Prenatal Exposure to Air Pollution and the Development of Noncognitive Skills

Beate Thies*

University of Mannheim

beate.thies@uni-mannheim.de

November 10, 2022

click here for most recent version

Abstract

Noncognitive skills are important predictors for important life outcomes like education, health and earnings. This paper provides causal evidence on the effect of in-utero exposure to air pollution on noncognitive abilities in childhood. I use the meteorological phenomenon of thermal inversions to address the endogeneity in exposure to particulate matter, and data from a representative household survey in Germany to measure noncognitive abilities. I find that an increase in particulate matter concentration by one unit during the prenatal period raises neuroticism at age 5-10 by 7% of a standard deviation. This implies that affected children are less emotionally stable, more fearful and less self-confident. Back of the envelope computations indicate that a one standard deviation increase in particulate matter reduces adult earnings by 0.24%-0.29% through its impact on neuroticism alone.

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

JEL-codes: J13, J24, Q53

*I am grateful to Ulrich Wagner, Felix Holub, Vera Huwe and Dana Kassem for helpful feedback and suggestions. I thank seminar audiences at the University of Mannheim as well as participants in the CRC TR 224 Summer School on the Development of Cognitive and Non-Cognitive Skills in Childhood and Adolescence and the 2021 AURÖ Young Researchers' Workshop for their comments.

1. Introduction

Air pollution imposes high costs on society since it adversely affects several dimensions of human health and well-being. Poor air quality is also an obstacle to social mobility: Isen, Rossin-Slater, and Walker (2017) find that exposure to air pollution while in utero and during the first year of life reduces earnings and employment during adulthood. Voorheis (2017) confirms this result and finds that educational outcomes are negatively affected as well, while recent evidence by Colmer and Voorheis (2020) suggests that these adverse impacts even extend to the next generation. Since low-income families and minorities often live in more polluted neighbourhoods than more affluent groups (see e.g. Banzhaf et al., 2019; Currie et al., 2020; Glatter-Götz et al., 2019; Rüttenauer, 2018, for evidence on the US and Europe), these long-run effects of gestational pollution exposure not only impose a substantial economic cost on society, but also inhibit equality of opportunity.

Optimal policy responses to this issue might depend on the mechanisms driving the adverse long-run effects of early pollution exposure. Educational achievement and and labor market success are functions of human capital, whose core components are cognitive and noncognitive skills.[1] While the predictive power of the two types of skills for educational attainment and labour market performance is comparable, they differ crucially in how they respond to intervention programs and investments: There is growing evidence that noncognitive skills are malleable up until adulthood and can be improved by way of low-cost interventions implemented in the classroom- or even work-environment (Adhvaryu et al., 2018; Alan et al., 2019; Sorrenti et al., 2020), whereas cognition is less malleable, especially after school start age (e.g. Almlund et al., 2011; Cunha et al., 2010). Hence, understanding which mechanisms drive adverse long-term effects of early-life pollution exposure, and how important the respective channels are in quantitative terms, is paramount when deciding about feasible and appropriate policy responses: If long-term effects were driven purely by reduced cognitive skills, the only option to avoid them would be to reduce air pollution. If, on the other hand, noncognitive abilities play a

[1] While cognitive ability captures intelligence, the ability to reason and understand complex ideas, noncognitive skills - also known under names such as socio-emotional skills, soft skills, or personality traits - comprise a variety of abilities that are weakly correlated with intelligence, such as social competencies, emotional stability and persistence.

relevant role as well, long-term effects can also be alleviated ex-post via intervention programs or investments targeting these abilities. Given that it can be extremely costly or even impossible to reduce pollution levels in some circumstances (e.g. pollution arising from another jurisdiction via transboundary atmospheric transport or from natural sources like wildfires), investigating whether alternative options exist to compensate individuals ex-post for exposure to high levels of pollution while in utero, is highly policy-relevant.

The existing evidence regarding these channels is incomplete. A number of studies find that prenatal pollution exposure causes worse performance in standardized achievement tests taken in primary and high school (Bharadwaj et al., 2017; Sanders, 2012), as well as tests of fluid intelligence (Molina, 2021), pointing strongly towards cognitive ability as a relevant channel.[2] Regarding the second main component of human capital, noncognitive abilities, causal evidence is missing. Therefore, the aim of this paper is to answer the question whether in-utero exposure to air pollution has a causal impact on noncognitive abilities, and to assess how important this potential channel is, relative to the cognitive ability mechanism.

I employ data on noncognitive abilities during childhood from the German Socio-Economic Panel (Goebel et al., 2019). Specifically, the survey includes mother-reported Big Five personality traits (Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism) assessed at ages 2-10 for children born between 2000 and 2015. I combine this with data on particulate matter with a diameter of less than 10 μm (PM10) measured at outdoor monitors and reanalysis data on meteorological conditions. To address the issues of endogeneity and measurement error in particulate matter exposure, I exploit plausibly exogenous variation in thermal inversions, a meteorological phenomenon that deteriorates air quality (following e.g. Arceo et al., 2016; Jans et al., 2018; Molina, 2021).

Results show that prenatal pollution exposure raises neuroticism, or put differently, reduces emotional stability. A 1 unit increase in gestational PM10 exposure raises Neuroticism measured at ages 5 to 10 by approximately 7% of a standard deviation. The effect

[2] In fact, while IQ and school achievement test scores are commonly considered as measures of cognitive ability, certain personality traits, esp. Conscientiousness and Emotional Stability, were found to have predictive power for these outcomes as well. However, cognitive skills can typically explain a larger part in the variation of these outcomes (e.g. Almlund et al., 2011; Borghans et al., 2008; Moutafi et al., 2006)

is mainly driven by exposure during the second and third trimester of pregnancy, and by increases in Neuroticism at the upper part of the distribution. Other dimensions of the Big Five are not affected. The effect on Neuroticism is of the same order of magnitude as the impact on measures of cognitive ability found in earlier work. Since existing research established a negative correlation between Neuroticism and labour market outcomes, it is a plausible channel underlying the long-run adverse effects of early-life pollution exposure. Back-of-the envelope computations imply that an increase of PM10 by one standard deviation could cause reductions in earnings by roughly 0.24 – 0.29% via the increase in Neuroticism.

This study contributes to the literature on the long-run consequences of early-life exposure to air pollution by providing causal evidence for a new relevant outcome, namely noncognitive abilities. Using data on cohorts born in Germany after 2000, I study a setting with relatively low baseline pollution levels, similar to concentrations prevailing today in many developed countries.[3] I thus shed light on a different part of the dose-response function than existing papers on the effects of gestational pollution exposure on skills, educational or labour market outcomes which are based on data from developing countries (Bharadwaj et al., 2017; Molina, 2021; Rosales-Rueda and Triyana, 2019) or cohorts born in the US during the 1970s-1980s (Isen et al., 2017; Sanders, 2012; Voorheis, 2017), i.e. settings with much higher air pollution levels.

The paper also adds to the literature on the development of non-cognitive skills. Related studies e.g. analyze the effects of family income (Akee et al., 2018), birth order (Black et al., 2018), parents' labor market incentives (Hufe, 2020) or child care arrangements (Datta Gupta and Simonsen, 2010). 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 noncognitive ability. Most closely related is work by Grönqvist, Nilsson, and Robling (2020) who analyze how childhood exposure to lead affects adult human capital and crime, and identify negative effects non-cognitive skills as an important mechanism. Relative to this study, I focus on less toxic air pollutants, which - as an externality of economic production and traffic - are omnipresent in both developing and developed countries. Thus, I complement the existing literature by

[3] Mean gestational PM10 exposure in my sample is approximately 24 $\frac{\mu g}{m^3}$.

analyzing the role of this pertinent environmental factor in the formation of noncognitive skills.

The remainder of the paper is structured as follows: The next section provides background information on prenatal pollution exposure and noncognitive abilities. Section 3 describes the data used in the analysis. I present the empirical specification in section 4 and the results in section 5. Section 6 concludes.

2. Background: Noncognitive Skills and Prenatal Air Pollution Exposure

Noncognitive skills are important predictors of educational achievement and labor market outcomes, even beyond their impact on education (Heckman, Stixrud, and Urzua, 2006; Lindqvist and Vestman, 2011), and over recent decades the returns to these skills increased relative to the returns to cognitive ability (Deming, 2017; Edin et al., 2017). To capture the different dimensions of noncognitive ability, I rely on the Big Five personality traits, assessed during childhood, as outcome variables. The Big Five are a widely used taxonomy which was developed in personality psychology. Among the five 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). Thus, these two traits in particular are potential mediators for the documented long-run effects of pollution exposure. Focusing on noncognitive ability during childhood is based on ample evidence in developmental psychology showing that by school start age children already differ substantially in character traits and that these are predictive of their adult personality (e.g. Almlund et al., 2011; De Pauw, 2017; Deal et al., 2005). Moreover, childhood noncognitive ability is a significant determinant of school performance and educational outcomes (e.g. Carneiro et al., 2007; Currie and Stabile, 2006; Johnston et al., 2014).

