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

Beate Thies*

November 2025

Abstract

This paper estimates the effect of in-utero exposure to air pollution on socio-emotional ability in childhood, using data from a representative household survey in Germany and thermal inversions to address endogeneity. 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 by 13% and 18% of a standard deviation, respectively. The effects are driven by increases in the upper parts of the outcome distributions. This implies that affected children are less emotionally stable and suggests adverse impacts on mental health. The effects on emotional stability are more pronounced than impacts on measures of cognitive ability. Back-of-the-envelope computations indicate that a standard deviation increase in air pollution reduces adult earnings by 0.23%-0.74% through its impact on socio-emotional ability. These results provide a better understanding of how in-utero exposure to air pollution generates adverse long-run effects.

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, Ulrich Wagner and several seminar and workshop participants for helpful comments. I gratefully acknowledge financial support by the University of Mannheim's Graduate School of Economic and Social Sciences.

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 (2024) 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 (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, but also inhibit equality of opportunity.

Optimal policy responses might depend on the mechanisms driving the effects of early pollution exposure. As education and labor market outcomes are functions of human capital, its core components – cognitive and socio-emotional skills – are potential mediators. While cognitive ability captures intelligence, socio-emotional or non-cognitive skills comprise a variety of abilities that are weakly correlated with intelligence, such as social competencies, emotional stability and persistence. The evidence regarding these channels is incomplete. While prenatal pollution exposure has been shown to reduce scores in school tests (Bharadwaj et al., 2017; Sanders, 2012) and fluid intelligence (Molina, 2021), i.e. measures of cognitive skills, evidence regarding its effect on socio-emotional abilities is missing. Therefore, this paper investigates whether in-utero exposure to air pollution affects socio-emotional skills, and 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). To address endogeneity 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, 2021). An inversion is a meteorological phenomenon during which air temperature increases with altitude. The warm upper air layer acts like a ceiling preventing ground-level emissions from dispersing.

I find that prenatal air pollution exposure reduces emotional stability; a 1 unit increase in PM2.5 raises the Big Five trait neuroticism by 13% of a standard deviation. In line with this, it also increases the SDQ subscale for internalizing behavior, which captures adverse emotional symptoms and peer relationship problems, 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 emotional stability are economically relevant: For a standard deviation increase in PM2.5, the effect magnitude exceeds the impacts of several other interventions reported in the literature, e.g. a cash transfer or an extension of maternity leave. Since neuroticism and internalizing behavior have been found to negatively affect labor market outcomes, they are plausible channels underlying the long-run effects of early-life pollution exposure. Back-of-the envelope computations imply that a standard deviation increase in PM2.5 reduces earnings by roughly 0.23 – 0.74% via the deterioration in socio-emotional skills. The effect size is of the same order of magnitude as the impact of air pollution on cognitive ability found in earlier work. In this setting, I find no significant effects of prenatal pollution exposure on cognitive skills, which suggests that socio-emotional skills may be more sensitive to pollution than intelligence.

Using quantile regressions, I find that the effects on neuroticism and internalizing behavior are mainly driven by increases in the upper half of the distributions. This is especially concerning, since high levels of neuroticism and internalizing problems have been linked to mental health issues. As mental health disorders pose a huge economic burden – through medical spending and as leading causes of disability – this further underscores the economic relevance of the effects.

The main results remain qualitatively unchanged in a large number of robustness tests, e.g. adding controls for contemporaneous inversions or including sibling fixed effects.

Understanding which mechanisms drive long-term effects of early-life pollution exposure is paramount for designing policy responses. While the predictive power of cognitive and socio-emotional skills for labour market performance is comparable, they differ in how they respond to interventions and investments: Growing evidence shows that socio-emotional skills are malleable until adulthood and can be improved through 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 cognitive skills, the only option to avoid them in future cohorts would be to reduce air pollution. If socio-emotional abilities play a relevant role as well, long-term effects can also be alleviated *ex-post* via interventions targeting these abilities. Given that it can be extremely costly or infeasible to reduce pollution in some circumstances (e.g. pollution from another jurisdiction or natural sources), 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 first causal evidence regarding effects on 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, a setting with relatively low baseline pollution levels, I 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 use data from developing countries (Bharadwaj et al., 2017; Molina, 2021; Rosales-Rueda and Triyana, 2019) or US cohorts born during the 1970s-1980s (Isen et al., 2017; Voorheis, 2017), i.e. settings with higher pollution levels.

Understanding the development of socio-emotional skills is 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., 2022). 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., 2024). 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 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.

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

2. Background

To capture the different dimensions of socio-emotional ability, I rely on the Big Five personality traits and the SDQ at ages 5-10. By school start age, children's character traits exhibit substantial variation and are predictive of 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 (directed outward, linked to aggression and hyperactivity, and tends to create conflict with one's environment), and internalizing behavior (directed inward, leads to distress and includes symptoms like social withdrawal, fearfulness, and anxiety). It is used as a screening tool for behavioural problems and potential mental disorders in several countries.[1] Internalizing behavior has been linked to lower adult earnings, while externalizing behavior can increase earnings (Papageorge et al., 2019).

Personality neuroscience studies how personality traits depend on brain structure, 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, causing oxidative stress and neuroinflammation. Since the in-utero period is a phase of rapid brain growth, gestational exposure to air pollutants might cause irreversible damages by disrupting this process (de Prado Bert et al., 2018), especially Carbon monoxide (CO) and ultrafine particles, which can cross the placenta. Possible mechanisms are maternal systemic and placental oxidative stress and inflammation and impaired transport of oxygen and nutrients to the fetus (Johnson et al., 2021). Maternal cortisol levels might pose another link between air pollution and children's emotional skills. In a randomized experiment, Li et al. (2017) find that acute exposure to PM2.5 raises levels of cortisol and other stress hormones, pointing to activation of the hypothalamus-pituitary-adrenal axis. Cortisol and activity in the hypothalamus-pituitary-adrenal axis in turn show relatively robust associations with neuroticism (Allen and DeYoung, 2017). This is in line with quasi-experimental evidence that maternal stress during pregnancy has adverse long-run impacts on children's mental health (Persson and Rossin-Slater, 2018).

[1] Examples include the UK and Australia where high scores are interpreted as indicators for probable mental disorders.

3. Data

3.1 Child Outcomes

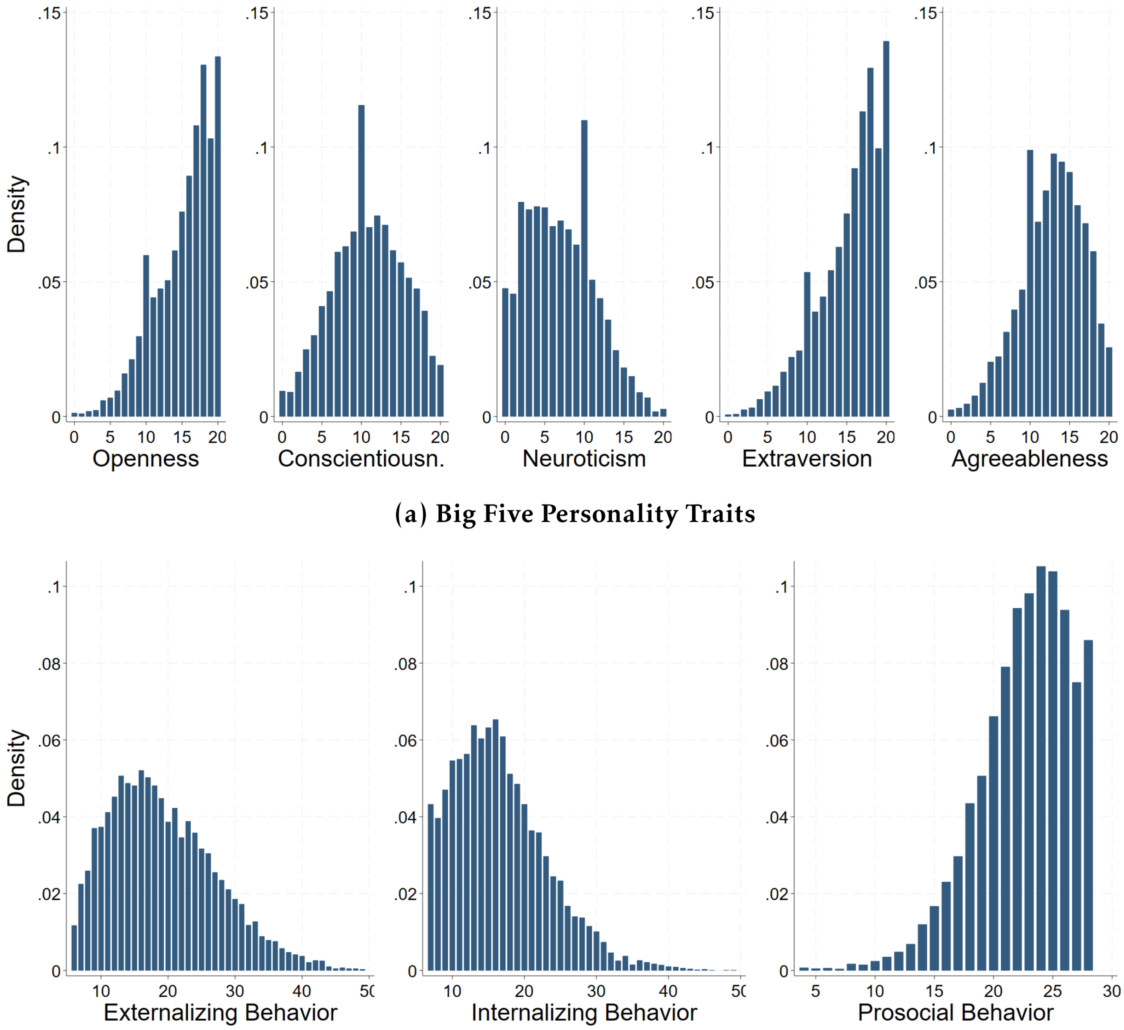
I collect information on children's socio-emotional ability from the German Socio-Economic Panel (SOEP), a representative household panel survey based at the German Institute for Economic Research, which has been running since 1984 (Goebel et al., 2019). To counter panel attrition and increase sample size, several refresher samples have been added since. Moreover, to maintain sample representativeness, adaptations have been made over time to reflect changes in the composition of the resident German population, e.g., including East Germany or adding a boost sample of migrants. Each year, approximately 15,000 households are interviewed.[2]

Socio-emotional skills are assessed at ages 5-6 and 9-10 in the *Mother-and-child* questionnaires. Each Big Five personality trait is assessed with two questions on child behavior, which mothers answer on a scale from 0 to 10 (Appendix Table A.1). Notably, the items underlying openness are likely to capture at least partially cognitive ability (e.g., *My child is quick at learning new things*). The SDQ comprises five domains: emotional symptoms and peer relationship problems – which can be aggregated into *internalizing behavior* –, hyperactivity/inattention and conduct problems – which can be aggregated into *externalizing behavior* – , and prosocial behavior. Mothers answer two to five questions per domain on a scale from one to seven (Appendix Table A.1). While internalizing behavior is linked to the Big Five trait neuroticism, externalizing behavior is associated with high extraversion, low agreeableness, and low conscientiousness (Mezquita et al., 2015). 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). The SDQ scores are available for fewer observations than the Big Five because the "Families in Germany" Study, which was integrated into the SOEP in 2014, included only the Big Five.

Figure 1 shows the outcome distributions (not standardized). All variables show substantial variation, indicating that they pick up differences in child personality and behavior. However, the distributions of the Big Five exhibit spikes at the intermediate value. 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, I still address this potential issue in robustness checks, including sibling fixed effect models and dropping uninformative

[2] For detailed information on the survey see Appendix D.1.

