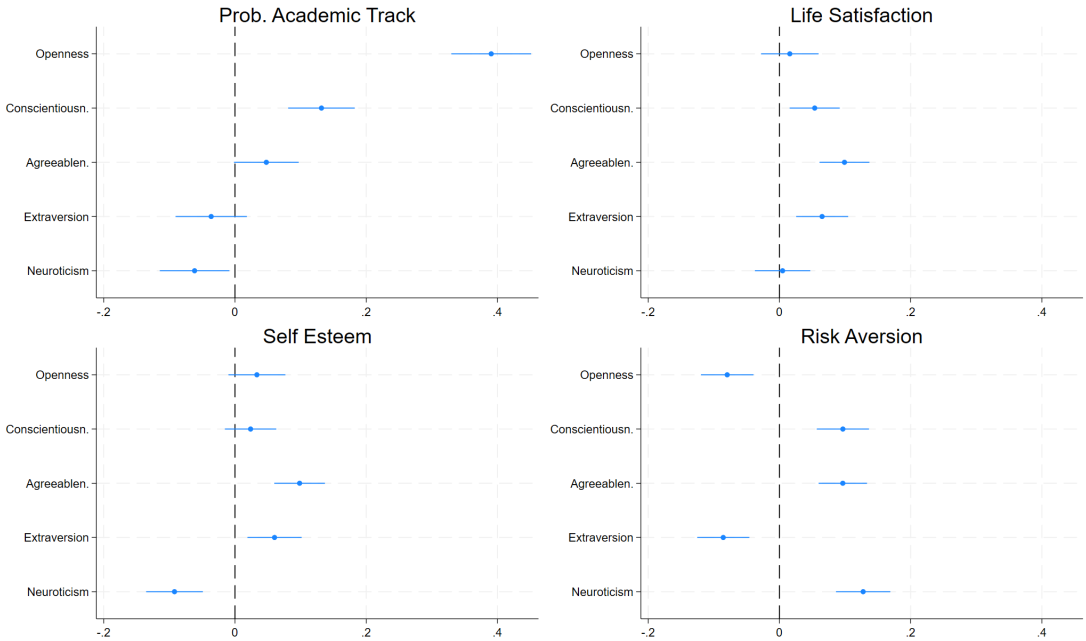


Figure B.2: Mother-reported Big Five and other Child Outcomes

Note: Plots show partial correlations between standardized, mother-assessed Big Five and other child outcomes, after controlling for parental education, a single-parent household dummy, child gender and migration background. Upper left plot: Correlations between the Big Five at age 5-6 years and the mother-assessed probability that the child will graduate from the academic track of the German school system, measured on an 7-point scale when the child is 7-8 years old. Remaining plots depict correlations between the Big Five when the child is aged 9-10 years and standardized child-reported life satisfaction, self esteem, and risk aversion at age 11-12, respectively. Sample sizes range from 3,202 to 4,143. 95%-Confidence Intervalls are based on robust standard errors. Data: SOEP, version 38.

C Robustness Tests

Table C.1 shows that the main results are robust to changes in sample construction and weighting. Column 1 replicates the baseline results. Then, I increase the minimum number of observations for a county to be included in the sample from 15 to either 20 or 30 (columns 2 and 3). In column 4, I show results from unweighted regressions, instead of weighting observations by the inverse of the number of times an individual is observed. In column 5, I keep only one observation per individual. If a child is observed at both 5-6 and 9-10 years, only data from the older age group is used. In column 6, I drop data from 2020 as some of the interviews that year were conducted during Covid19-induced lockdowns.

Table C.1: Robustness: Sample Construction

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Neuroticism						
PM2.5	.135**	.150**	.193***	.121**	.159**	.144**
in utero	(.066)	(.067)	(.075)	(.058)	(.072)	(.065)
Observations	11,242	10,096	8,013	11,242	8,244	10,258
Counties	219	167	102	219	219	205
1st Stage F-Stat.	101	91	70	107	103	101
Panel B: Internalizing Behavior						
PM2.5	.185**	.173**	.186**	.173***	.200***	.214***
in utero	(.074)	(.079)	(.090)	(.066)	(.078)	(.075)
Observations	7,119	6,278	4,606	7,119	5,443	6,320
Counties	160	121	67	160	160	148
1st Stage F-Stat.	131	105	70	130	138	122
Minimum obs. per county	15	20	30	15	15	15
Obs. per individual	all	all	all	all	1	all
Weights	✓	✓	✓	x	x	✓
Years	all	all	all	all	all	w/o 2020

Note: The table displays 2SLS-estimates of the effect of a 1 µg/m^3 increase in prenatal PM2.5 exposure on Neuroticism (Panel A) and internalizing behavior (Panel B). The outcomes are standardized within age groups. Column 1 replicates the baseline result. Columns 2 and 3 vary the minimum number of individuals required within a county. In columns 4 regressions are not weighted by the inverse of the number of observations per individual. In column 5 only one observations for the same individual is included, if multiple are available. In Column 6, observations from the survey year 2020 are dropped. All regressions control for county, year, month, and age group fixed effects, parental education, child gender, migration background, age in months and its square, weather controls (cubic functions of temperature and precipitation, and linear functions of relative humidity and wind speed). Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table C.2 tests robustness to the set of included covariates and fixed effects. In Column 1, I include additional background controls, namely an indicator for single parent households, dummies for maternal age at birth (in 5 year steps), and birth order dummies. In Column 2, I add a lead and lag of the inversion instrument and weather controls to account for potential autocorrelation. In Columns 3 and 4, I vary the included weather controls, and in Columns 5 and 6, I use more stringent fixed effects than in the baseline model.