Findings from brain lesion studies, neuroimaging and psychopharmacological research imply that the source of all noncognitive abilities lies in the brain (e.g. Almlund et al., 2011). Personality neuroscience uses neuroimaging techniques to identify how personality traits depend on brain structure and function as well as levels of hormones and neurotransmitters. While the field is relatively new and the high costs of neuroimaging often restrict sample sizes, some findings already emerged as relatively robust, e.g. the association of Extraversion with dopamine or the association of Neuroticism with cortisol and activity in the hippocampus (Allen and DeYoung, 2017). As skills that depend on the functioning of the brain, noncognitive abilities might plausibly be affected by exposure to air pollution, as the latter not only causes respiratory and cardiovascular diseases, but can also induce damage to the central nervous system. The medical literature shows that very small particles reach the brain tissue where they cause oxidative stress and neuroinflammation. Since brain formation and growth proceed very rapidly during the prenatal period, air pollution exposure during this critical time window might cause irreversible damages to the nervous system by disrupting these processes (de Prado Bert et al., 2018). The pollutants that are most likely to cause permanent reductions in cognitive and noncognitive ability in this way are carbon monoxide (CO) and ultrafine particles, as both can cross the placenta, and thus pose most harm to the foetus. Possible mechanisms are maternal systemic and placental oxidative stress and inflammation and impaired transport of oxygen and nutrients to the fetus (Levy, 2015; Johnson et al., 2021). Recent evidence from brain imaging indeed suggests that prenatal air pollution exposure is associated with a reduction in white matter volume and changes to brain structure in humans (de Prado Bert et al., 2018; Beckwith et al., 2020), however given very small sample sizes and non-random variation in pollution exposure, such associations cannot be interpreted as reflecting causal effects.[4]

Given these premises, several epidemiological studies investigate the correlation between in-utero or early childhood exposure to air pollution and mental health issues (e.g. ADHD or Autism) or behavioral problems in childhood, i.e. outcomes related to noncognitive ability.[5] Annavarapu and Kathi (2016), Myhre et al. (2018) and Xu et al. (2016) provide reviews of this literature and point out common caveats: Most studies are based on cross-sectional comparisons between individuals living in different places, and thus

[4]For complementary evidence from animal studies using random variation in pollution exposure see e.g. Woodward et al. (2018) or Costa et al. (2014) for a review.

[5]Mental health issues are often interpreted as extreme realizations of personality traits (Almlund et al., 2011; Widiger et al., 2017) and, especially at child age, frequently measured by the same concepts as noncognitive ability (Currie and Almond, 2011; Johnston et al., 2014).

prone to omitted variable bias. Hence, while suggestive of a relationship between early air pollution exposure and noncognitive ability, the results of this literature do not reflect causal effects.

3. Data

3.1 Socio-economic Panel: Noncognitive skills

Estimating the impact of early life pollution exposure on the formation of noncognitive ability requires data that not only includes information on individuals' socioemotional skills but also on their location and time of birth. Besides, the data must cover a sufficiently large number of individuals from multiple birth cohorts, since only temporal variation in particulate matter and inversions is used in the research design. Many surveys that contain extensive information about noncognitive ability do not meet these requirements as they cover only single birth cohorts, e.g. the British longitudinal cohort studies or the studies from the U.S. Early Childhood Longitudinal Study program. Other data sets lack the necessary disaggregated geographical information, such that individual pollution exposure cannot be determined. Lastly, frequently used administrative data sources on cognitive and noncognitive skills, e.g. the Swedish military enlistment data, cover individuals who were born in a place and at a time for which accurate measurements of air pollution are unavailable.[6]

A data source that addresses most of these demands is the German Socio-Economic Panel (SOEP), a large household panel survey started in 1984, which covers roughly 15,000 households and 30,000 individuals (SOEP, 2019). It includes mother-reported Big Five personality traits for all children aged 2-10 in SOEP households which I use as measures of childhood noncognitive abilities. The relevant information is based on the "Mother-and-child" questionnaires which were introduced in 2005 for 2- to 3-year-old children, in 2008 for 5- to 6-year-olds and in 2012 for 9- to 10-year-olds. Mothers an-

[6] The enlistment data offers a large sample size and high quality measures of noncognitive skills based on interviews with a trained psychologist, but is not suitable for my analysis, because air quality data for Sweden is not available at a large scale before 1980. Due to falling demand for conscripts, the share of a birth cohort that was enlisted fell to roughly 70% for mid-1980 born cohorts, implying potential selection bias problems when analyzing cohorts born after 1980. (Grönqvist, Öckert, and Vlachos, 2017)

swer questions on their child's behavior on a scale from 0 to 10. Each question can be mapped into one of the Big Five domains. Mothers' of 2-3 year old children are asked to answer only one question per trait and neuroticism is not yet included. When children are in the older two age groups, two questions are included per domain and all five traits are assessed. The questions are presented in appendix table A.1. One important thing to note about the questionnaires is that the items intended to measure openness are likely to capture at least in part the child's cognitive ability (e.g. *'My child is quick at learning new things vs. needs more time'*). Thus, the main focus of my analysis will lie on the other four traits which isolate non-cognitive skills. To construct the outcomes of interest I collect the relevant information from the 2005-2018 mother-and-child questionnaires, I recode items where necessary to ensure that high values reflect higher realizations of the respective trait, add up values for items within each domain and standardize the resulting scores within each age group (2-3, 5-6 and 9-10). For children observed at multiple ages, I keep only the most recent observation, as personality differences become more pronounced with age and my main interest is in long-run effects of pollution exposure. The relevant information is available for individuals born between 2000 and 2012 for Neuroticism (*Sample II*) and for those born 2000-2015 for the other four traits (*Sample I*), i.e. the included birth cohorts extend over more than a decade.

To assign pollution exposure to individuals, I rely on information on year and month of birth as well as county of residence. For the majority of children (71% in *Sample I* and 65% in *Sample II*), I can identify the county of birth, as the county of residence during the year of birth. In the remaining cases, the households entered the panel after the child was already born such that the place of birth is unknown. As a proxy for county of birth, I assign the county of residence during the first wave the household was interviewed. Wrong assignments can induce measurement error in early-life particulate matter exposure which is not addressed by the instrumental variable strategy. Since measurement error causes attenuation bias, any results would reflect a lower bound. I restrict the sample to children born in a county with at least 20 observations in my sample since effects are identified from variation in environmental conditions during early life between children born within the same county. This restriction is intended to ensure that the comparison groups are large enough to avoid spurious results generated by outliers. This yields 9,470 individuals across 192 counties in *Sample I* and 6,548 individuals across 155 counties in

Sample II. Lastly, the SOEP provides a multitude of relevant demographic and socioeconomic background variables, e.g. the child's gender, age in month and migration history, whether it lives in a single-parent household, parental education, and the number of siblings.[7] Basic sample characteristics are summarized in the middle part of table 1.

Table 1: Summary Statistics

	Sample I (B5 \ Neuro)	*Sample II* (Neuroticism)
Observations	9,470	6,548
Counties	192	155
Years of Birth	2000-2015	2000-2012
Age	6.9 (2.9)	8.2 (2.0)
Migration Background [%]	31.7	29.7
College-educated mothers [%]	26.8	26.7
Single-parent households [%]	15.6	17.6
PM10 in-utero [$\frac{\mu g}{m^3}$]	24.1 (5.0)	24.6 (5.1)
NO2 in-utero [$\frac{\mu g}{m^3}$]	28.8 (10.6)	29.0 (10.8)
CO in-utero [$\frac{\mu g}{m^3}$]	460.2 (193.3)	480.7 (203.5)
Inversions in-utero [%]	39.9 (7.1)	39.7 (7.0)

Note: Summary statistics based on data from the SOEP, version 35. The table reports mean values and standard deviations (in parentheses).

Given that the Five Factor Model is the most common taxonomy of personality and widely employed in economics, using the Big Five personality traits as main outcomes allows to benchmark my results against other studies and to conduct back-of-the-envelope calculations translating effects on noncognitive abilities into effects on earnings. However, the fact that these variables are based on maternal assessments and short scales with only

[7] To measure parental education I construct dummy variables reflecting a low, medium or high level of education for mothers and fathers. Low education is defined as less than high school (Abitur) or vocational training, medium education is defined as having completed high school or vocational training, but no tertiary degree, and high education is defined as having completed a tertiary degree. I also include children in the sample when information on parental education is missing, and define separate dummies for these cases.

one or two items per domain might raise concerns about measurement error. Thus, in the following I present some descriptive evidence to illustrate that these measures do contain substantial information regarding the children's non-cognitive skills.