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



(a) Big Five Personality Traits

(b) SDQ subscales

Notes: Possible values for the outcome variables range from 0 to 20 for the Big Five Traits, from 7 to 49 for Externalizing and Internalizing Behavior, and from 4 to 28 for Prosocial Behavior. Distributions are based on all available observations. Panel (a) is based on 13,300 observations, and panel (b) on 9,900 observations.

answers.

For children aged 9-10, mothers also answer two questions about whether the child likes to learn and performs well at school (on a 4-point scale), and report their grades in math and language classes. I combine these variables into a *learning and school performance index* and a *grade index* to capture cognitive ability (Appendix Table A.2). Further, the data include demographic and socioeconomic background variables, e.g. the child's gender and migration history, and parental education[3].

To assign pollution exposure to individuals, I rely on information on year and

[3] I construct dummy variables reflecting different levels of education. For details see Appendix D.1.

Table 1: Summary Statistics

	Big Five				SDQ			
Observations	11,242				7,119			
Unique children	8,253				5,464			
Birth cohorts	2000-2014				2002-2014			
Counties	219				160			
Average County Size	1,015 km^2				988 km^2			
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

Notes: 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.

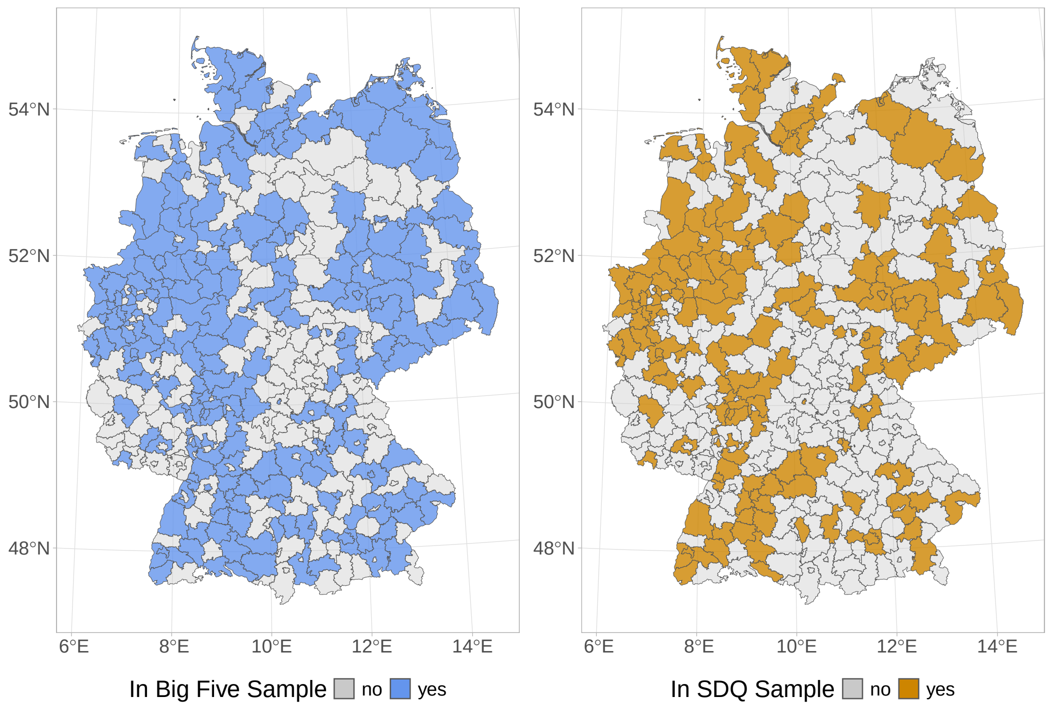
month of birth as well as county of residence – the most granular geographical unit available in the data.[4] For more than two thirds of the children, I can identify the county of residence during the year of birth. In the remaining cases, the households entered the panel after the child was born. As a proxy for county of birth, I assign the county of residence during the first interview. Incorrect assignments can induce measurement error in pollution exposure. Since classical measurement error causes attenuation bias, any results would reflect a lower bound.

I restrict the sample to counties with at least 15 children in the data. Since effects are identified from variation in prenatal PM2.5 concentration between children born in the same county, this serves to reduce the influence of outliers. This yields 11,242 observations of 8,253 individuals for the Big Five, and 7,119 observations of 5,464 individuals for the SDQ. Sample characteristics are displayed in Table 1, and the counties of birth included in the samples are depicted in Figure 2.

Validating the Outcomes. In Appendix B, I demonstrate that the mother-reported Big Five and SDQ contain substantial information about children's socio-emotional skills. I show that the maternal assessments are significantly and positively correlated with paternal assessments at the same age, and with self assessments during adolescence. For a standard deviation increase in mother-reported neuroticism at age 9-10, for example, self-reported neuroticism at age 16-17 increases by 0.16 standard deviations ($p < 0.01$). Moreover, the mother-reported Big Five are significant predictors of the mother-assessed probability that the child will graduate from the

[4] Germany is divided into 401 counties, the administrative division between municipality and federal state.

Figure 2: Counties of Birth



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 interest because very small particles can cross the placenta. Since PM2.5 monitoring only started in 2008 in Germany, I 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. For a small number of 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.

I also collect daily readings of PM10 (particulate matter with a diameter < 10 µm), CO, and NO2 (nitrogen dioxides) from outdoor monitors operated by Germany's federal environmental agency (Umweltbundesamt, 2024). PM10 contains 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 harm 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 are depicted in Table 1.

Thermal Inversions. Under normal conditions, air temperature decreases with altitude. Emissions released at the ground level rise and disperse in the air. A thermal inversion occurs when temperature instead *increases* with altitude. The warm upper air acts like a ceiling that traps emissions close to the ground, raising surface-level pollution concentrations. Following Jans et al. (2018) and Molina (2021), I exclusively consider night-time inversions when constructing the IV to minimize concerns about the exclusion restriction, as they are even less likely to be observable and to induce behavioral responses than day-time inversions. I collect hourly reanalysis data on surface level and upper air temperature for the years 1999-2021 from the European Centre for Medium-Range Weather Forecasts (ECMWF), reported 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 upper air temperature exceeds surface temperature, the county experiences a night-time inversion. The IV 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 2021 from the ECMWF. 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`.

[5] NO2 [CO] concentrations are measured consistently at 252 [64] stations over the sample period.

4. Empirical Strategy

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

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

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. County fixed effects, θ_c, account for persistent differences in pollution and skill levels across locations, month-of-birth fixed effects, θ_m, control for seasonality in air quality, and year-of-birth fixed effects, θ_y, 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 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 cubic 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, i.e. children observed at 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 and parental income, which might be invested into the child's skill development. Moreover, individual pollution exposure is measured with error, causing attenuation bias. Thus, I use an IV strategy based on thermal inversions, following Colmer et al. (2021) and Molina (2021). While inversions exhibit seasonality, 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 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.

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

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

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

Table 2: First Stage Results

	PM2.5	PM10	NO2	CO
Panel A: Big Five Sample				
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 B: SDQ Sample				
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

Notes: The table depicts OLS estimates of the parameter α in Equation (2), reflecting the effects of a standard deviation increase in inversion frequency during the nine-month in-utero period (SD = 0.07) on concentration of different air pollutants during the same period. Outcomes are measured in µg/m^3. 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

The first stage results show that a standard deviation increase in prenatal inversion frequency induces a rise in prenatal PM2.5 concentrations by 0.36 to 0.45 µg/m^3, or 17-21% of a standard deviation (Table 2). The F-Statistics are large, exceeding or just slightly below the threshold of 104 for valid t-ratio inference (Lee et al., 2022). Inversions also significantly increase PM10, NO2 and CO at ground monitors, but for CO, the F-Statistic is very low. With only one instrument, I cannot disentangle the 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 for CO is weak.

To test the credibility of the exogeneity assumption, I show that inversion frequency is uncorrelated with predetermined family characteristics, e.g. parental education or maternal age at birth (Appendix Table A.3). Therefore, my preferred specification controls for seasonality in inversions and composition of births using month-of-birth fixed effects to maximize the identifying variation. In robustness checks, I show that using more stringent fixed effects for month-of-birth × federal state leaves the main results unchanged.

In the following section, I test several versions of Equation (1), raising the probability to incorrectly reject at least one null hypothesis. Thus, in addition to conventional inference, I report sharpened q-values (Benjamini et al., 2006; Anderson, 2008) to control the false discovery rate for all the 38 hypotheses tested in Tables 3 to 6.

5. Results

Table 3 displays 2SLS estimates of the effect of in-utero exposure to PM2.5 on the Big Five personality traits. I find that an increase in PM2.5 by 1 μg/m^3 causes neuroticism to increase by 13.5% of a standard deviation.[6] This implies that children exposed to higher levels of air pollution while in utero are less emotionally stable, i.e. more fearful and less self-confident. I cannot reject that prenatal particulate matter exposure has no effect on the other traits. The largest point estimate emerges for openness, which likely captures cognitive skills. The negative estimate is in line with existing results on the impact of prenatal air pollution exposure on cognitive ability. The fact that it is insignificant and smaller than the effect on neuroticism might imply that emotional stability is more sensitive to air pollution than cognitive skills, which I explore further below. The remaining point estimates are substantially smaller in absolute terms.

Panel B displays 2SLS estimates of the effect of gestational PM2.5 exposure on child behaviors assessed in the SDQ. Internalizing behavior – the outcome related to neuroticism – increases by 18.5% of a standard deviation in response to a unit increase in PM2.5, implying that affected children are shyer, more fearful and have more problems in interacting with peers. For externalizing behavior and prosocial behavior estimates are small and insignificant.[7]

In sum, the results imply that prenatal PM2.5 exposure reduces emotional stability during childhood.[8] Is the magnitude of this effect economically meaningful? In Figure 3a I compare the impact of a standard deviation increase in PM2.5 (2.1 μg/m^3) on neuroticism (+0.28 SD) to the effects of four other factors that have been found to cause significant changes in this outcome during childhood or adolescence. The comparison indicates that the effect of prenatal pollution exposure is of relevant magnitude, exceeding e.g. effects of an extension of maternity leave by three months or an unconditional cash transfer of $3,500.

To approximate the long-run earnings impact of air pollution via the reduction

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

[7]The fact that the SDQ was not administered in all interviews, together with the restriction to counties with at least 15 children, results in the different number of observations and counties between Panels A and B.

[8]For comparison, Appendix Table A.4 depicts OLS estimates.

Table 3: 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)
	[.352]	[.725]	[.884]	[.809]	[.097]
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)
	[.725]	[.089]	[.809]
Observations	7,183	7,119	7,188
Counties	163	160	162
1st Stage F-Stat.	135	131	134

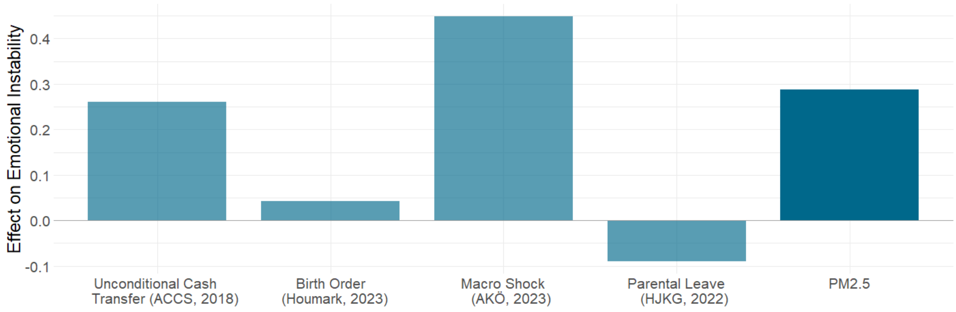
Notes: The Table displays 2SLS-estimates of the parameter β in Equation (1), reflecting the effects of a 1 µg/m^3 increase in PM2.5 concentration during the in-utero period on the Big Five (Panel A) and on the SDQ sub-scores (Panel B). Outcomes are standardized within age groups. Regressions are weighted by the inverse of the number of times a child is observed in the data. Standard errors clustered at the county level in parentheses. */**/*** indicates significance at the 10%/5%/1% level based on conventional p-values. Sharpened q-values in brackets adjust for multiple hypotheses testing.