Table C.2: Robustness: Model Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Neuroticism						
PM2.5	.139*	.170*	.116*	.133**	.149*	.159**
in utero	(.075)	(.090)	(.064)	(.067)	(.082)	(.068)
Observations	10,935	11,242	11,242	11,242	11,242	11,242
Counties	193	219	219	219	219	219
1st Stage F-Stat.	93	55	119	98	65	136
Panel B: Internalizing Behavior						
PM2.5	.182**	.188*	.157**	.201***	.194**	.195**
in utero	(.080)	(.097)	(.068)	(.076)	(.097)	(.080)
Observations	5,869	7,119	7,119	7,119	7,119	7,119
Counties	142	160	160	160	160	160
1st Stage F-Stat.	122	86	173	125	76	163
Specification	background controls	1 lead & lag of IV	cubic weather	linear weather	year-quarter fixed effects	region-month & region-year FEs

Note: The table displays 2SLS-estimates of the effect of a 1 µg/m^3 increase in prenatal PM2.5 exposure on Neuroticism (Panel A) and internalizing behavior (Panel B). The outcomes are standardized within age groups. Relative to the baseline specification, Column 1 adds more background controls (dummies for single parent household, birth order and maternal age bins); Columns 2 adds a lead and lag of the inversion instrument, as well as corresponding leads and lags of the weather controls; Column 3 and 4 vary the included weather variables; Column 5 includes birth year × quarter instead of birth year fixed effects; Column 6 includes East Germany × birth year and East Germany × birth month instead of birth year and birth month fixed effects. All regressions are weighted by the inverse of the number of times a child is observed in the data. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

In Table C.3, I show that the positive impact of particulate matter on neuroticism and internalizing behavior can be replicated with monitor-measured PM10 concentration instead of satellite-assessed PM2.5 concentration as regressor. The fact that the coefficients are smaller than the estimated effects of a unit increase in PM2.5 is unsurprising since PM10 includes less harmful, larger particles.

I also test whether the results for Neuroticism are affected by 'uninformative' survey answers by mothers who have difficulties in assessing their children's non-cognitive skills. This concern arises from the fact that intermediate values are chosen 'too often' when mothers assess their children's personality traits. Table C.4 presents results when dropping observations where mothers picked the intermediate value of 5 for all items underlying conscientiousness and neuroticism, the two outcomes that show most excess mass for the middle value. As an alternative strategy, I drop the 10% or 20% of mothers with the lowest variance of answers across all ten survey items relating to the Big Five.

Respondents with high a capability to memorize and recall relevant information about their child's behavior are likely to deviate more in their answers from the mean across different items.

Table C.3: Robustness: PM10 as Regressor

	Neuroticism	Internalizing Behavior
PM10	.097***	.100*
postnatal period	(.033)	(.051)
Observations	7,810	5,795
Counties	162	139
F-Statistic	87	68

Notes: The table shows 2SLS results from regressions of Neuroticism and internalizing behavior on average, monitor-measured PM10 concentration during the in-utero period, instrumented with inversion frequency during the same period. Regression controls for county, year, month, and age group fixed effects, weather conditions, child's gender, age in months and its square, parental education, and migration background. The regressions are weighted by the inverse of the number of times a child is observed in the data. Standard errors clustered at the county level in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table C.4: Robustness: Uninformative Answers

	Dependent Variable: Neuroticism			
	(1)	(2)	(3)	(4)
PM2.5	.135**	.128*	.153**	.130*
in utero	(.066)	(.067)	(.074)	(.075)
Sample	baseline	drop those with many "fives"	drop bottom 10% (variance)	drop bottom 20% (variance)
Observations	11,242	11,160	9,773	8,030
Counties	219	219	204	177
1st Stage F-Stat.	101	100	92	84

Note: The table displays 2SLS-estimates of the effect of a 1 µg/m^3 increase in prenatal PM2.5 exposure on age-standardized Neuroticism . Column 1 replicates the baseline result. In column 2, observations for which all items underlying Neuroticism and Conscientiousness have a value of 5 are dropped. In columns 3 and 4 the 10% or 20% of observations with the lowest variance across items underlying the Big Five in the SOEP are dropped. All regressions are weighted by the inverse of the number of times a child is observed in the data. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01