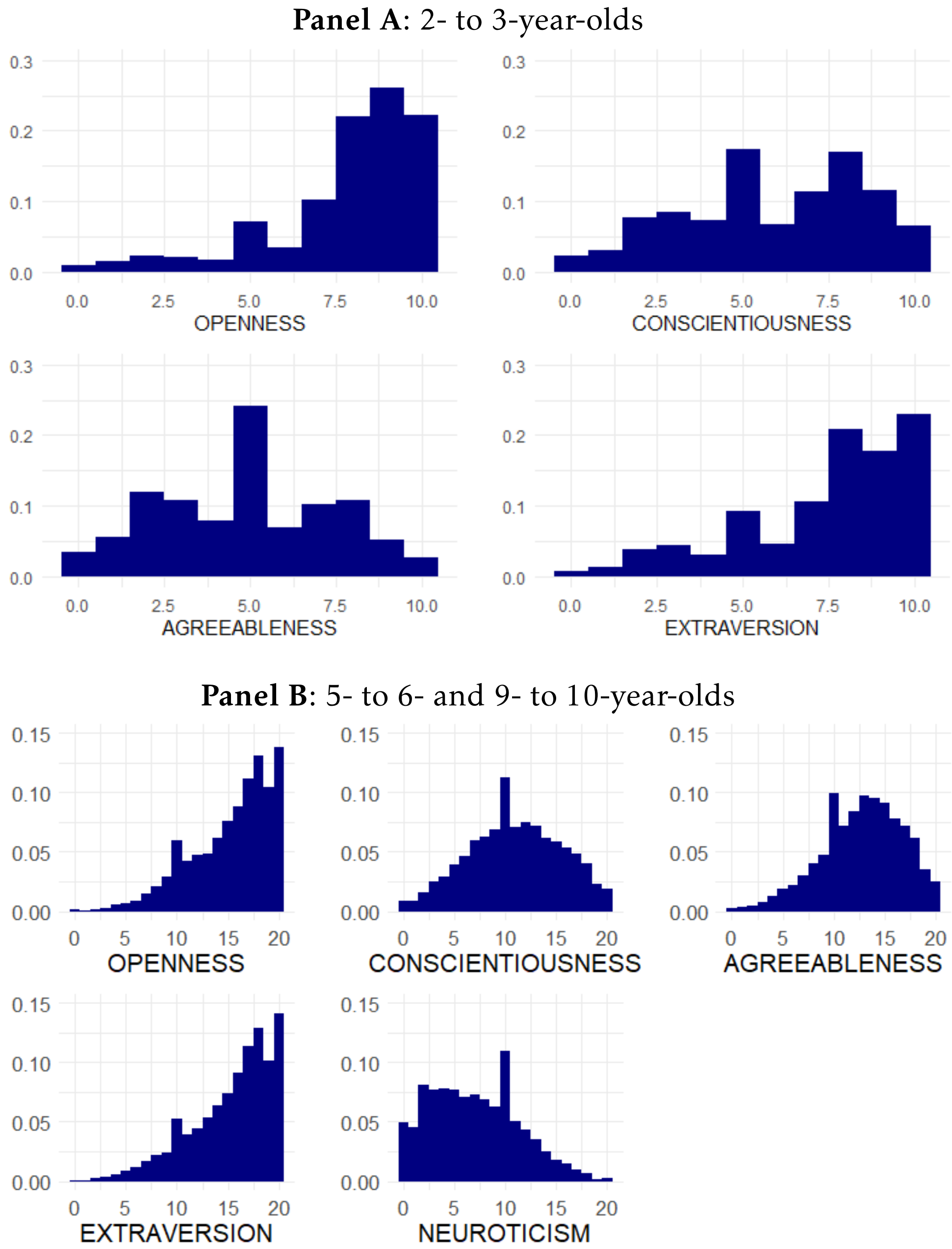


Figure 1: Distribution of Big Five personality traits by age group.

Note: The distributions are based on all available observations in the SOEP, i.e. including multiple observations for the same child and children born in counties with less than 20 individuals. Panel A is based on 8,037 observations, Panel B on 11,657.

Firstly, Figure 1 shows the distribution of the Big Five traits by age group (not standardized). Panel A shows distributions for the age range from 2-3 years. While there is some clustering at high values for Openness and Extraversion which are picked by 20%

to 25% of all mothers, all potential values do occur with positive frequency across all four outcomes. In the older age groups, the distributions exhibit even more variation as mothers answer two instead of one question per domain (Panel B). One important thing to note is that the distributions for all five traits exhibit a spike at the intermediate value of 10, with different intensity across outcomes. A potential explanation for this is that some mothers have difficulties when trying to assess their children's noncognitive skills and thus opt for the middle values. However, the vast majority of respondents do not follow that strategy, suggesting that their answers are informative about the personality of their children. In figure A.1, the distributions are depicted by maternal education. The variables exhibit strong variation in all socio-economic groups, but less educated mothers tend to choose intermediate values slightly more often than highly educated mothers. Given that I control for parental education in the regressions and conduct falsification tests showing that the instrument is not significantly correlated with family characteristics, this should not affect the results. In the main analysis, I thus use the data as it is, but in a robustness check, I drop observations where mothers opted for the intermediate values across several items.

To further explore the validy of the outcome variables, I show correlations of the mother-reported Big Five with father-, and self-reported personality traits in figure 2.[8] 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. All variables are standardized to have mean zero and standard deviation one within the respective age-by-respondent cell. All correlations are positive and statistically significant, implying not only that mother-reports are in line with reports by others for the same age range, but also predictive for noncognitive ability during adolescence, at an age relatively close to transition into the labor market or higher education.

In appendix figure A.2, I also show partial correlations between the mother-reported Big Five and measures of the child's school performance, well-being and preferences. In

[8]Father-reported Big Five are available only for a subsample of 1,144 children at age 9-10. Self-reported Big Five are available from age 11-12 onward. Most of the adolescents observed at these ages live in households who were not yet part of the SOEP during their year of birth. For these reasons, I rely solely on the mother-reported Big Five in the main analysis.

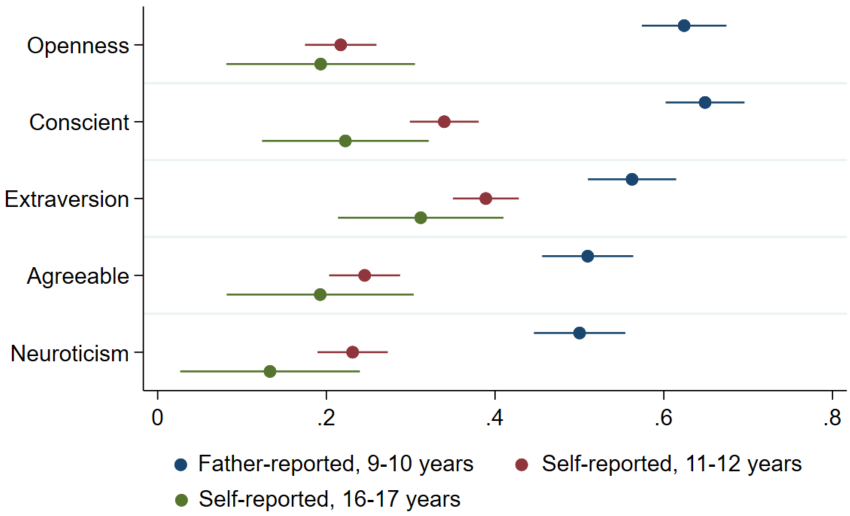


Figure 2: Mother-, father- and self-reported Big Five

Note: The figure presents correlations between mother-assessed and either father- or self-reported Big Five personality traits. Father-reported Big Five are measured with the same scales as mother-reported variables based on sample sizes ranging from 1,085 to 1,097 depending on the domain. Self-reported Big Five at ages 11-12 and 16-17 are measured with three items per domain and available for 2,215 to 2,266 and 355 to 361 individuals, respectively who were also observed at age 9-10. 95%-confidence intervalls are based on heteroscedasticity robust standard errors.

each case I control for parental education, a single-parent household dummy, child gender and migration background. Firstly, mother-reported Big Five when the child is aged 5-6 years are relevant predictors of the mother-assessed probability that the child will graduate from the academic track of the German school system when the child is 7 to 8 years old (i.e. before track choice is made). Openness is a strong positive predictor, which is unsurprising, given that openness as measured in the SOEP likely reflects cognitive skills to some extent. More importantly, conscientiousness and neuroticism are also relevant predictors, with positive and negative sign, respectively. Child-reported life satisfaction and risk aversion at ages 11-12 are both correlated with mother-assessed Big Five at age 9-10: Agreeableness, Extraversion and Conscientiousness are positively associated with life satisfaction. All five traits are correlated with risk aversion, with the strongest, positive association for Neuroticism. The signs of these correlations are plausible, providing further support for the validity of the mother-reported noncognitive skills.

3.2 Pollution, Weather and Inversions

I obtain data on daily PM10 concentration measured at outdoor monitors between 2001 and 2016 from the federal environmental agency (Umweltbundesamt [UBA]). I use data from 181 stations which were active throughout the full time period. A map of the monitor locations is provided in figure A.3. To assign PM10 to counties, I use inverse distance weighting, based on stations within a radius of 60km around the county centroid.[9] The median distance between the county centroid and the assigned stations is 25.8 km. I aggregate daily PM10 concentrations to trimesters (= 90 days) and in-utero periods (= 270 days). I keep only observations with less than 115 missing daily values during the in-utero period and at most 40 missing values per trimester. Given that the European Union only introduced binding limit values for PM10 in 2005, there were relatively few measurement stations in Germany during the early 2000s, with the number of monitors rising sharply only in 2003. Hence, the dataset does not cover the full sample period (the oldest individual in the sample were born in 2000) and only 190 out of the 192 relevant counties. On top of that, there is a non-trivial number of missing measurements in the data. In total, for 14.5% of the final sample, the in-utero period PM10 concentration is missing. Hence, I will report both 2SLS and reduced form results, since the first are informative about the quantitative effect of pollution on noncognitive abilities, but the second are based on a larger sample. I also collect data on nitrogen dioxide (NO2) and carbon monoxide (CO) concentration from UBA. They are common co-pollutants of PM10, and CO can also cause damage to the unborn child.[10] I proceed in the same way as described above to transform monitor-by-day observations into county level concentrations during the periods of interest.

To construct the instrumental variable based on thermal inversion periods, I employ reanalysis data from the European Centre for Medium-Range Weather Forecasts (ECMWF)

[9] I restrict the maximum number of stations per county to three since using more stations increases the number of missing values. In a small number of large cities, which are also counties, there are more than three monitors within the city boundaries. In these cases I assign all monitors within at most 20km distance.

[10] NO2 concentrations are measured at 230 distinct stations over the relevant time period, with a median distance of 21km between county centroids and monitors. The number of active CO monitors is 61. Data can be assigned to a subsample of 122 counties, with a median distance of 32.3 km between county centroids and CO monitors.

on surface level and upper air temperature for the years 1999-2016. During normal times, air temperature decreases with altitude. A thermal inversion occurs when the relationship between temperature at different altitude levels is reversed, i.e. temperature increases with altitude. The data are available at an hourly level on a regular 0.25° latitude x 0.25° longitude grid.[11] Upper air temperature is measured as temperature at a pressure level 50 hPa below the surface level pressure in the county, which corresponds to approximately 400-500m higher altitude. I then assign data from grid points to counties, based on the inverse distance weighting method, including all grid points within a 30km radius around a county centroid. To determine the occurrence of night-time inversions, I average both surface level and upper air temperature between 2 am and 6 am. If the difference between nightly upper air temperature and surface level temperature is positive, the county experienced a night-time inversion on the particular day. I aggregate the daily data to the time periods of interest (trimesters or in-utero period). I define the instrument as the share of days with a night-time inversion during the time period of interest.

Lastly, I collect reanalysis data on meteorological conditions for the time period 1999-2016 from the ECMWF, to include as control variables. Monthly average temperature, precipitation and wind speed are available on a 0.1° x 0.1° grid.[12] I aggregate data to counties, using all grid points falling into a county, or in case of small counties without a point on its territory, I assign data from up to 10 closest points within 40km distance using inverse distance weighting. I aggregate the weather data from the monthly level to the time periods of interest.