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 results implies that a standard deviation increase in prenatal PM2.5 concentration reduces annual earnings by .23%-.27% through its effect on neuroticism.[9] Similarly, combining the estimated effect of PM2.5 on internalizing behavior with estimates of its effect on earnings (Papageorge et al., 2019) implies an earnings reduction by .63%-.74%.[10] I account for the fact that my outcomes are measured at an earlier age than in Fletcher (2013) and Papageorge et al. (2019) by scaling the results with the correlation between neuroticism [internalizing behavior] in childhood and at ages 16-17 [ages 11-12] (see Figure B.1). Importantly, the effects via neuroticism and internalizing behavior are not additive, because the variables cap-

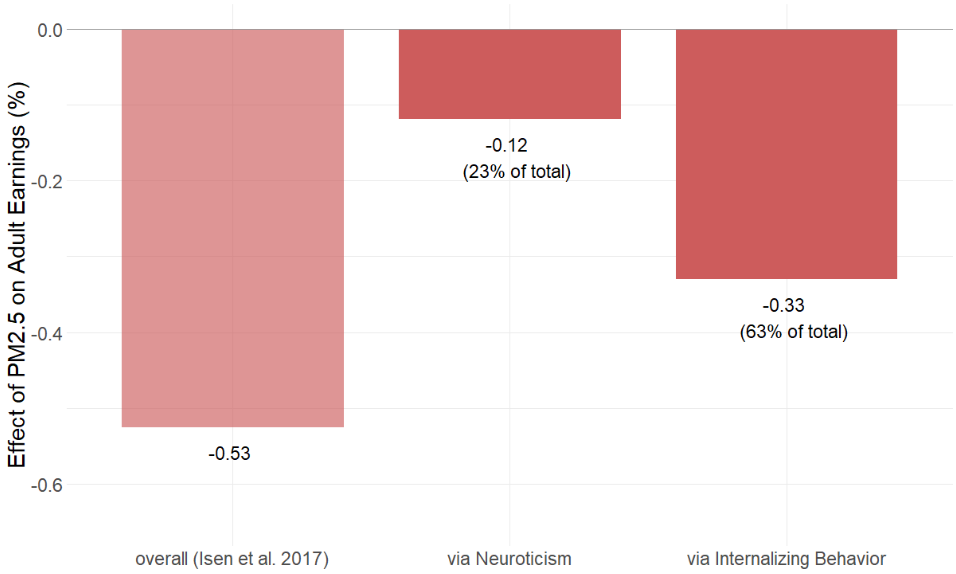
[9] A SD increase in gestational PM2.5 exposure is estimated to raise neuroticism during childhood by 2.1 x .135 = 0.28 SD. To approximate effects on neuroticism in early adulthood, I multiply this with the correlation between neuroticism in childhood and at ages 16-17 (Figure B.1): 0.28 x 0.16 = 0.045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

[10] 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.

Figure 3: Effect Magnitude



(a) Change in Emotional Instability for different Treatments



(b) Implied Earnings Effects of PM2.5

Notes: Panel (a) illustrates the magnitude of the effects of different treatments on emotional instability, measured in standard deviations of the outcome. The effect of PM2.5 on neuroticism is based on Table 3 and refers to a standard deviation increase in prenatal PM2.5. It is compared to the effects of an unconditional cash transfer (of $3,500 on average) (Akee et al., 2018), birth order (Houmark, 2023), and a parental leave extension (by 3.2 months) (Houmark et al., 2024) on neuroticism, and the (inverse) effect of a macro shock on self-confidence (Azmat et al., 2023). Panel (b) shows estimated percentage changes in adult earnings due to a 1 $\mu g/m^3$ increase in prenatal PM2.5 via its effects on neuroticism and internalizing behavior, respectively. Estimates are derived by combining results from Table (3) 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).

ture related characteristics, and are thus correlated ($\hat{\rho} = .5$), but rather provide two estimates of the true earnings effect through emotional instability. For context, Isen et al. (2017) find that a reduction in total suspended particles (TSP, comprising PM10 and larger particles) during the year of birth by 10 $\mu g/m^3$ increases adult earnings by 1%-1.4%. In the US the average TSP-to-PM2.5 ratio is estimated to be 4.38 (Voorheis, 2017). While a comparison across different settings has limitations, it suggests that the socio-emotional ability channel plays a *quantitatively* relevant

Table 4: Comparing Effects on Socio-emotional and Cognitive Ability

	Learning & School Performance Index		Grade Index		Combined Index		Neuroticism		Internalizing Behavior	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
PM2.5	−.063	−.066	.026	.001	−0.024	−0.045	.170*	.228**	.175**	.219**
in utero	(.098)	(.102)	(.089)	(.090)	(0.095)	(0.097)	(.089)	(.093)	(.087)	(.097)
								{.102}**	{.089}**	{.1016}**
	[.713]	[.713]	[.809]	[.884]	[.809]	[.725]	[.097]	[.089]	[.097]	[.089]
Observations	5272	4568	4767	3972	4665	3883	5364	4598	3509	2928
Counties	205	146	198	132	195	130	209	145	157	106
1st Stage F-St.	55	47	58	49	56	46	62	51	96	69

Notes: The table displays 2SLS-estimates of the effect of a 1 µg/m^3 increase in prenatal PM2.5 exposure on measures of cognitive performance and emotional instability. Dependent variables are an index based on two questions about whether the child likes to learn and performs well at school (columns 1 and 2), an index based on school grades in Math and German (columns 3 and 4), an index combining all four aforementioned variables (columns 5 and 6), neuroticism, and internalizing behavior (columns 7 to 10). For details on the cognitive outcomes see Appendix Table A.2. All dependent variables are mesured at age 9-10 and standardized. Cognitive outcomes are coded such that higher values imply better skills. Odd-numbered [even-numbered] columns are based on observations in counties with at least 10 [15] children. Standard errors clustered at the county level in parentheses. tF-adjusted standard errors (Lee et al., 2022) in curly brackets. */**/*** indicates significance at the 10%/5%/1% level based on conventional p-values. Sharpened q-values in squared brackets adjust for multiple hypotheses testing.

role (see Figure 3b).

Comparison to Effects on Cognitive Ability. Next, I compare the effect of prenatal air pollution exposure on neuroticism and internalizing behavior to its impact on cognitive ability. 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 0.024 SD.[11] Sanders (2012) finds that a 10 µg/m^3 increase in TSP exposure during the year of birth reduces high school math test scores by 0.06 SD. While these effects are smaller than my results, the outcomes were assessed at older ages. I also estimate the effect of prenatal pollution on cognitive skills in this setting. Table 4 depicts 2SLS estimates of the effect of prenatal PM2.5 on three proxies of cognitive ability, which are assessed at ages 9-10, standardized, and coded such that higher values imply better skills.[12] While I find a negative point estimates of −0.066 SD for learning and school performance, it is not significant and considerably smaller than the effects on neuroticism and internalizing behavior in both the main sample and in this subsample, i.e. only assessed at ages 9-10 (columns 7-10). Effects on school grades and a combined cognitive ability index are even smaller. While I cannot rule out that the absence of a significant effect on cognition arises because I rely on coarser measures than previous studies,

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

[12]These variables are uncorrelated with neuroticism and only weakly correlated with internalizing behavior (Appendix Figure A.1). As they are only assessed at age 9-10, I report results when including counties with at least 15 children (baseline), and when including counties with at least 10 children to increase sample size.

the items used to measure neuroticism are of similar nature.[13] Importantly, these findings imply that the reduction in emotional stability cannot be explained by air pollution reducing cognitive skills which in turn decreases self-confidence. Instead, they suggest that emotional stability might be more sensitive to air pollution than intelligence. Overall, relative to cognitive skills, socio-emotional ability – specifically emotional stability – seems to be a relevant channel through which gestational pollution exposure generates adverse long-run effects.

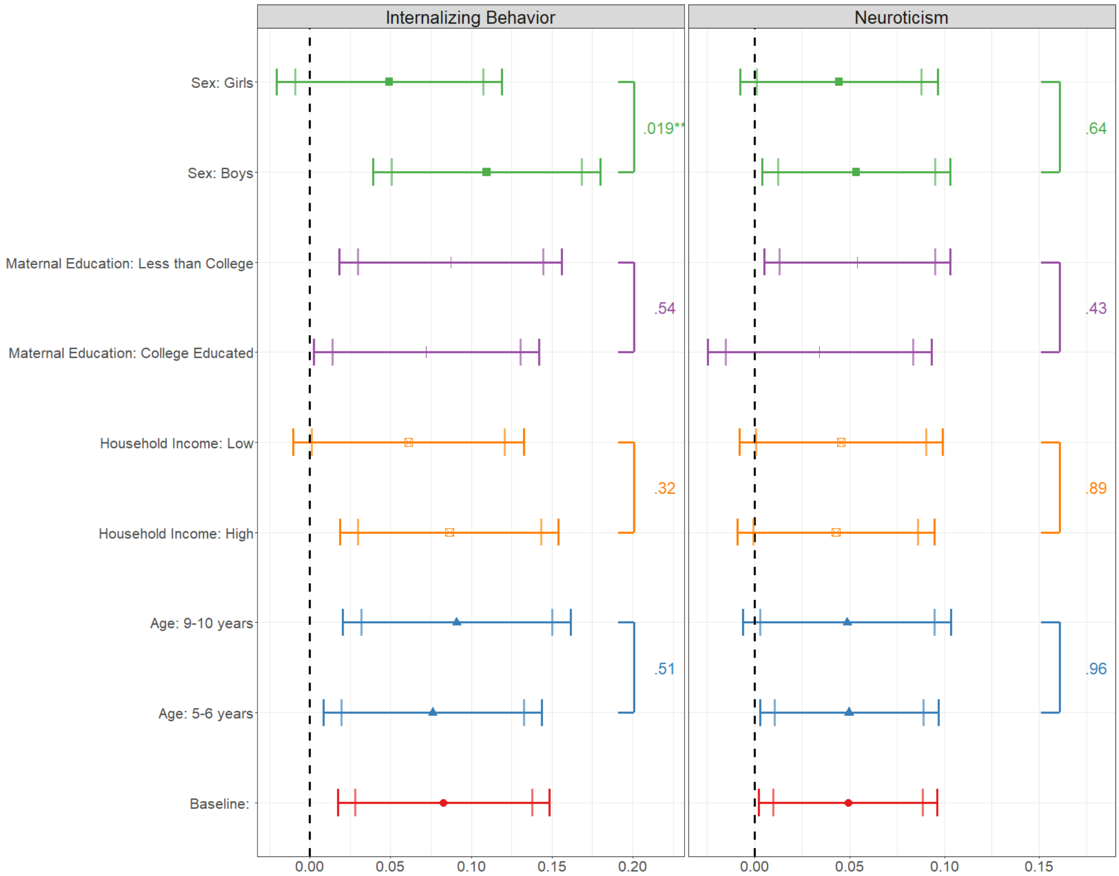
Heterogeneity. I analyze heterogeneity in the effect of in-utero exposure to particulate matter on emotional stability with respect to child gender, age (5-6 vs. 9-10 years), maternal education (college degree vs. less than college), and household income (above vs. below median). To do so, I run the reduced form regression and include an interaction between inversion frequency and the characteristic of interest. Results are illustrated in Figure 4. The plot depicts group specific marginal effects for every subgroup, which are given as the sum of the main effect and, if applicable, the interaction effect. I find that the effect of inversions on internalizing behavior is more than twice as large for boys as compared to girls, but no comparable pattern for neuroticism. There is no significant heterogeneity by age, household income, or maternal education.

I also examine heterogeneity in effect magnitude along the distribution of the outcomes, using conditional quantile regressions. To account for endogeneity, I adopt a two-step control function approach (Lee, 2007), which I describe in detail in appendix D.2. The effects of prenatal PM2.5 exposure on both outcomes are larger for children who, conditional on covariates, have higher scores, i.e. who are conditionally emotionally less stable (Figure 5).[14] For neuroticism, for instance, the point estimate at the 0.9 quantile is almost three times as large as the insignificant estimate at the 0.1 quantile. Appendix Figure A.2 presents additional results from reduced form *unconditional* quantile regressions (Firpo et al., 2009), confirming that the effects of inversions are also strongest in the upper part of the unconditional distribution. The finding that the average effect is driven by shifts in the upper half of the distributions is noteworthy as it implies potential effects on mental health. High scores of the SDQ are widely used as indicators for probable mental disorders. Similarly, in personality psychology, mental disorders are conceptualized as extreme realizations of the Big Five traits. High values of neuroticism show robust

[13] Moreover, I find significant estimates for the effect of average temperature during the in-utero period on the grade index, suggesting that the lack of significance for PM2.5 is not due to the nature of the outcome variables.

[14] For computational reasons, in these estimations I use only one observation per individual, at the oldest age available.

Figure 4: Effect Heterogeneity

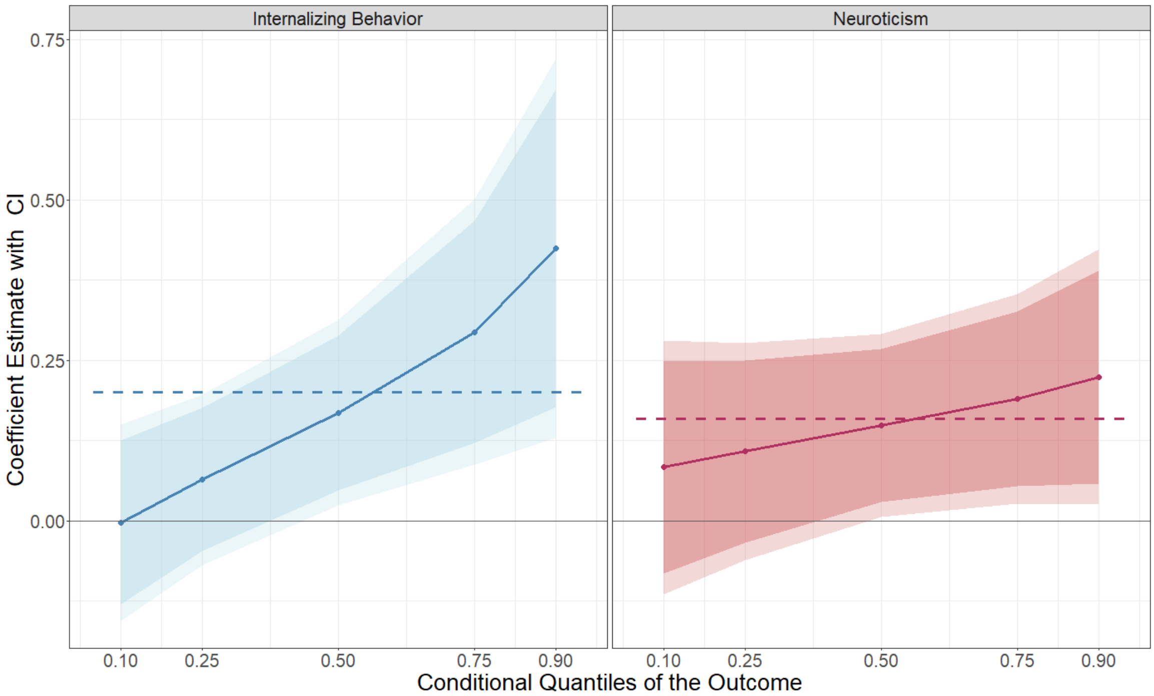


Notes: The plots depict effects of inversion frequency on internalizing behavior and neuroticism for different demographic groups. Estimates in the same color and column are from the same regression. Regressions include inversion frequency and its interaction with an indicator for a specific demographic group, as well as fixed effects and controls as in Equation (1). All estimates depict the marginal effect of a standard deviation increase in inversion frequency in the subgroup described on the left. Error bars: 90%- and 95%-confidence intervals. P-values on the right refer to the test of difference between the marginal effects in the two subgroups.

and consistent correlations with depression and anxiety disorders (Almlund et al., 2011). Hence, prenatal pollution exposure seems to not just reduce an important socio-emotional skill within the range of "normal" variations in personality, but might 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 background characteristics as before and trimester-specific weather controls. I find that the overall effects are driven by the second and third trimesters (Appendix Table A.5), in line with findings on the effect of air pollution on cognitive ability (Molina, 2021) and on performance in school tests (Bharadwaj et al.,

Figure 5: Effects along the Distribution of the Outcomes



Notes: Plots show estimates of the effect of an increase in prenatal PM2.5 concentration by 1 µg/m^3 from conditional quantile regressions using a control function approach, along with 90%- and 95%-confidence intervals. Standard errors are bootstrapped with 500 replications and clustered at the county level. Horizontal dashed lines: 2SLS point estimate of the effect on the conditional mean of the outcome.

2017). Since I lack information on the precise duration of the in-utero period, these patterns are only suggestive.[15]

Additionally, I investigate whether postnatal air pollution exposure affects emotional stability. Appendix Table A.6 shows results from 2SLS and reduced form regressions, which yield insignificant coefficients on PM2.5 concentration and inversion frequency during the first nine months *after* birth for both outcomes of interest. As a placebo test, the table also report results for the effect of pollution and inversions during the nine months *before* conception, respectively. The absence of an effect during this period supports the exogeneity assumption.

Finally, I investigate effects of exposure to inversions in the months preceding the interview. Contemporaneous exposure to air pollution has been found to affect cognitive ability and crime (e.g., Voorheis, 2017). To test whether this is also true for socio-emotional skills, I measure inversion frequency during the last 30 or last 180 days before the interview, respectively, i.e. periods that mothers likely recall well and might base their assessments on. Table 5 presents results from regressions including both prenatal PM2.5 (2SLS) or inversions (reduced form) and contempo-

[15] I discuss this issue in more detail in the robustness section.

Table 5: **Prenatal and Contemporaneous Temperature Inversions**

	Neuroticism				*Internalizing Behavior*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM2.5	.162**		.157**		.202***		0.194**	
in utero	(.073)		(.072)		(.077)		(.077)	
	[.089]		[.089]		[.089]		[.089]	
Inversions		.058**		.056**		.087**		.084**
in-utero		(.026)		(.026)		(.034)		(.034)
		[.089]		[.089]		[.089]		[.089]
Inversions 30 days	.013	.017			.012	.012		
preceding interview	(.017)	(.018)			(.023)	(.024)		
	[.609]	[.431]			[.725]	[.725]		
Inversions 180 days			0.028	.039			-.003	.005
preceding interview			(.024)	(.024)			(.033)	(.034)
			[.334]	[.17]			[.852]	[.845]
Observations	11,116	11,116	11,116	11,116	6,968	6,968	6,968	6,968
1st Stage F-Stat	92		95		122		124	

Notes: The Table presents estimated effects of inversion-induced air pollution during the prenatal period and during the months preceding the interview on emotional stability. Coefficients show effects of a unit increase in PM2.5 concentration and of a standard deviation increase in inversion frequency, respectively. Columns (1), (3), (5) and (7) show 2SLS estimates, while columns (2), (4), (6) and (8) show reduced-form OLS estimates. All specifications control for family background variables, weather conditions during the prenatal period and during the relevant period before the interview date; fixed effects for year, month and county of birth as well as interview year, interview month and county of residence during the interview (see Appendix D.3 for details). 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. */**/*** indicates significance at the 10%/5%/1% level based on conventional p-values. Sharpened q-values in brackets adjust for multiple hypotheses testing.

raneous inversions. While the estimates for the prenatal period are similar to the baseline, I find no evidence for effects of inversion-induced pollution prior to the interview.[16]

Robustness. Results of robustness tests are presented in Appendix C. While I restrict the main sample to counties with at least 15 children, the effects of prenatal PM2.5 exposure on neuroticism and internalizing behavior are robust to varying the minimum number of observations per county (Table C.1).

In Table C.2, I show that results are robust to dropping the weights given by the inverse of the number of times an individual is observed, or using only one observation per individual. Results are also robust to dropping data from 2020 (Covid-19 pandemic), and to dropping individuals with imputed county of birth.

In Table C.3, I estimate the models for the Big Five on the subsample of children for whom the SDQ is available, to check whether results are affected by differences

[16]Appendix Table A.8 further adds interactions between prenatal and contemporaneous exposure, Table A.7 shows effects of contemporaneous inversion frequency on contemporaneous PM10, and Table A.9 confirms the absence of effects of contemporaneous inversions on emotional stability across various specifications.

in sample and county composition. The only notable change to the full sample is that the effect on openness turns significant at the 10% level and larger in magnitude, but it remains smaller than the estimate for neuroticism.

Table C.4 shows that results are robust to (i) adding controls for single parent households, maternal age at birth, and birth order, (ii) adding a lead and lag of the instrument and weather controls to account for potential autocorrelation, (iii) varying the included weather controls, and (iv) using more stringent fixed effects.

Results are also robust to using measures of exposure to high levels of particulate matter instead of average concentrations (Table C.5). Effects of maximum monthly PM2.5 during pregnancy are smaller than effects of average prenatal PM2.5 concentration, suggesting that both exposure to high concentrations and changes in air quality at lower levels contribute to the overall effect. The main results are also robust to alternative methods for assigning PM2.5 to counties, specifically assigning PM2.5 concentration in the grid cell overlapping the county population center rather than the mean across all cells covering the county (Table C.6).

The PM2.5 reanalysis data is constructed using a chemical transport model that incorporates the planetary boundary layer height – which is related to temperature inversions. This might give rise to a mechanical relationship between the IV and the endogenous regressor, leading to overestimation of instrument relevance. In Table C.7, I check that the impact of air pollution on emotional instability is robust to replacing PM2.5 with monitor-measured PM10 or NO2 concentrations, and the F-Statistics remain high (68 to 87).[17] Additionally, I find positive effects on neuroticism and internalizing behavior with an alternative instrument based on wind direction, albeit somewhat smaller in magnitude (Tables C.8 and C.9).

Because the sample includes individuals with imputed place of birth, I abstain from using survey weights, which might give disproportionate influence to observations with incorrectly assigned counties, leading to attenuation bias. In Table C.10, I show that the main findings remain qualitatively unchanged with survey weights, after dropping individuals with imputed county of birth, even though results lose significance, likely due to the drop in sample size.[18]

The attribution of prenatal pollution exposure is based on the month of birth, as information on the exact pregnancy period is unavailable. Since gestational age might be endogenous to air pollution, this could lead to measurement error in prenatal PM2.5 exposure, especially in exposure during the first trimester. Using data

[17]While NO2 is likely not a main driver of the effects of prenatal pollution on child development, it is a co-pollutant of PM2.5, and measured at a larger number of monitors than PM10.