The pollution and inversion variables for the sample are summarized in the bottom part of table 1. I report concentrations of PM10, NO2 and CO as well as the frequency of inversions for the in-utero period, which is defined as the 270 day or 9-month period ending with but including the month of birth.

[11] Specifically the two data products I use are *ERA5 hourly data on single levels from 1979 to present* and *ERA5 hourly data on pressure levels from 1979 to present*

[12] Product: *ERA5-Land monthly averaged data from 1981 to present*

4. Empirical Strategy

To identify the effect of air pollution exposure on children's socioemotional skills, I rely on an approach based on thermal inversions, following e.g. Arceo et al. (2016); Colmer et al. (2021); Jans et al. (2018) and Molina (2021).

The baseline model is given by:

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

where NC^{a}_{icym} denotes a measure of non-cognitive ability of individual i, born in county c in month m of year y. The age group at which the skills are assessed is represented by a. PM_{cym} is local PM10 concentration during the gestational period. The model includes county fixed effects θ_c to account for persistent difference in pollution and skill levels across locations, month-of-birth fixed effects θ_m to control for seasonality in air quality, and year-of-birth fixed effects θ_y to capture changes in cognitive skills over time which affect individuals in all counties equally. θ_a controls for age-specific trends in outcomes. The model further includes individual and family background characteristics $\mathbf{X}_i$ (child gender, age in month and its square, migration background and dummies for parental education levels) and meteorological conditions, $\mathbf{W}_{cym}$. These include third order polynomials in temperature, precipitation and wind speed during the gestational period. Standard errors are clustered at the county level.

For the OLS estimate of β to be consistent, the unobserved determinants of childhood non-cognitive abilities summarized in u_{icyma} must not be correlated with PM10 levels conditional on fixed effects and control variables. This assumption is probably violated, e.g. due to region specific economic shocks affecting both air quality as well as parental income which might be spent on investments into the child's skill development. Secondly, individual pollution exposure is measured with error which leads to attenuation bias. To address these issues, I rely on the meteorological phenomenon of thermal inversions to extract exogenous variation in air quality. Under normal conditions, air temperature decreases with altitude. Emissions released at the ground level rise and disperse in the air. During an inversion, air temperature increases with altitude, i.e. upper air layers are warmer than ground level air. The warm upper air layer acts like a ceiling that prevents ground level pollution from rising and dispersing. As pollutants are trapped beneath the warm air layer, their surface level concentration increases.

Thermal inversions are a meteorological phenomenon. They exhibit a seasonal pattern with inversions occurring more frequently in the winter as compared to the summer. However, conditional on month-of-birth fixed effects and weather controls, it is as good as random whether the specific combination of meteorological conditions occurs that gives rise to a thermal inversion. Importantly, the frequency of inversions should be plausibly uncorrelated with local business cycles. At pollution levels that were common in Germany during my sample period, inversions usually do not lead to visible smog events or extremely poor air quality, and are thus unlikely to trigger avoidance behavior. Besides, following e.g. Jans et al. (2018) and Molina (2021), I exclusively consider nighttime inversions, which should be even less likely to induce any behavioral responses.

Specifically, the first stage model is given by:

$$PM_{icym} = \alpha Inv_{cym} + \gamma' \mathbf{X}_i + \delta' \mathbf{W}_{cym} + \theta_c + \theta_y + \theta_m + \theta_a + u_{icyma} \tag{2}$$

where Inv_{cym} is the share of days on which a nighttime inversion occurred during the period of interest. In addition to the variables mentioned above, I add one lead and one lag of the instrument to $\mathbf{W}_{cym}$ in both the first and second stage model, to account for autocorrelation in inversion frequency.

The first stage results in table 2 replicate the finding from previous studies that thermal inversions are a strong instrument for PM10, thus satisfying the relevance condition. The effect magnitude in columns 1 and 2 imply that mean PM10 concentration during a 9 month period increases by 1.1 and 1.2 $\frac{\mu g}{m^3}$, respectively, in the two analysis samples, for a standard deviation increase in inversion frequency (.07).[13] The associated F-Statistics are large. In columns 3 to 6, I show results when replacing PM10 by either NO2 or CO as outcome. Inversions have a somewhat smaller, but highly significant effect on NO2, which is a common co-pollutant of PM10. For CO, the effect is also positive, but less significant, such that the F-Statistic is clearly below the common threshold for a sufficiently strong instrument. In sum, thermal inversions increase the concentration of all three pollutants considered in the table, and with only one instrument, I cannot cleanly disentangle their

[13]The coefficients in the table depict the effect for moving from no inversions at all to having an inversion each night during the period of interest. In the data, however, none of these two extreme cases occurs, and the actual variation in inversion frequency is substantially smaller.

distinct effects on noncognitive skills. However, the first stage results suggest that PM10 is most likely to drive any results in the second stage, because (i) unlike CO and fine particles, NO2 is not commonly considered as a major risk factor in the medical literature, and (ii) the first stage effect for CO is weak.

Table 2: First Stage Results

	PM10	PM10	NO2	NO2	CO	CO
Share Inversions	15.86***	17.82***	10.85***	9.87***	122.12	187.93**
in utero	(1.611)	(1.985)	(1.223)	(1.535)	(87.679)	(90.747)
Observations	7,645	4,911	8,880	6,099	5,494	3,679
F-Statistic	98.68	81.36	73.63	41.47	1.91	3.91
Sample	I	II	I	II	I	II

Note: The table depicts estimates of the effects of inversion frequency during the nine-month in-utero period on pollution concentration during that period from model (2). All outcomes are measured in $\frac{\mu g}{m^3}$. Regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, weather controls and individual background characteristics as described in the text. Standard errors clustered at the county level are reported in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01

The key identifying assumption underlying the 2SLS and reduced form estimations is that, conditional on the included covariates, the frequency of thermal inversions during the in-utero period affects noncognitive skills in childhood only through their effect on air pollution. This assumption will be violated if families of children exposed to more inversion periods during early life differ systematically from families of children exposed to fewer inversions along unobservable characteristics.

As a test of the identifying assumption, I check whether inversion frequency is systematically correlated with observed predetermined family background variables. The results of these falsification checks are displayed in table 3. The estimated coefficients reflect the change in outcomes for a one standard deviation increase in inversion frequency (0.07) and reveal that family characteristics show no consistent pattern with respect to inversion frequency: For instance, fathers of children exposed to more inversions while in-utero

Table 3: Falsification Tests

	Maternal Age at birth	Migration History	Tertiary degree (Mother)	Less than High School (Mother)	Tertiary degree (Father)	Less than High School (Father)
Panel A: Sample I						
Share Inversions	-.1691	-.0054	-.0177	.0080	-.0150	.0137
in utero	(.1570)	(.0117)	(.0113)	(.0098)	(.0163)	(.0128)
Observations	9,200	9,440	9,201	9,201	6,284	6,284
Panel B: Sample II						
Share Inversions	-.1861	-.0076	-.0042	-.0014	-.0064	.0130
in utero	(.1963)	(.0144)	(.0136)	(.0116)	(.0184)	(.0151)
Observations	6,326	6,548	6,380	6,380	4,157	4,157

Note: The table depicts coefficients from OLS regressions of family characteristics on the share of inversions during the child's in-utero period. Maternal age is measured in years. The other outcomes are dummy variables. Regressions control for cubic functions of precipitation, temperature and wind speed. Standard errors clustered at the county level are reported in parentheses. Significance levels: *p<0.1; **p<0.05; ***p<0.01

tend to be slightly less educated, but at the same time, on average, slightly fewer of these children have a migration history. Importantly, none of the tested variables (maternal and paternal education, maternal age at birth and migration background) are significantly correlated with the instrument and the point estimates are generally small in magnitude.

5. Results

This sections starts by presenting the 2SLS and reduced form results for the effects of in-utero exposure to air pollution on the Big Five personality traits. The main finding is that worse air quality reduces emotional stability. Thereafter, I explore this result in terms of heterogeneity and critical phases of exposure, and test its robustness to changes in sample construction, the set of included covariates and the choice of the instrument.

5.1 Main results

Panel A of Table 4 displays 2SLS estimates of the effect of in-utero exposure to PM10 on the Big Five personality traits. A large and significant impact is found for neuroticism which increases by 7.1% of a standard deviation for an increase in mean PM10 concentration by one $\frac{\mu g}{m^3}$. This implies that individuals exposed to higher levels of air pollution during the prenatal period are less emotionally stable in childhood, i.e. more fearful and less self-confident. The second largest point estimate emerges for openness. As mentioned earlier, openness as measured in the SOEP most likely partially reflects cognitive skills. Thus, the negative estimate is in line with existing results on the negative impact of in-utero exposure to air pollution on later life cognitive ability. The fact that the negative effect is not statistically significant is unsurprising given that the questions asked to assess openness are only crude measures of cognitive ability as compared to detailed tests used in previous studies. The remaining three traits are not affected by gestational particulate matter levels. The point estimates are close to zero (all below 1.7% of a standard deviation), and not statistically significant.

These results are confirmed by the estimates from the reduced form models, which are presented in panel B. As mentioned above, the reduced form can be estimated on larger samples. Coefficients are multiplied by 0.07 to reflect the estimated effects for a one standard deviation increase in inversion frequency. Based on the first stage results this corresponds to an increase in PM10 by a bit more than one $\frac{\mu g}{m^3}$. In line with the 2SLS estimates, the only significant result emerges for neuroticism, which rises by 6.7% of a standard deviation in response to the increase in in-utero inversion frequency. This is very similar in terms of magnitude to the 2SLS estimate, implying that the relationship holds in counties with and without PM10 monitors. Regarding the other four traits, the reduced form confirms the absence of any effects. Again, apart from the negative but insignificant coefficient for openness, the point estimates are small in magnitude. A potential explanation for these null effects could be that these traits are measured on a different sample than neuroticism, including younger children. Possibly, personality differences in 2-3-year-olds are not yet pronounced enough or the measures used in the SOEP for this age group are too crude to uncover the impact of prenatal pollution exposure. Therefore, in table A.2, I repeat the estimation for these four outcomes on *sample II*, i.e. the sample used for the analysis of neuroticism, including only children aged 5 or older. The results are similar, again showing no significant effect on any of the four traits. The pattern is also unchanged, showing the largest negative point estimates for openness and relatively small estimates for the other three outcomes.

5.2 Additional Results

Having established that in-utero exposure to particulate matter increases neuroticism in childhood, this paragraph further examines this effect in terms of critical windows of exposure and effect heterogeneity.

Effect Heterogeneity. I analyze effect heterogeneity with respect to child and family characteristics, namely child gender, age at which neuroticism is assessed (5-6 vs. 9-10 years), and current household income (above vs. below median). I do not find evidence for heterogeneity in the effect of inversions on neuroticism along any of the three dimensions. Results are illustrated in figure A.4. In all three cases, the coefficient on the interaction term is close to zero and insignificant, whereas the main effect remains statistically significant and of similar magnitude as in the baseline model. In summary, the effects of prenatal air pollution exposure emerge already by school start age, affect both genders

Table 4: Impact of prenatal PM10 exposure on the Big Five

	Openness	Consc.'ness	Extraversion	Agreeableness	Neuroticism
Panel A: 2SLS					
PM10	-.0361	.005	.0033	.0163	.0714**
in utero	(.0288)	(.0279)	(.0258)	(.0261)	.0314
Observations	7,655	7,645	7,666	7,614	4,911
Counties	163	163	164	162	122
1st Stage F-Stat.	98.	96.9	98.7	95.1	80.5
Panel B: Reduced Form					
Inversions	-.0423	-.0182	-.0034	.0231	.0667**
in utero	(.0279)	(.0264)	(.0261)	(.0265)	(.0323)
Observations	9,454	9,440	9,445	9,423	6,548
Counties	192	192	192	192	155

Note: Panel A displays 2SLS-estimates of the effect of a 1 $\frac{\mu g}{m^3}$ increase in PM10 concentration during the in-utero period on Big Five personality traits. PM10 is instrumented by inversion frequency. Panel B shows the OLS estimates of the effect of an increase in the share of days with a nighttime inversion during the in-utero period by one standard deviation (+7%) on the outcomes. Outcomes are standardized within age groups. All regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, controls for parental education, child gender, migration background, age in months and its square, cubic functions of temperature, wind speed and precipitation, plus one lead and one lag of the instrument. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

equally and are not dampened by parental resources.

Secondly, I examine heterogeneity in effect magnitude along the distribution of neuroticism. I construct two indicator variables for the value of neuroticism falling below the first quartile and above the third quartile, respectively. Figure 3 shows 2SLS and reduced form results. While all point estimates have the expected sign, i.e. on average air pollution exposure reduces the probability to fall below the lower quartile and increases the probability to fall above the upper quartile, only the latter effect is statistically significant and also larger in magnitude (-1.5 pp vs +3 pp). This result is noteworthy as it implies that prenatal air pollution might affect child mental health: In modern personality psychology, mental disorders are conceptualized as extreme realizations of the Big Five traits. High values of neuroticism show particularly robust and consistent correlations with a range of mental health issues, e.g. depression and anxiety disorders. (Almlund et al., 2011; Widiger et al., 2017) Hence, the fact that the main result is mostly driven by increases in neuroticism in the upper part of the distribution suggests that prenatal particulate matter exposure might not just reduce an important noncognitive skill within the range of 'normal' variations in personality, but could even give rise to mental health issues. It should be stressed that this is only suggestive, and with the available data, I am unable to explore this in more depth.[14]

Trimester level effects. To assess critical windows of exposure, I split the in-utero period into three trimesters, and run the reduced form regression including variables measuring inversion frequency separately for each trimester. Besides, to analyze whether postnatal air pollution exposure also generates adverse long-run effects on neuroticism, the regression includes inversions during the first nine month after birth, broken down into three-month periods. Symmetrically, I include inversion frequency during three "placebo trimesters", i.e. three-month periods preceding conception. Assuming that air pollution affects childhood neuroticism solely through physiological channels, the frequency of thermal inversions before conception should not have a significant impact. Regressions include the same fixed effects, individual and family characteristics as before and trimester-specific, quadratic weather controls. Table 5 presents the results. Sig-

[14]The SOEP questionnaires do not include instruments measuring child mental health, and formal diagnosis of mental disorders is a very rare outcome that cannot sensibly be analyzed given the sample size.

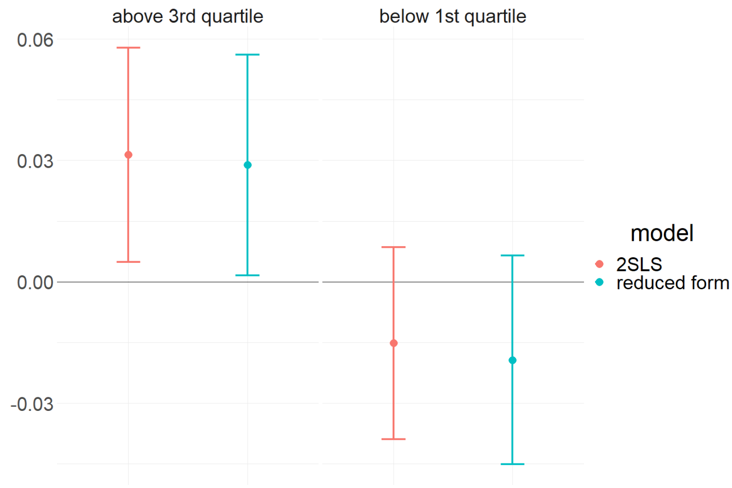


Figure 3: Effect Heterogeneity along the distribution of Neuroticism

Note: Red colored dots depict 2SLS estimates of the effect of a 1 $\frac{\mu g}{m^3}$ increase in PM10 concentration during the in-utero period on indicator variables taking the value one if age standardized neuroticism falls above the 75th percentile of the sample distribution (left) or below the 25th percentile of the sample distribution (right), respectively. Blue colored dots depict OLS estimates of the effect of a one standard deviation increase in the share of days with a nighttime inversion during the in-utero period on the same outcomes. All regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, controls for parental education, child gender, migration background, age in months and its square, cubic functions of temperature, wind speed and precipitation, plus one lead and one lag of the instrument. Bars represent 95%-confidence intervals, based on standard errors clustered at the county level.

nificant positive effects are found for inversion frequency during the second and third trimester of pregnancy. Inversions during the first trimester and the first nine month after birth have no significant impact on neuroticism. This result resembles the findings on the effect of air pollution on performance in school achievement tests in Bharadwaj et al. (2017) (effects mainly in the third trimester) and cognitive ability in Molina (2021) (significant effects only in the second trimester). Reassuringly, the coefficients on the three placebo trimesters are insignificant and also of smaller magnitude than the effects of in-utero exposure.

Postnatal exposure. The reduced form results just presented indicate that neuroticism is not affected by postnatal exposure to air pollution. To corroborate this finding, I analyse the effects of air pollution during the first nine month after birth in the main 2SLS and reduced form models, for neuroticism as well as the other noncognitive skills. Since early life is also a critical period of development, I run these regressions to make sure not to miss any relevant effects. 2SLS and reduced form models are defined as before, with the

Table 5: Neuroticism - Effects by Trimester

	Neuroticism
Inversions	-.0050
9-7 months before conception	(.0165)
Inversions	.0227
6-4 months before conception	(.0167)
Inversions	.0179
3-1 months before conception	(.0163)
Inversions	-.0079
1st trimester	(.0187)
Inversions	.0327**
2nd trimester	(.0164)
Inversions	.0384**
3rd trimester	(.0169)
Inversions	-.0231
1-3 months post birth	(.0176)
Inversions	.0171
4-6 months post birth	(.0163)
Inversions	-.0220
7-9 months post birth	(.0169)
Observations	6,548

Note: The table shows OLS estimates of the effect of an increase in the share of days with a nighttime inversion during the in-utero period by one standard deviation (+7%) on Neuroticism (age-standardized). Regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, as well as covariates for parental education, child gender, migration background, age in months and its square, plus trimester-specific, quadratic functions of temperature, wind speed and precipitation. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

only difference that weather controls are included for both the pre- and postnatal period to control for any possible long-run effects of early life weather conditions. I find no effects of postnatal PM10 exposure on any of the five outcomes. As mentioned above, the finding that noncognitive ability, in particular neuroticism, is only affected by prenatal pollution exposure is in line with results for cognitive ability. The result is also plausible given the fact that both types of skills are governed by brain structure and function which develops most rapidly in the gestational period. However, adverse long-run effects on educational and labor market outcomes where found for exposure during both the in-utero period and the first year of life (Isen et al., 2017; Voorheis, 2017). This implies that other components of human capital, e.g. physical health, might also mediate a part of the adverse long-run effects. This is in line with results by Klauber et al. (2021) who follow children from birth to age five and find that in-utero exposure to particulate matter causes subtle but persistent damage to respiratory health.

Robustness Checks. The main finding from the preceding paragraphs is that in-utero exposure to particulate matter increases neuroticism at ages 5-10. The similarity of the 2SLS and reduced form results as well as the absence of an effect during the placebo trimesters provide a first confirmation that the finding is not driven by correlated unobservables. However, given the modest sample size and the large number of hypotheses tested, I subject this finding to a number of additional robustness checks.