[18]The drop in sample size is even larger than in the robustness check dropping observations with unknown place of birth, because survey weights are zero for individuals who did not participate in any intermediate interview wave (*temporary dropouts*).

Table 6: Models with Family Fixed Effects

	Neuroticism		Internalizing Behavior	
PM2.5	.145		.317***	
in-utero	(.096)		(.121)	
	[.196]		[.089]	
Inversion Frequency		.053		.125**
in-utero		(.043)		(.057)
		[.322]		[.089]
Observations	8,045	8,045	5,191	5,191
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. Standard errors clustered at the county level in parentheses. */**/*** indicates significance at the 10%/5%/1% level based on conventional p-values. Sharpened q-values in brackets adjust for multiple hypotheses testing.

on gestational age from survey modules answered by mothers of newborns, I find suggestive evidence that prenatal air pollution indeed leads to a higher likelihood of preterm birth (Table C.11.) To test whether this causes bias in the estimated effects on emotional stability, I re-estimate the main models on the subsample of children with data on gestational age and socio-emotional skills, and check how estimates change when dropping cases of pre-term birth (Table C.12). I find no systematic changes in the 2SLS estimates for prenatal PM2.5 concentration. For trimester level inversion frequency, the estimates are overall less precise in this small subsample, but the pattern of larger point estimates in later trimesters is qualitatively unchanged.

Moreover, I show that results are robust to clustering standard errors at a higher administrative level to account for potential spatial correlation in the error terms (Table C.13), and to dropping "uninformative" answers in the Big Five sample (Table C.14).

Finally, I incorporate the IV approach into a model with family fixed effects. This accounts for potential differences in mothers' subjective assessment of child behavior by exploiting only variation in outcomes and inversion-induced air pollution between children of the same mother. For both neuroticism and internalizing behavior, the model yields estimates that are positive and larger than in the baseline model, but the effect on neuroticism is not statistically significant (Table 6). This is not surprising given the reduced identifying variation. However, finding positive and sizable estimates alleviates concerns about systematic differences in reporting standards across mothers.

6. Conclusion

Exploiting quasi-random variation in thermal inversions, I show that in-utero exposure to air pollution reduces children's emotional stability, an important component of their socio-emotional skills. These effects exceed the impacts of gestational pollution exposure on cognitive skills in this setting, and are a plausible channel contributing to long-run earnings effects of air pollution. Given that I study a setting with moderate pollution levels, this channel will likely remain relevant despite overall improvements in air quality in many countries.

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 like wild fires, could at least in part be compensated 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 raise the probability of developing mental health issues. A more comprehensive analysis of this relationship 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). 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). Childhood circumstances and adult outcomes: Act ii. *Journal of Economic Literature 56*(4), 1360–1446.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association 103*(484), 1481–1495.

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. *IZA Discussion Paper*.

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

Benjamini, Y., A. M. Krieger, and D. Yekutieli (2006). Adaptive linear step-up procedures that control the false discovery rate. *Biometrika 93*(3), 491–507.

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 (2024). The Grandkids Aren't Alright: The Intergenerational Effects of Prenatal Pollution Exposure. CEP Discussion Paper.

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). 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 (2022). The rising return to noncognitive skill. *American Economic Journal: Applied Economics 14*(2), 78–100.

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.

Firpo, S., N. M. Fortin, and T. Lemieux (2009). Unconditional quantile regressions. *Econometrica 77*(3), 953–973.

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.

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*.

Houmark, M. A., C. M. L. Jørgensen, I. L. Kristiansen, and M. Gensowski (2024). Effects of extending paid parental leave on children's socio-emotional skills and well-being in adolescence. *European Economic Review 170*, 104880.

Hufe, P. (2024). The Parental Wage Gap and the Development of Socio-Emotional Skills in Children. *CESifo Working Paper no. 11111*.

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.

Johnson, N. M., A. R. Hoffmann, J. C. Behlen, C. Lau, D. Pendleton, N. Harvey, R. Shore, Y. Li, J. Chen, Y. Tian, et al. (2021). Air pollution and children's health—a review of adverse effects associated with prenatal exposure from fine to ultrafine particulate matter. *Environmental health and preventive medicine 26*, 1–29.

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). Valid t-ratio inference for iv. *American Economic Review 112*(10), 3260–90.

Lee, S. (2007). Endogeneity in quantile regression models: A control function approach. *Journal of Econometrics 141*(2), 1131–1158.

Li, H., J. Cai, R. Chen, Z. Zhao, Z. Ying, L. Wang, J. Chen, K. Hao, P. L. Kinney, H. Chen, et al. (2017). Particulate matter exposure and stress hormone levels: a randomized, double-blind, crossover trial of air purification. *Circulation 136*(7), 618–627.

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. Nber working paper.

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). The Causal Impact of Socio-Emotional Skills Training on Educational Success. *The Review of Economic Studies*.

Umweltbundesamt (2024). Luftdaten. 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. 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

<table>
<tr><td colspan="3">Panel A: Big Five</td></tr>
<tr><td colspan="2">Dimension</td><td>How would you rank your child in comparison to other children of the same age? My child is ...</td></tr>
<tr><td colspan="2">Openness</td><td>not that interested -– hungry for knowledge
understands quickly — needs more time</td></tr>
<tr><td colspan="2">Conscientiousness</td><td>focused – easily distracted
tidy -– untidy</td></tr>
<tr><td colspan="2">Extraversion</td><td>talkative -– quiet
withdrawn -– sociable</td></tr>
<tr><td colspan="2">Agreeableness</td><td>obstinate – compliant
good-natured — irritable</td></tr>
<tr><td colspan="2">Neuroticism</td><td>self-confident -– insecure
fearful — fearless</td></tr>
<tr><td colspan="3">Panel B: Strength and Difficulties Questionnaire</td></tr>
<tr><td>Aggregate Scale</td><td>Subscale</td><td>To what extent do the following statements apply to your child?</td></tr>
<tr><td rowspan="2">Internalizing Behavior</td><td>Emotional Problems</td><td>is often unhappy or dejected
is nervous/clingy in new situations, loses self-confidence easily
has many fears, becomes frightened easily</td></tr>
<tr><td>Peer Problems</td><td>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</td></tr>
<tr><td rowspan="2">Externalizing Behavior</td><td>Hyperactivity</td><td>is agitated, hyperactive, cannot sit still
is fidgety
is easily distracted and lacks concentration
finishes tasks, is able to concentrate
thinks before acting</td></tr>
<tr><td>Conduct Problems</td><td>often has tantrums, has a temper
quarrels a lot with other children, picks on them</td></tr>
<tr><td>Prosocial Behavior</td><td>Prosocial Behavior</td><td>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</td></tr>
</table>

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: Cognitive Skills in the SOEP Mother-and-child-questionnaires

Outcomes		Items	Scales
Combined Index	Learning & School Performance Index	Child keeps up well with its lessons.	1 (fully agree) to 4 (disagree)
		Child enjoys learning.	1 (fully agree) to 4 (disagree)
	Grade Index	Math Grade	1 (best) to 6 (worst)
		Language Grade	1 (best) to 6 (worst)

Note: Items in the SOEP used to measure children's cognitive ability. Questions are answered by mothers of children aged 9 to 10, i.e. in fourth grade of elementary school. Indices are constructed by adding up individual items and standardizing. Items are recoded such that higher values of the indices reflect higher cognitive ability.

Table A.3: 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 B: 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

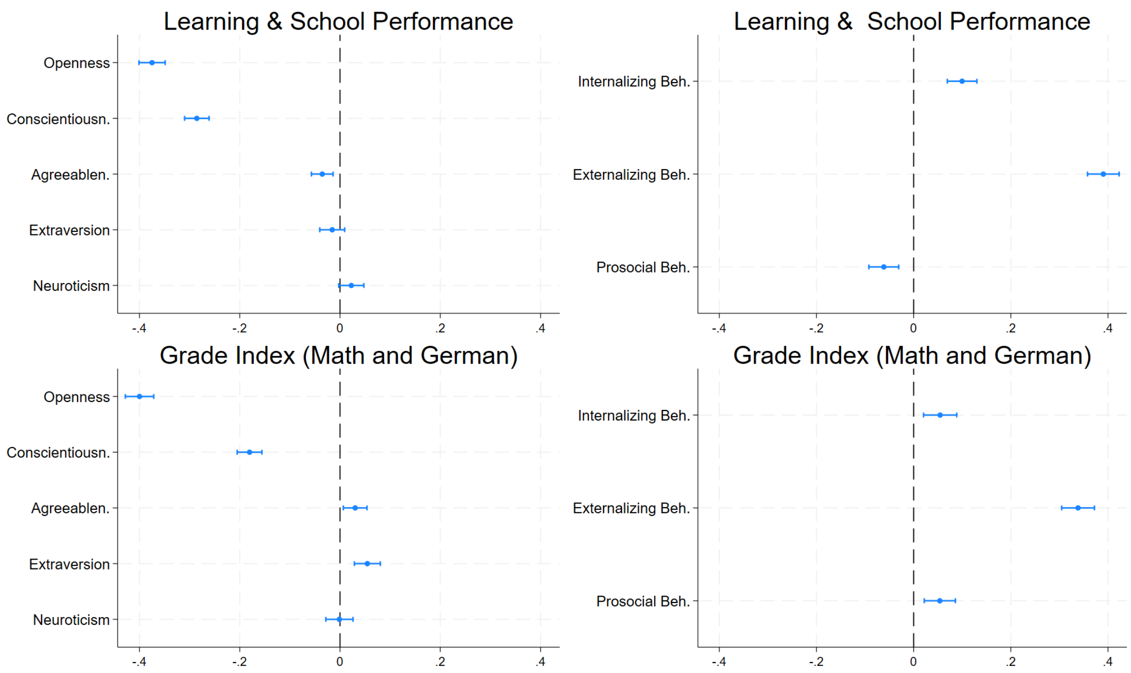
Table A.4: OLS estimates: Prenatal PM2.5 and Childhood Socio-emotional Skills

Panel A: Big Five	Openness	Conscientious	Extraversion	Agreeable	Neuroticism
PM2.5	−.010	.024*	−.031**	.010	.034**
in-utero	(.013)	(.013)	(.013)	(.013)	(.015)
Observations	11,243	11,243	11,239	11,193	11,242
Counties	219	219	219	218	219

Panel B: SDQ	Externalizing Behavior	Internalizing Behavior	Prosocial Behavior
PM2.5	−.003	.037**	−.003
in-utero	(.016)	(.016)	(.019)
Observations	7,183	7,119	7,188
Counties	163	160	162

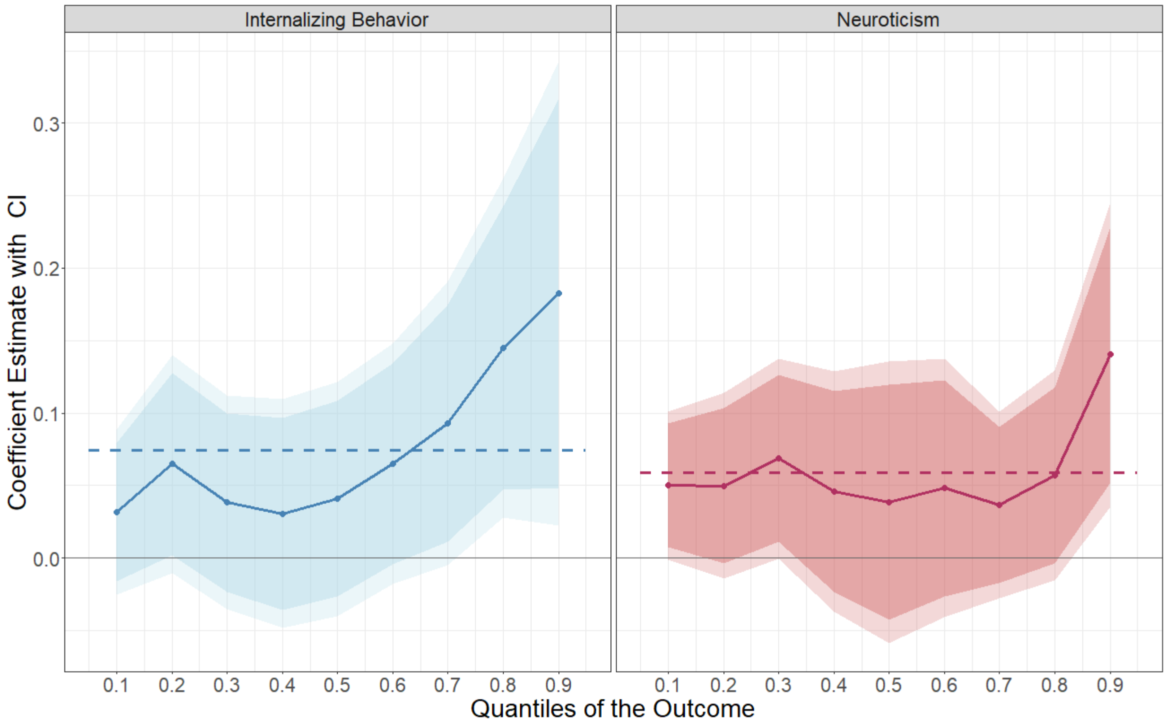
Note: The table displays OLS-estimates of the parameter β in Equation (1), reflecting the effects of a 1 $\mu g/m^3$ increase in PM2.5 concentration during the in-utero period on the Big Five (Panel A) and on externalizing, internalizing, and prosocial behavior based on the Strength and Difficulties Questionnaire (Panel B). 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

Figure A.1: Partial Correlations of Socio-Emotional Skills and Measures of Cognitive Ability



Note: Plots show partial correlations between standardized, mother-assessed socio-emotional skills and standardized measures of the child's cognitive ability, after controlling for parental education, a single-parent household dummy, child gender and migration background. Upper plots: Correlations between the Big Five (left) or behaviors from the SDQ (right) and an index combining mother reports on whether the child likes to learn and performs well at school (both items are answered on a 4-point scale). Bottom plots: Correlations between the Big Five (left) or behaviors from the SDQ (right) and an index combining the child's school grades in Math and German classes. All variables are assessed at age 9-10. Higher values indicate *lower* cognitive skills (unlike in Table 2). 95%-Confidence Intervalls are based on robust standard errors. Data: SOEP, version 38.

Figure A.2: Effects of Inversion Frequency along the Unconditional Distribution of the Outcomes



Note: Plots show estimates of the effect of a standard deviation increase in prenatal inversion frequency from unconditional quantile regressions (Firpo et al., 2009), along with 90%- and 95%-confidence intervals. Standard errors are bootstrapped with 500 replications and clustered at the county level. Horizontal dashed lines: OLS point estimate of the average effect on the respective outcome.

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.

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

	Neuroticism			Internalizing Behavior		
	(1)	(2)	(3)	(4)	(5)	(6)
PM2.5	−.029			−.110		
postnatal period	(.075)			(.070)		
PM2.5		0.057			.003	
before conception		(.080)			(.077)	
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	11,242	7,119	7,119	7,119
F-Statistic	109	57		127	96	

Notes: Columns 1 and 4 show 2SLS estimates of the effect of average PM2.5 concentration during the first nine months *after* the month of birth, instrumented with inversion frequency during the same period. Controls include the same fixed effects and background characteristics as the main model, plus weather conditions during the exposure period, as well as inversion frequency and weather conditions during the in-utero period. Columns 2 and 5 show 2SLS estimates of the effect of average PM2.5 concentration during the first nine months *before conception*, instrumented with inversion frequency during the same period. Controls include the same fixed effects and background characteristics as the main model, plus weather conditions during the exposure period and inversion frequency and weather conditions during the pre-conception period. Columns 3 and 6 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

Table A.7: Effects of Contemporaneous Inversions on Contemporaneous PM10

	PM_{10} 30 days preced. interv.	PM_{10} 180 days preced. interv.	PM_{10} 30 days preced. interv.	PM_{10} 180 days preced. interv.
Inversions 30 days preceding interview	1.58*** (.093)		1.93*** (.103)	
Inversions 180 days preceding interview		0.41*** (.095)		0.41*** (.100)
F-Statistic	291	18.4	349	16.6
Observations	10,415	10,267	6,528	6,446

Notes: The table shows OLS estimates of the effect of a standard deviation increase in contemporaneous inversion frequency on contemporaneous monitor-measured PM10 concentrations. All specifications control for family background variables, weather conditions during the prenatal period and during the relevant period before the interview date; fixed effects for year, month and county of birth as well as interview year, interview month and county of residence during the interview. 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 depicted in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.8: In-utero and Contemporaneous Pollution: Interaction Effects

	Neuroticism		*Internalizing Behavior*	
	(1)	(2)	(3)	(4)
Inversions in-utero	.058** (.026)	.056** (.026)	.087** (.034)	.084** (.034)
Inversions 30 days preceding interview	.018 (.018)		.012 (.024)	
Inversions 180 days preceding interview		.039 (.025)		.005 (.034)
Inversions in-utero × Inv. 30 days pre. interview	-.010 (.012)		.003 (.014)	
Inversions in-utero × Inv. 180 days pre. interview		-.005 (.012)		-.002 (.016)
Observations	11,116	11,116	6,968	6,968

Notes: The Table shows estimated effects of thermal inversions during the prenatal period and during the months preceding the interview on emotional stability, as well as their interaction. Coefficients show OLS estimates of the effects of a standard deviation increase in inversion frequency. All specifications control for family background variables, weather conditions during the prenatal period and during the relevant period before the interview date; fixed effects for year, month and county of birth as well as interview year, interview month and county of residence during the interview. 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 depicted in parentheses. $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A.9: Effects of Contemporaneous Inversions on Emotional Stability

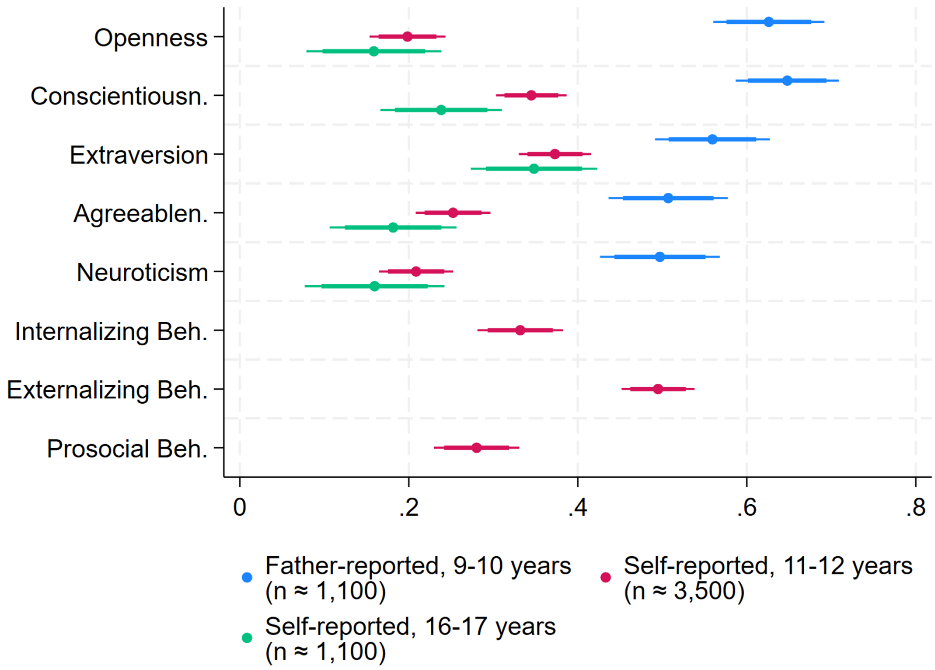
	Neuroticism			*Internalizing Beh.*		
	(1)	(2)	(3)	(4)	(5)	(6)
Inversions 30 days preceding interview	.0099 (.016)		0.020 (.030)	.004 (.019)		.007 (.039)
Inversions 180 days preceding interview		.025 (.023)			.008 (.029)	
Inversion on interview date	-.006 (.021)	-.005 (.021)	.047 (.043)	.0025 (.030)	.004 (.029)	.044 (.061)
Inversions 30 days following interview	.010 (.016)	.016 (.016)	.030 (.040)	-.032 (.021)	-.026 (.020)	.005 (.039)
County FE	✓	✓		✓	✓	
Interview Month FE	✓	✓	✓	✓	✓	✓
Interview Year FE	✓	✓	✓	✓	✓	✓
Child FE			✓			✓
Observations	11,878	11,878	6,698	7,791	7,791	4,160

Notes: The table shows OLS estimates of the effects of temperature inversions shortly before the interview on Neuroticism and Internalizing Behavior. Coefficients show effects of a standard deviation increase in inversion frequency during the 30 days or last 180 days preceding the interview, and the first 30 days following the interview, respectively. "Inversion on interview date" is a dummy variable that takes value one if a nighttime inversion occurred on the day of the interview. All specifications control for family background variables, weather conditions during the relevant periods, and fixed effects as indicated in the table. Columns (3) and (6) are estimated only on the subsample of children observed at both ages 5-6 and 9-10. Standard errors clustered at the county level are depicted 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, controlling 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 (i.e.

highest) 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.

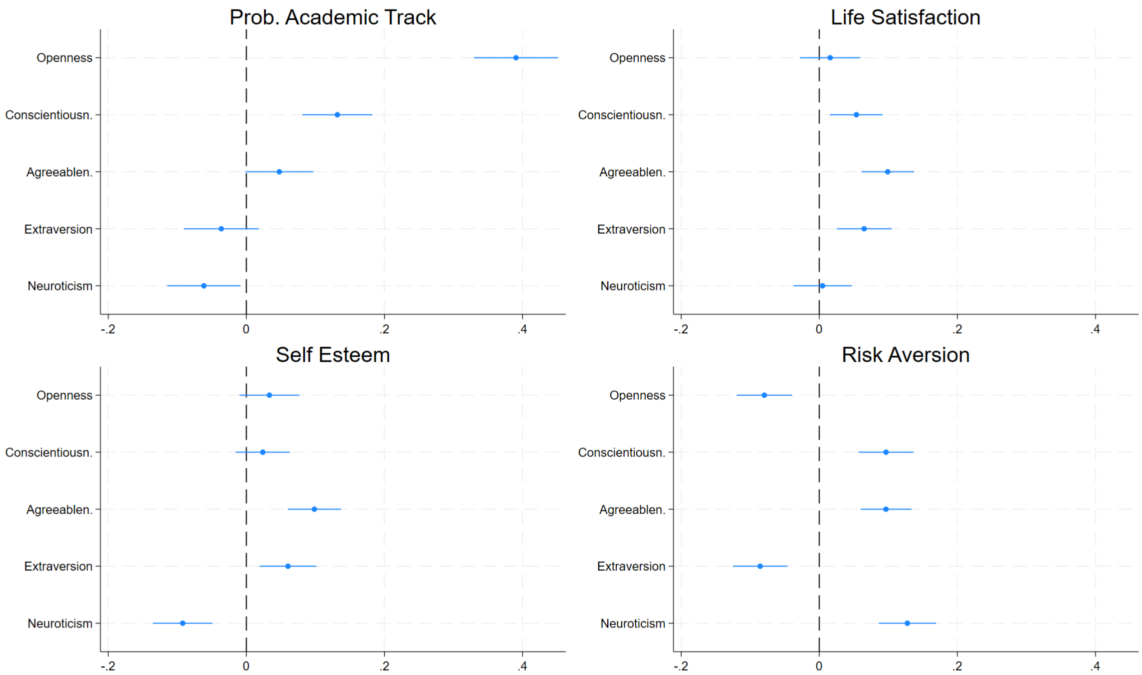


Figure B.2: Mother-reported Big Five and other Child Outcomes

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: Robustness: Minimum Number of Individuals per County