First, I test sensivity to changes in sample construction. Results are reported in appendix table A.4. Panel A shows 2SLS estimates, while panel B presents reduced form estimates. Column 1 replicates the baseline result. In columns 2 and 3, I vary the minimum number of observations for a county to be included in the sample to either 15 (column 2) or 25 (column 3) instead of 20. While the number of included individuals and counties changes notably across the three columns, the point estimates are positive and significant in all cases, with their size varying slightly between 6.65% and 8.88%. The fact that the estimated coefficient gets larger and more significant for the larger cut-off value is plausible as it is easier to precisely estimate the effect of pollution exposure if the within-county comparison group is larger. In columns 4 and 5, I use all available observations in the SOEP, instead of keeping only one observation per individual. If a child is observed at both 5-6 and 9-10 years of age, both observations are included. Column 4 reports results from an unweighted regression. In column 5, I weight observations by the inverse of the number of observations per individual. All estimates are similar in quantitative terms to the baseline result, implying that the finding is not driven by the specific choice of the analysis sample.

Secondly, I analyze the robustness of the results to the set of included covariates. Results are displayed in appendix table A.5. The baseline result is replicated in column 1. Column 2 shows estimates from a specification that includes a more comprehensive set of background controls. Specifically, I add an indicator for a single parent household, dummies for maternal age at birth (in 5 year steps), and birth order dummies. Columns 4 and 5 depict results from specifications including more and less leads and lags of the instrument, respectively. In all three model versions, the estimated effects are similar in both magnitude and significance to the baseline finding. This holds for both the 2SLS and reduced form results.

Next, I test whether the result could be just a statistical artifact, arising because certain mothers have difficulties in assessing their children's non-cognitive skills. This concern arises from the fact that intermediate values are chosen 'too often' when mothers assess their children's personality traits, and more so by less-educated mothers (see section 3.1). Column 2 of table A.6 presents results from estimation after dropping observations from the sample where mothers picked the intermediate value of 5 for all items underlying conscientiousness and neuroticism, the two outcomes that showed most excess mass for the middle value (see figure 1). The estimated effect remains statistically significant and drops by less than .2 percentage points relative to the baseline result in column 1 in both the 2SLS- and reduced form specifications. Since this approach only drops a small number of observations from the sample, I also employ an alternative strategy: I compute the variance in responses across the ten questions underlying the Big Five personality traits in the SOEP, after demeaning all items in the full sample. I then drop the 5% (10%) of the sample with the lowest variance. This reflects the idea that mothers who are good at assessing their children's personality will deviate more from the population mean. This approach is partly based on Falk, Neuber, and Strack (2021).[15] The results are depicted

[15] In the paper, Falk et al. (2021) build a choice model of survey response behavior where respondents have imperfect self-knowledge. They develop an estimator of self-knowledge and show that regressions using self-reported risk-attitudes or non-cognitive skills on either the left or right side yield larger estimated

in columns 3 and 4. The effect of exposure to air pollution during the in-utero period remains statistically significant in both cases. Both in the reduced form and the 2SLS specification, it increases slightly in magnitude relative to the baseline model. Overall, responses by mothers who have difficulties in evaluating their child's noncognitive ability seem not to bias the results.

In a final robustness check, I re-run the 2SLS and reduced form models with a different instrument, the inverse of the planetary boundary layer height (PBLH). This instrument was recently used by Godzinski and Castillo (2021). The planetary boundary layer (PBL) is the lowest part of the troposhere, i.e. the air layer directly above the surface. Pollution emitted at the ground level disperses within this air layer. The higher the PBL, the larger the volume of air in which emissions can dissipate, leading to lower ground level concentration of air pollutants. Hence, for lower values of the PBLH, or higher values of the inverse of the PBLH, pollution concentration near the surface increases. This implies that PBLH exploits the same physical mechanism to extract exogenous variation in air pollution as thermal inversions and thus also affects multiple air pollutants. However, while the main instrument only measures whether on any given day a thermal inversion either occurs or does not occur, the inverse of the PBLH is a continuous variable accounting for the strength of the meteorological phenomenon. To construct the alternative instrument, I collect data on average monthly PBLH (in 1000 meters) from the ECMWF, aggregate this to counties and the time period of interest as described in section 3.2, and then take the inverse of the resulting variable.[16] The correlation of inversion frequency and inverse PBLH is 0.266 in *sample II*. Table A.7 shows 2SLS and reduced form results based on the

coefficients and R^2 in the subsample of individuals with higher self-knowledge. In my application, mothers answer questions about their children's personality. Hence, the term self-knowledge is not suitable here, but the general issue is the same: mothers differ in their capability to memorize and recall relevant information about child behavior, or in their knowledge about 'normal' child behavior. Mothers with the lowest levels of this type of knowledge will likely opt for intermediate values. As the proposed estimator of self-knowledge requires panel data, which is not available for all individuals in my sample, I can only estimate the between-variance, the numerator of the estimator for self-knowledge by Falk et al. (2021) which rises monotonically in self-knowledge in their model.

[16]The mean value of the inverse PBLH in *sample II* is 1.76, with a standard deviation of 0.2. The effect of a one unit increase in the inverse PBLH on PM10 concentration during the in-utero period is 4.48. Hence a one standard deviation increase in the IV raises PM10 by 0.9 $\frac{\mu g}{m}$.

inverse PBLH. Estimated coefficients are very similar in magnitude to the baseline results. The estimates are only significant at the 10% level, which likely follows from the fact that the inverse PBLH is not as strong an instrument as the frequency of nighttime inversions (F-Statistic of 18.5). Overall, I view this as a confirmation of the main results, indicating that the effect on neuroticism is not driven by unobservable variables correlated with inversion frequency.

Discussion. The main result of the analysis is that neuroticism increases by approximately 7% of a standard deviation for a one unit increase in prenatal PM10 concentration. Assessing whether it is plausible from a medical perspective to find an effect of in-utero exposure to particulate matter on neuroticism but not the other dimensions of non-cognitive ability, and discussing potential channels for this effect, is difficult given that the research on neurobiological origins of personality traits is still in a rather early stage (Allen and DeYoung, 2017). However, the existing results suggest maternal cortisol levels as a potential pathway between particulate matter exposure during pregnancy and the child's level of neuroticism. Li et al. (2017) find in a randomized experiment with college students in Shanghai that acute exposure to PM2.5 causes an increase in 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) Cortisol as potential link between air pollution and neuroticism is also in line with the result by Persson and Rossin-Slater (2018) that maternal stress during pregnancy has a causal impact on the mental health of children, including adult anxiety and depression.

To assess the magnitude of the estimated effect, a comparison to related studies examining the impact of gestational pollution exposure on later life cognitive ability is helpful. Results by Molina (2021) imply that a one $\frac{\mu g}{m^3}$ increase in PM10 concentration during the second trimester reduces cognitive ability in adulthood, measured by the Raven's test, by 2.4% of a standard deviation.[17] Sanders (2012) finds that among cohorts born during the

[17] This number is derived by combining quasi first stage- and reduced form results in Molina (2021). Hence, it should be interpreted with caution. However, for the purpose of assessing whether my results are of a comparable order of magnitude, I think computing this estimate is justified and helpful. PM10 levels

early 1980s in the US, a 10 $\frac{\mu g}{m^3}$ increase in total suspended particulate (TSP, which includes both PM10 and larger, less harmful particles) exposure during the students' year of birth reduces their high school math test scores by 6% of a standard deviation. Given an average PM10-to-TSP ratio across North America of approximately 0.5 (Cicero-Fernandez et al., 1993; Van der Meulen et al., 1987) , this implies that a 1 $\frac{\mu g}{m^3}$ increase in PM10 concentration would reduce test scores by approximately 1.2% of a standard deviation. To account for the fact that the mean age of survey respondents studied in Molina (2021) is 17, and test takers considered in Sanders (2012) are aged 15-18, whereas children in my sample are assessed between age 5 and 10, I re-scale my main 2SLS estimate (table 4, column 5) with the correlation between neuroticism in childhood and at ages 16-17 (figure 2). This implies that a 1 $\frac{\mu g}{m^3}$ increase in prenatal PM10 concentration raises neuroticism at that age range by 7% x 0.137 = 1% of a standard deviation. While my analysis is conducted in a very different setting with substantially lower baseline pollution, the result is of the same order of magnitude as the effects found on cognitive skills. Hence, non-cognitive ability is a plausible additional mechanism driving a part of the adverse effects of in-utero exposure to air pollution on labor market outcomes.

To further assess the relevance of this channel, I conduct a back-of-the-envelope calculation to approximate the earnings impact of poor air quality via the increase in neuroticism. Using survey data and a sibling fixed effects approach, Fletcher (2013) finds that a one standard deviation increase in neuroticism measured in young adulthood reduces annual earnings by 5-6%. As mentioned above, my results imply that a 1 $\frac{\mu g}{m^3}$ increase in gestational PM10 exposure raises neuroticism during young adulthood by 1% of a standard deviation. Combining these results indicates that a one standard deviation (5 $\frac{\mu g}{m^3}$) increase in prenatal PM10 concentration reduces annual earnings by .24%-.29% through its adverse effect on emotional stability. Isen et al. (2017) exploit the US clean air act and find that a reduction in early life exposure to TSP by 10 $\frac{\mu g}{m^3}$ increases adult earnings by 1%. While a comparison of results derived from different settings and time periods of course has limitations, it at least suggests that the noncognitive ability channel is of relevant size. Given that my estimates are likely attenuated due to some missing information on the place of birth, the channel could account for roughly a third of the full effect.

during the first and last trimester of pregnancy have no effect on the Raven's test score.

6. Conclusion

This paper provides causal evidence on the effect of in-utero exposure to air pollution on noncognitive ability in childhood. Using the meteorological phenomenon of thermal inversions to address the endogeneity in exposure to particulate matter, I find that exposure to PM10 during the prenatal period reduces emotional stability at age 5-10 in a sample of children born in Germany since the year 2000. In terms of magnitude, an increase in PM10 concentration by 1 $\frac{\mu g}{m^3}$ raises neuroticism by 7% of a standard deviation. The result proves robust to changes in the model specification, analysis sample and instrument. Back of the envelope computations imply that an increase in PM10 by 5 $\frac{\mu g}{m^3}$, i.e. one standard deviation, reduces adult earnings by .24%-.29% just through its impact on neuroticism.