	(1)	(2)	(3)	(4)	(5)
Panel A: Neuroticism					
PM2.5	.135**	.150**	.193***	.129**	.115*
in utero	(.066)	(.067)	(.075)	(.064)	(.063)
Observations	11,242	10,096	8,013	12,191	12,982
Counties	219	167	102	279	395
1st Stage F-Stat.	101	91	70	115	120
Panel B: Internalizing Behavior					
PM2.5	.185**	.173**	.186**	.167**	.140**
in utero	(.074)	(.079)	(.090)	(.072)	(.068)
Observations	7,119	6,278	4,606	8,278	9,207
Counties	160	121	67	237	388
1st Stage F-Stat.	131	105	70	154	171
Minimum obs. per county	15	20	30	10	1

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. In Columns 2 to 5, the minimum number of individuals required for a county to be included in the analysis is varied. 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: Robustness: Weights and Sample Construction

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Neuroticism						
PM2.5	.135**	.121**	.159**	.144**	.194**	0.178**
in utero	(.066)	(.058)	(.072)	(.065)	(.077)	(.076)
Observations	11,242	11,242	8,244	10,258	6,751	7,816
Counties	219	219	219	205	151	218
1st Stage F-Stat.	101	107	103	101	103	119
Panel B: Internalizing Behavior						
PM2.5	.185**	.173***	.200***	.214***	.146*	.183**
in utero	(.074)	(.066)	(.078)	(.075)	(.088)	(.085)
Observations	7,119	7,119	5,443	6,320	4,843	5,891
Counties	160	160	160	148	118	185
1st Stage F-Stat.	131	130	138	122	101	125
Obs. per individual	all	all	1	all	all	all
Weights	✓	x	x	✓	✓	✓
Years	all	all	all	w/o 2020	all	all
Individuals with imputed county of birth	✓	✓	✓	✓	x	x
Minimum Obs. per county	15	15	15	15	15	10

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. In columns 2 regressions are not weighted by the inverse of the number of observations per individual. In column 3 only one observations for the same individual is included, if multiple are available. In Column 4, observations from the survey year 2020 are dropped. In Columns 5 and 6, individuals with imputed place of birth 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.3: Results for Big Five on Subsample with SDQ data

	Openness	Consc.	Extrav.	Agreeab.	Neuroticism
PM2.5	-0.128*	-0.017	-0.083	0.015	0.153**
in-utero	(.068)	(.066)	(.063)	(.068)	(.064)
Observations	7161	7163	7135	7149	7160
Counties	161	161	160	161	161
F-Statistic	132	133	130	130	132

Notes: Table depicts 2SLS estimates of the effect of a unit increase in PM2.5 concentration on the Big Five Personality Traits. Models include controls and fixed effects as described in Equation (1). Models are estimated only on the subsample of individuals for whom also the SDQ was administered. 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. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$

Table C.4: Robustness: Model Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Neuroticism								
PM2.5	.139*	.170*	.116*	.113*	.133**	.149*	.159**	.120**
in utero	(.075)	(.090)	(.064)	(0.065)	(.067)	(.082)	(.068)	(.061)
Observations	10,935	11,242	11,242	11,242	11,242	11,242	11,242	11,242
Counties	193	219	219	219	219	219	219	219
1st Stage F-Stat.	93	55	119	121	98	65	136	131
Panel B: Internalizing Behavior								
PM2.5	.182**	.188*	.157**	.172**	.201***	.194**	.195**	.169***
in utero	(.080)	(.097)	(.068)	(0.070)	(.076)	(.097)	(.080)	(.063)
Observations	5,869	7,119	7,119	7,119	7,119	7,119	7,119	7,119
Counties	142	160	160	160	160	160	160	160
1st Stage F-Stat.	122	86	173	175	125	76	163	171
Specification	background controls	1 lead & lag of IV	cubic weather	quadratic weather	linear weather	year-quarter fixed effects	east-month & east-year FEs	state-month FEs

Note: The table displays 2SLS-estimates of the effect of a 1 $\mu g/m^3$ 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); Column 2 adds a lead and lag of the instrument, as well as corresponding leads and lags of the weather controls; Columns 3 to 5 vary the included weather variables; Column 6 includes birth year × quarter instead of birth year fixed effects; Column 7 includes East Germany × birth year and East Germany × birth month instead of birth year and birth month fixed effects; Column 8 includes state × birth month instead of 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. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table C.5: Robustness: Alternative Ways to Measure Air Pollution

	Neurocticism			*Internalizing Behavior*		
	(1)	(2)	(3)	(4)	(5)	(6)
maximum monthly $PM_{2.5}$ concentration	0.049** (.024)			0.063** (.026)		
number of months with $PM_{2.5} > 20\ \frac{\mu g}{m^3}$		0.206** (.101)			0.345** (.150)	
share of days with $PM10 > 40\ \frac{\mu g}{m^3}$			5.80*** (2.03)			6.01* (3.07)
Counties	219	219	162	160	160	139
1st Stage F-stat.	119	63	49	136	37	35
Observations	11,242	11,242	7,810	7,119	7,119	5,795

Notes: The Table displays 2SLS-estimates of the parameter β in Equation (1). Each estimated coefficient is derived from a separate regression. The air pollution measurements listed in the table are instrumented by inversion frequency. The pollution measurements used are the maximum monthly PM2.5 concentration during the in-utero period (row 1); the number of months during the in-utero period with a PM2.5 concentration above 20 $\mu g/m^3$ (row 2); and the share of days during the in-utero period with a PM10 concentration above 40 $\mu g/m^3$ (row 3). 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.$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C.6: Robustness: Air Pollution Assignment to Counties

	Open	Consci.	Extrav.	Agreea.	Neurot.	Externalizing Beh.	Internalizing Beh.	Prosocial Beh.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: PM2.5 measured at county population center								
$PM_{2.5}$ in-utero	-0.077 (.069)	-0.043 (.071)	-0.001 (.070)	0.019 (.069)	0.134** (.065)	0.031 (0.071)	0.182** (0.073)	-0.021 (0.070)
1st Stage F	97	97	96	95	98	129	126	128
Panel B: PM2.5, Weather and Inversions measured at county population center								
$PM_{2.5}$ in-utero	-0.093 (.069)	-0.057 (.071)	0.024 (.072)	0.056 (.068)	0.132** (.067)	-0.009 (.067)	0.165** (.070)	0.023 (.066)
1st Stage F	95	95	94	93	96	137	133	136
Observations	11,243	11,243	11,239	11,193	11,242	7,183	7,119	7,188
Counties	219	219	219	218	219	163	160	162

Notes: The Table displays 2SLS-estimates of the parameter β in Equation (1). In Panel A, instead of average PM2.5 concentration across all grid-cells overlapping the county, PM2.5 concentration measured at the population center of the county of birth is used to construct the main regressor. In Panel B, also weather controls and inversion frequency are measured at the population centroid rather than across the entire county and around the geographic centroid, respectively. 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.$^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table C.7: Robustness: Monitor-measured Pollution

	Neuroticism		Internalizing Behavior	
PM10	.097***		.100*	
in utero	(.033)		(.051)	
NO2		.086**		.104**
in utero		(.039)		(.042)
Observations	7,810	10,590	5,795	6,865
Counties	162	207	139	155
F-Statistic	87	79	68	70

Notes: The table shows 2SLS results from regressions of Neuroticism and internalizing behavior on average, monitor-measured PM10 or NO2 concentration during the in-utero period, respectively, 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.8: Robustness: IV Approach based on Wind Direction

	Openness	*Conscientiousn.*	*Extraversion*	*Agreeableness*	*Neuroticism*
$PM_{2.5}$	−0.035	0.022	−0.067	−0.021	0.104*
in-utero	(.050)	(.048)	(.054)	(.048)	(.061)
1st Stage F-stat.	47	49	48	46	47
Counties	219	219	219	218	219
Observations	11,243	11,243	11,239	11,193	11,242
		Externalizing Beh.	*Internalizing Beh.*	*Prosocial Beh.*	
$PM_{2.5}$		0.000	0.113*	-0.079	
in-utero		(.065)	(.060)	(.068)	
1st Stage F-stat.		55	56	56	
Counties		160	163	162	
Observations		7,182	7,119	7,188	

Notes: The Table displays 2SLS-estimates of the parameter β in Equation (1), reflecting the effects of a 1 µg/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 wind direction. Details on the IV approach are given in Appendix D.4. As instrumental variable, the share of days during the in-utero period with eastern winds (from high pollution areas in Eastern Europe) is used. The first stage relationship between eastern winds and PM2.5 is allowed to vary by the east–west coordinate of the county centroid, and between north and south Germany (see Table C.9). 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

Table C.9: First Stage Results for Wind Direction IV Approach

	PM2.5	*PM2.5*
Share Days with Eastern Wind	.240***	.262***
	(.064)	(.071)
Share Days with Eastern Wind	.039***	.039***
× Easting (100km)	(.011)	(.011)
Share Days with Eastern Wind	−.168	.085
× South	(.164)	(.178)
Share Days with Eastern Wind	−.033	-.051*
× South × Easting (100km)	(.026)	(.030)
Sample	Big Five	SDQ
Observations	11,242	7,119

Notes: The Table displays OLS estimates from a regression of PM2.5 concentration during the in-utero period (in µg/m^3) on wind direction variables used in the IV approach based on wind direction. Details on the IV approach are given in Appendix D.4. Coefficients depict the estimated effect of an increase in the frequency of eastern winds by one standard deviation. *South* denotes a dummy variable that is one if the county of birth is located in Bavaria or Baden-Württemberg. *Easting* refers to the east-west coordinate of the county centroid measured in 100km east of the projection origin under the ETRS89 projected Coordinate Reference System. Controls include cubic weather conditions, individual background variables, and fixed effects for the county, year, and month of birth as well as the age at assessment. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table C.10: Robustness: Regressions with Survey Weights

	Neuroticism		*Internalizing Beh.*	
	(1)	(2)	(3)	(4)
$PM_{2.5}$	.090	.164	.160*	.167
in-utero	(.103)	(.112)	(.091)	(.104)
Counties	196	198	139	158
1st Stage F-stat.	109	97	111	87
Observations	9,496	6,556	5,740	4,703
Individuals with imputed county of birth	✓	x	✓	x
Min. Obs/County	15	10	15	10

Notes: The Table displays 2SLS-estimates of the parameter β in Equation (1), reflecting the effects of a 1 µg/m^3 increase in PM2.5 concentration during the in-utero period on neuroticism and internalizing behavior. PM2.5 is instrumented by inversion frequency. Outcomes are standardized within age groups. In contrast to the main results in Table 3, the regressions are weighted by survey weights. *p<0.1

Table C.11: PM2.5 and Gestational Age

	Preterm birth (1)	Weeks of gestation (2)
PM2.5	0.050*	-0.214
in-utero	(0.027)	(0.200)
Number of observations	7,253	7,253
Dep. Var. Mean	.12	39
1st Stage F-Stat.	115	115
Counties	189	189

Notes: Estimations are based on data from SOEP mother-and-child questionnaires given to mothers of newborns. The table depicts 2SLS estimates of the effect of prenatal PM2.5 levels on an indicator for birth before the 37th week of pregnancy (column 1) and on weeks of gestation (column 2). PM2.5 is instrumented with frequency of thermal inversions. Controls include birth order dummies, dummies for parental education, maternal age at birth, an indicator for a single parent family, an indicator for maternal employment, migration background, child sex, and meteorological conditions as well as month, year and county of birth fixed effects. Observations with missing control variables are included, and separate dummies for missing controls are added. Standard errors are clustered at the county level. * indicates significance at the 10% level.

Table C.12: PM2.5 and Emotional Stability: Dropping Cases of Preterm Birth

	Neuroticism				*Internalizing Behavior*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: 2SLS estimates								
PM2.5	.151	.161	.206**	.231**	.231**	.174*	.201**	.190*
in-utero	(.095)	(.112)	(.100)	(.115)	(.094)	(.102)	(.103)	(.113)
1st Stage F-Stat.	63	49	46	38	60	48	44	37
Panel B: Reduced Form								
Inversion Frequency	-0.004	-0.005	0.025	0.039	0.011	0.016	0.028	0.050
1st trimester	(.035)	(.038)	(.036)	(.039)	(.037)	(.042)	(.044)	(.047)
Inversion Frequency	0.043	0.040	0.052	0.046	0.070*	0.038	0.038	0.007
2nd trimester	(.036)	(.039)	(.041)	(.048)	(.036)	(.040)	(.043)	(.049)
Inversion Frequency	0.050	0.046	0.053	0.066	0.088**	0.078*	0.087*	0.108**
3rd trimester	(.032)	(.037)	(.034)	(.042)	(.041)	(.046)	(.047)	(.050)
Minimum obs./county	10	10	15	15	10	10	15	15
Preterm births dropped		✓		✓		✓		✓
Number of observations	4,412	3,763	3,518	3,013	3,977	3,364	3,108	2,671
Counties	130	114	79	71	127	109	75	68

Notes: Estimates are based on the subsample of observations with information on gestational age available. As indicated at the bottom, individuals born before the 37th week of pregnancy are dropped in some columns. Panel A depicts 2SLS estimates of the effect of prenatal PM2.5 levels on age group-standardized Neuroticism and Internalizing Behavior. PM2.5 is instrumented with frequency of thermal inversions. Models include controls and fixed effects as described in Equation (1). Panel B depicts OLS estimates 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. Regressions include fixed effects and background controls as in Equation (1) and weather controls for each trimester. Standard errors are clustered at the county level. */**/*** indicates significance at the 10%/5%/1% level.

Table C.13: Robustness: Clustering Standard Errors at the District Level

	Neuroticism	Internalizing Behavior
PM2.5	.135**	.185***
in-utero period	(.062)	(.051)
Observations	11,242	7,119
Counties	219	160
Clusters *(Admin. districts)*	48	46

Notes: Note: The Table displays 2SLS-estimates of the parameter β in Equation (1), reflecting the effects of a 1 µg/m^3 increase in PM2.5 concentration during the in-utero period on neuroticism and internalizing behavior. 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. In contrast to the main results in Table 3, standard errors (in parentheses) are clustered at the level of administrative districts, rather than the county level. Districts are an administrative unit between counties and federal states (NUTS2 regions according to the European Union's nomenclature of territorial units for statistics). *p<0.1; **p<0.05; ***p<0.01

Table C.14: 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. This is based on the idea that respondents with high capability to recall information about their child's behavior are likely to deviate more in their answers from the mean across different items (Falk, Neuber, and Strack, 2021). 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

D Additional Information

D.1 Data

The German Socio-Economic Panel. The SOEP is a long-running household panel survey, which is representative of the resident German population, and covers numerous topics such as household composition, occupational biographies, employment, health, and personality traits. The survey is based at the German Institute for Economic Research (DIW Berlin). First started in 1984, the survey was expanded to include East Germany in 1990. To counter panel attrition and increase sample size, several refresher samples have been added since. Moreover, to maintain sample representativeness, adaptations have been made over time to reflect the changes in the composition of the overall population, e.g., a boost sample of migrants was added. Thus, the SOEP is an unbalanced panel. Samples are multi-stage random samples with regional clusters. The households are selected by random-walk routines. The SOEP has been widely used by researchers to study topics such as the impacts of parental wage gaps on child development (Hufe, 2024), the intergenerational transmission of risk preferences (Dohmen, Falk, Huffman, and Sunde, 2011), or the consequences of government surveillance in the GDR (Lichter, Löffler, and Siegloch, 2020).

Children's socio-emotional skills are assessed in the SOEP *Mother-and-child* questionnaires which were introduced in 2008 for children aged 5-6, and in 2012 for children aged 9-10. I use version 38 of the SOEP which includes interview waves up to the year 2020, i.e. I am using information from the Mother-and-child questionnaires adminstered between 2008 and 2020.

Covariates for parental education. I define separate dummies for low, medium and high 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 and define separate dummies for children with missing information on parental education.

Geograhic information. Using the SOEPremote system of remote computer access, it is possible to work with geographic information on household locations at the level of counties. Information on the county of birth is not reported directly for any household member. Whenever possible, I identify mother's county of residence during the child's year of birth in order to assign gestational air pollution exposure to children. For households that entered the survey after the child's year of birth, I identify the county of residence during the first wave the household was interviewed, and use this as a proxy for the child's county of birth.

D.2 Quantile Regressions

For the analysis of effect heterogeneity along the outcome distribution (Figure 5), I follow Lee (2007) who proposes a two-step control function approach. The first step consists of a linear quantile regression of the endogeneous regressor, PM2.5, on the exogenous covariates and the excluded instrument. The residuals from this regression are recovered and a flexible function $g(\hat{\epsilon}(q)_{icyma})$ of these residuals is added to the second stage quantile regression, to capture the endogeneity of air pollution. In the main specification, I use a median quantile regression, and add a third order polynomial of the residuals to the second stage. Results are highly similar when using different quantiles, or a least squares first stage; and when using a third-order B-spline with three knots of the residuals.

Specifically, the models estimated for different quantiles q are given as:

$$\begin{aligned} SocEm_{icyma} =& \beta(q)PM_{cym} + \gamma'(q)\mathbf{X}_i + \delta'(q)\mathbf{W}_{cym} + \\ & \theta_c(q) + \theta_y(q) + \theta_m(q) + \theta_a(q) + g(\hat{\epsilon}(q)_{icyma}) + u(q)_{icyma} \end{aligned}$$

Standard errors are bootstrapped at the county level with 500 replications. The quantile regression with fixed effects is implemented using the `xtqreg` command based on method of moments approach by (Machado and Santos Silva, 2019).

The resulting estimates indicate that individuals who have low emotional stability, given their individual and family characteristics, place and time of birth, are most affected by prenatal air pollution. This is complemented with unconditional quantile regressions (Figure A.2), to analyse effects on the quantiles of the unconditional outcome distribution. Since there is no IV estimator for unconditional quantile effects that is applicable in this setting, I present reduced form results. They are implemented as recentered influence function regressions using the command `rifhdreg` (Firpo et al., 2009).

D.3 Controlling for Contemporaneous Inversions

Maternal assessments might be strongly affected by recent behavior, such that an omitted variable issue could arise if contemporaneous air pollution exposure affects child behavior and inversion frequency in a county is auto-correlated.

To address this, in Table 5 I show results from regression models that include variables measuring contemporaneous inversion frequency in addition to prenatal conditions, specifically the number of days with night-time inversions during the last 30 days or the last 180 days preceding the interview. This is based on the assumption that when answering the survey questionnaire, mothers likely recall important events and child behaviors during these recent periods very well. Specifically the models estimated are of the following form:

$$\begin{aligned}Y_{icymjd} =& \beta_1 PM_{cym}^{in-utero} + \beta_2 Inv_{jd}^{pre-interview} + \boldsymbol{\gamma}_1 \boldsymbol{W}_{cym}^{in-utero} + \boldsymbol{\gamma}_2 \boldsymbol{W}_{jd}^{pre-interview} + \\ & \boldsymbol{\delta X}_i + \theta_y + \theta_m + \theta_c + \theta_j + \theta_{y(d)} + \theta_{m(d)} + \epsilon_{icymjd}\end{aligned}$$

where Y_{icymjd} refers to an outcome observed for individual i born in year y and month m in county c, who is assessed on interview date d in county of residence j. $PM_{cym}^{in-utero}$ is instrumented with inversion frequency during the in-utero period. $Inv_{jd}^{pre-interview}$ is the key variable added to capture contemporaneous exposure to inversion-induced pollution. The model further includes weather conditions during the in-utero period and individual background controls as in model (1), quadratic weather controls for the pre-interview period, and fixed effects for county-, year-, and month-of-birth, as well as year and month of the interview and county of residence during the interview. Data on contemporaneous inversions and contemporaneous weather conditions are collected from the ECMWF.

D.4 IV Approach based on Wind Direction

The alternative 2SLS approach based on wind direction focuses on the broad pattern of inter-regional air pollution transport in Europe: Eastern winds typically carry polluted air from relatively highly polluted Eastern European countries into Germany, whereas the predominant western winds bring cleaner air originating from the Atlantic Ocean and less polluted countries in western Europe. Especially Poland, one of Germany's eastern neighbors, was consistently among the most polluted countries in Europe during the sample period. I construct a variable that measures the share of days during the in-utero period with eastern winds (daily average wind direction between 0° and 180°), i.e. the design exploits time-variation in the frequency of eastern winds, controlling for seasonality and time-invariant local characteristics. To take into account that the effect of eastern winds on local air pollution depends on the proximity to the Eastern European source regions, I interact this variable with the east-west coordinate of the county centroid. Additionally, the relationship between eastern winds, longitude and air pollution is allowed to vary between the north and south of Germany, since – as mentioned above – air originating from Poland in the north-east is expected to carry most pollution. The first stage model is given below.

$$
\begin{aligned}
PM_{cym} = &\alpha_1 Share\ East\ Wind_{cym} + \alpha_2 Share\ East\ Wind_{cym} \times Easting_c + \\
&\alpha_3 Share\ East\ Wind_{cym} \times South_c + \alpha_4 Share\ East\ Wind_{cym} \times Easting_c \times South_c + \\
&\gamma \mathbf{X}_i + \delta \mathbf{W}_{cym} + \theta_c + \theta_y + \theta_m + \theta_a + u_{icyma}
\end{aligned}
$$

REFERENCES (APPENDIX)

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2011, 11). The intergenerational transmission of risk and trust attitudes. *The Review of Economic Studies 79*(2), 645–677.

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

Firpo, S., N. M. Fortin, and T. Lemieux (2009). Unconditional quantile regressions. *Econometrica 77*(3), 953–973.

Hufe, P. (2024). The Parental Wage Gap and the Development of Socio-Emotional Skills in Children. *CESifo Working Paper no. 11111*.

Lee, S. (2007). Endogeneity in quantile regression models: A control function approach. *Journal of Econometrics 141*(2), 1131–1158.

Lichter, A., M. Löffler, and S. Siegloch (2020). The long-term costs of government surveillance: Insights from stasi spying in east germany. *Journal of the European Economic Association 19*(2), 741–789.

Machado, J. A. and J. Santos Silva (2019). Quantiles via moments. *Journal of Econometrics 213*(1), 145–173.