The finding is important in light of recent evidence showing that the labour market returns to noncognitive skills have increased, relative to the returns to cognitive skills (Edin et al., 2017; Deming, 2017). This suggests that the magnitude of this channel might become even larger over time.

The results are obtained from a setting with relatively low baseline pollution levels, especially in comparison to other studies on long-run effects of in-utero exposure. Hence even at air quality levels below current limit values in developed countries, adverse long-run effects on skills, and thus most likely also later life labor market outcomes, arise.

The finding has also important policy implications in light of the fact that noncognitive ability is 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 early life exposure to air pollution on educational attainment and adult earnings. Individuals born in places during periods of poor air quality, e.g. from natural sources such as wild fires, could at least in part be compensated for these bad starting conditions by way of targeted interventions. This result might guide policy-makers in allocating scarce resources for programs fostering skill development.

Finding that the rise in neuroticism is mainly driven by increases at the upper end of the distribution suggests that in-utero exposure to air pollution might not only reduce emotional stability within the range of 'normal' variations in personality, but also induce mental health problems. This suggestive relationship between in-utero exposure to air pollution and later life mental health should be explored more in future research.

Another interesting avenues for future work would be to replicate this analysis in other settings, e.g. a developing country, to investigate the external validity of the result, or using data on adults to examine how the effect develops over the life-cycle. Lastly, one caveat of this analysis is that the instrument does not allow to disentangle the effects of different pollutants, most importantly CO and PM10. Repeating the analysis with a different instrument that overcomes this issue, e.g. changes in relevant policies, can generate important insights for policy-makers deciding about regulation of air pollutants.

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 (2018). The Skills to Pay the Bills: Returns to on-the-job Soft Skills Training. NBER Working Paper, National Bureau of Economic Research.

Akee, R., W. Copeland, E. J. Costello, and E. Simeonova (2018, March). 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*, 319–352.

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.

Annavarapu, R. N. and S. Kathi (2016). Cognitive Disorders in Children Associated with Urban Vehicular Emissions. *Environmental Pollution 208*, 74–78. Special Issue: Urban Health and Wellbeing.

Arceo, E., R. Hanna, and P. Oliva (2016, 01). 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.

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

Beckwith, T., K. Cecil, M. Altaye, R. Severs, C. Wolfe, Z. Percy, T. Maloney, K. Yolton, G. LeMasters, K. Brunst, and P. Ryan (2020, 01). Reduced Gray Matter Volume and Cortical Thickness Associated with Traffic-related Air Pollution in a Longitudinally Studied Pediatric Cohort. *PLOS ONE 15*(1), 1–19.

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.

Borghans, L., H. Meijers, and B. Ter Weel (2008). The Role of Noncognitive Skills in Explaining Cognitive Test Scores. *Economic Inquiry 46*(1), 2–12.

Carneiro, P., C. Crawford, and A. Goodman (2007). The Impact of Early Cognitive and Non-Cognitive Skills on Later Outcomes. CEE DP 92.

Cicero-Fernandez, P., W. Thistlewaite, Y. Falcon, and I. Guzman (1993). Tsp, pm10 and pm10/tsp ratios in the mexico city metropolitan area: a temporal and spatial approach. *Journal of exposure analysis and environmental epidemiology 3*, 1–14.

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*, 102511.

Colmer, J. and J. Voorheis (2020, December). The Grandkids aren't alright: The Intergenerational Effects of Prenatal Pollution Exposure. CEP Discussion Papers 1733,

Centre for Economic Performance, LSE.
Costa, L. G., T. B. Cole, J. Coburn, Y.-C. Chang, K. Dao, and P. Roque (2014). Neurotoxicants are in the Air: Convergence of Human, Animal, and in vitro Studies on the Effects of Air Pollution on the Brain. *BioMed Research International 2014*.
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. and D. Almond (2011). Chapter 15 - Human Capital Development Before Age Five. Volume 4 of *Handbook of Labor Economics*, pp. 1315–1486. Elsevier.
Currie, J. and M. Stabile (2006). "Child Mental Health and Human Capital Accumulation: The Case of ADHD". *Journal of Health Economics 25*(6), 1094–1118.
Currie, J., J. Voorheis, and R. Walker (2020). What Caused Racial Disparities in Particulate Exposure to Fall? New evidence from the Clean Air Act and Satellite-based Measures of Air Quality.
Datta Gupta, N. and M. Simonsen (2010). Non-cognitive Child Outcomes and Universal High Quality Child Care. *Journal of Public Economics 94*(1), 30–43.
De Pauw, S. S. (2017). Childhood Personality and Temperament. *The Oxford Handbook of the Five Factor Model*, 243–280.
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*(3), 351–364.
Deal, J. E., C. F. Halverson Jr, V. Havill, and R. P. Martin (2005). Temperament Factors as Longitudinal Predictors of Young Adult Personality. *Merrill-Palmer Quarterly (1982-)*, 315–334.
Deming, D. J. (2017, 06). The Growing Importance of Social Skills in the Labor Market. *The Quarterly Journal of Economics 132*(4), 1593–1640.
Edin, P.-A., P. Fredriksson, M. Nybom, and B. Öckert (2017). The Rising Return to Non-Cognitive Skill. Discussion paper, IZA.
Falk, A., T. Neuber, and P. Strack (2021). Limited Self-Knowledge and Survey Response Behavior. Discussion Paper 307, CRC TR 224.
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.
Glatter-Götz, H., P. Mohai, W. Haas, and C. Plutzar (2019, jul). Environmental Inequality in Austria: Do Inhabitants' Socioeconomic Characteristics Differ Depending on their Proximity to Industrial Polluters? *Environmental Research Letters 14*(7), 074007.
Godzinski, A. and M. S. Castillo (2021). Disentangling the Effects of Air Pollutants with many Instruments. *Journal of Environmental Economics and Management 109*, 102489.
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, E., B. Öckert, and J. Vlachos (2017). The Intergenerational Transmission of Cognitive and Noncognitive Abilities. *Journal of Human Resources 52*(4), 887–918.
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.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior. *Journal of Labor Economics 24*(3), 411–482.

Hufe, P. (2020). The Parental Wage Gap and the Development of Socio-emotional Skills in Children.

Isen, A., M. Rossin-Slater, and W. R. Walker (2017). Every Breath You Take—Every Dollar You'll Make: The Long-term Consequences of the Clean Air Act of 1970. *Journal of Political Economy 125*(3), 848–902.

Jans, J., P. Johansson, and J. P. Nilsson (2018). Economic Status, Air Quality, and Child Health: Evidence from Inversion Episodes. *Journal of Health Economics 61*, 220–232.

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), 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 (2021). Killing Prescriptions Softly: Low Emission Zones and Child Health from Birth to School. Discussion paper.

Levy, R. J. (2015). Carbon monoxide pollution and neurodevelopment: A public health concern. *Neurotoxicology and Teratology 49*, 31–40.

Li, H., J. Cai, R. Chen, Z. Zhao, Z. Ying, L. Wang, J. Chen, K. Hao, P. L. Kinney, H. Chen, and H. Kan (2017). Particulate Matter Exposure and Stress Hormone Levels: A Randomized, Double-blind, Crossover Trial of Air Purification. *Circulation 136*(7), 618–627.

Lindqvist, E. and R. Vestman (2011). The Labor Market Returns to Cognitive and Noncognitive ability: Evidence from the Swedish Enlistment. *American Economic Journal: Applied Economics 3*(1), 101–28.

Molina, T. (2021). Pollution, Ability, and Gender-specific Investment Responses to Shocks. *Journal of the European Economic Association 19*(1), 580–619.

Moutafi, J., A. Furnham, and I. Tsaousis (2006). Is the Relationship between Intelligence and Trait Neuroticism Mediated by Test Anxiety? *Personality and Individual Differences 40*(3), 587–597.

Myhre, O., M. Låg, G. D. Villanger, B. Oftedal, J. Øvrevik, J. A. Holme, H. Aase, R. E. Paulsen, A. Bal-Price, and H. Dirven (2018). Early life Exposure to Air Pollution Particulate Matter (PM) as Risk Factor for Attention Deficit/Hyperactivity Disorder (ADHD): Need for Novel Strategies for Mechanisms and Causalities. *Toxicology and Applied Pharmacology 354*, 196–214.

Persson, P. and M. Rossin-Slater (2018). Family Ruptures, Stress, and the Mental Health of the Next Generation. *American Economic Review 108*(4-5), 1214–52.

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.

SOEP (2019). Socio-Economic Panel, data for years 1984-2018. DOI: 10.5684/soep-core.v35.

Sorrenti, G., U. Zölitz, D. Ribeaud, and M. Eisner (2020). The Causal Impact of Socio-emotional Skills Training on Educational Success. Working Paper 343, University of Zurich, Department of Economics.

Van der Meulen, A., B. Van Elzakker, and G. Van den Hooff (1987). Pm10: results of a one-year monitoring survey in the netherlands. *JAPCA 37*(7), 812–818.

Voorheis, J. (2017). Air Quality, Human Capital Formation and the Long-term Effects of Environmental Inequality at Birth. Working paper, Center for Economic Studies, US Census Bureau.

Widiger, T. A., W. L. Gore, C. Crego, S. L. Rojas, and J. R. Oltmanns (2017). Five Factor Model and Personality Disorder. *The Oxford Handbook of the Five Factor Model*, 449.

Woodward, N., A. Haghani, R. Johnson, T. Hsu, A. Saffari, C. Sioutas, S. Kanoski, C. Finch, and T. Morgan (2018). Prenatal and early life Exposure to Air Pollution Induced Hippocampal Vascular Leakage and Impaired Neurogenesis in Association with Behavioral Deficits. *Translational Psychiatry 8*(1), 1–10.

Xu, X., S. U. Ha, and R. Basnet (2016). A Review of Epidemiological Research on Adverse Neurological Effects of Exposure to Ambient Air Pollution. *Frontiers in Public Health 4*, 157.

A. Appendix

Table A.1: Big Five in the SOEP Mother-and-child-questionnaires

How would you rank your child in comparison to other children of the same age?
My child is ...

2-3 years	O	quick at learning new things - needs more time
	C	focused - easily distracted
	E	shy - outgoing
	A	obstinate - obedient
	N	-
5-6 years 9 -10 years	O	understands quickly – needs more time
		not that interested – hungry for knowledge
	C	focused - easily distracted
		tidy – untidy
	E	talkative – quiet
		withdrawn – sociable
	A	obstinate - compliant
		good-natured – irritable
	N	self-confident – insecure
		fearful – fearless

Table A.2: **Impact of prenatal PM10 exposure on the Big Five: Outcomes on Sample II**

	Openness	Conscien.'ness	Extraversion	Agreeableness
Panel A: 2SLS				
PM10	-.0398	.0179	-.0228	.0337
in utero	(.0274)	(.0310)	(.0280)	(.0315)
Observations	4,887	4,902	4,879	4,853
Counties	121	122	121	120
First Stage F-Stat	79.6	80.4	80.5	78.0
Panel B: Reduced Form				
Inversions	-.0458	-.0257	-.0171	.0411
in utero	(.0291)	(.0316)	(.0314)	(.0355)
Observations	6,541	6,534	6,532	6,517
Counties	155	155	155	155

Note: Panel A displays 2SLS-estimates of a 1 $\frac{\mu g}{m^3}$ increase in PM10 concentration during the in-utero period on Big Five personality traits, estimated on *sample II*, i.e. children assessed at ages 5-6 or 9-10. PM10 is instrumented by inversion frequency Panel B shows the OLS estimates of the effect of an increase in share of days with a nighttime inversion during the in-utero period by one standard deviation. Outcomes are age-standardized. All regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, controls for parental education, child gender, migration background, age and age squared, cubic functions of temperature, wind speed and precipitation, plus one lead and one lag of the instrument. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table A.3: Impact of postnatal PM10 exposure on the Big Five

	Openness	Consc.'ness	Extraversion	Agreeableness	Neuroticism
Panel A: 2SLS					
PM10	.0256	.0017	.0096	.0202	-.0251
post birth	(.0224)	(.0242)	(.0219)	(.0263)	(.0283)
Observations	8,081	8,070	8,074	8,016	5,354
Counties	169	169	169	167	131
1st Stage F-Stat.	108.2	107.7	108.4	106.8	87.6
Panel B: Reduced Form					
Inversion Frequency	.0397	-.00002	.0104	.0316	-.0181
post birth	(.0263)	(.0241)	(.0240)	(.0286)	(.0346)
Observations	9,454	9,440	9,445	9,423	6,548
Counties	192	192	192	192	155

Note: Panel A displays 2SLS-estimates of the effect of a 1 $\frac{\mu g}{m^3}$ increase in PM10 concentration during the nine-month period following the month of birth on Big Five personality traits. PM10 is instrumented by inversion frequency. Panel B shows the OLS estimates of the effect of an increase in the share of days with a nighttime inversion during the postnatal period by one standard deviation (+7%) on the outcomes. Outcomes are age-standardized. All regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, controls for parental education, child gender, migration background, age and age squared, cubic functions of temperature, wind speed and precipitation during both the postnatal and prenatal period, plus one lead and one lag of the instrument. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table A.4: Robustness - Sample Construction

	Dependent Variable: Neuroticism				
	(1)	(2)	(3)	(4)	(5)
Panel A					
PM10	.0714**	.0665**	.0888***	.0823***	.0725**
in utero	(.0314)	(.0281)	(.0343)	(.0289)	(.0294)
Minimum obs. per county	20	15	25	20	20
Obs. per individual	1	1	1	all	all
Weights	x	x	x	x	✓
Observations	4,911	5,600	4,293	6,572	6,572
Counties	122	163	94	122	122
1st Stage F-Stat.	80.7	106.9	63.2	89.0	81.0
Panel B					
Inversions	.0667**	.0697**	.0887***	.0768**	.0632**
in utero	(.0323)	(.0307)	(.0335)	(.0300)	(.0304)
Minimum obs. per county	20	15	25	20	20
Obs. per individual	1	1	1	all	all
Weights	x	x	x	x	✓
Observations	6,548	7,230	5,747	8,497	8,497
Counties	155	196	118	155	155

Note: The table displays 2SLS-estimates of the effect of a 1 $\frac{\mu g}{m^3}$ increase in in-utero PM10 exposure on Neuroticism (Panel A) and reduced form results (Panel B). The outcome is standardized within age groups. Column 1 replicates the baseline results. Columns 2 and 3 vary the minimum number of individuals required within a county. In columns 4 and 5 multiple observations for the same individual are included, if available. In column 5 observations are weighted by the inverse of the number of observations per individual. All regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, controls for parental education, child gender, migration background, age in months and its square, cubic functions of temperature, wind speed and precipitation, plus one lead and one lag of the instrument. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table A.5: Robustness - Alternative Specifications

	Dependent Variable: Neuroticism			
	(1)	(2)	(3)	(4)
Panel A				
PM10	.0714**	.0666**	.0726**	.0771**
in utero	(.0314)	(.0313)	(.0330)	(.0357)
Specification	*baseline*	*extended controls*	*no leads & lags*	*2 leads & lags*
Observations	4,911	4,882	4,911	4,911
Counties	122	122	122	122
1st Stage F-Stat.	80.7	81.3	96.2	43.5
Panel B				
Inversions	.0667**	.0640*	.0674**	.0841**
in utero	(.0323)	(.0324)	(.0326)	(.0352)
Specification	*baseline*	*extended controls*	*no leads & lags*	*2 leads & lags*
Observations	6,548	6,496	6,548	6,415
Counties	155	154	155	152

Note: The table displays 2SLS-estimates of the effect of a 1 $\frac{\mu g}{m^3}$ increase in in-utero PM10 exposure on Neuroticism (Panel A) and reduced form results (Panel B). The outcome is standardized within age groups. Column 1 replicates the baseline results. Column 2 adds dummies for single parent household, birth order and maternal age. Columns 3 and 4 vary the number of leads and lags of the instrument included in the regression. All regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, controls for parental education, child gender, migration background, age in months and its square, cubic functions of temperature, wind speed and precipitation. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table A.6: Robustness - Uninformative Anwers

	Dependent Variable: Neuroticism			
	(1)	(2)	(3)	(4)
Panel A				
PM10	.0714**	.0698**	.0717**	.0811**
in utero	(.0314)	(.0309)	(.0336)	(.0348)
Sample	*baseline*	*drop answers with many "fives"*	*drop bottom 5% (variance in answers)*	*drop bottom 10% (variance in answers)*
Observations	4,911	4,882	4,453	4,125
Counties	122	122	113	108
1st Stage F-Stat.	80.7	83.5	75.4	68.9
Panel B				
Inversions	.0667**	.0657**	.0696**	.0687*
in utero	(.0323)	(.0325)	(.0351)	(.0383)
Sample	*baseline*	*drop answers with many "fives"*	*drop bottom 5% (variance in answers)*	*drop bottom 10% (variance in answers)*
Observations	6,548	6,512	5,943	5,453
Counties	155	155	144	134

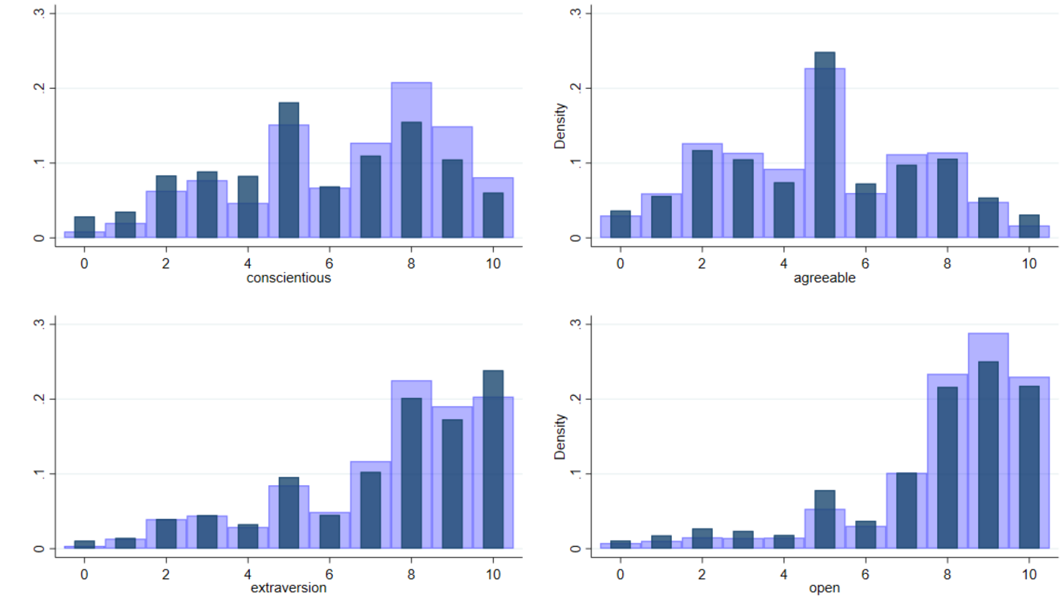
Note: The table displays 2SLS-estimates of the effect of a 1 $\frac{\mu g}{m^3}$ increase in in-utero PM10 exposure on Neuroticism (Panel A) and reduced form results (Panel B). The outcome is standardized within age groups. Column 1 replicates the baseline results. In column 2 observations for which all items underlying Neuroticism and Conscientiousness have a value of 5 are dropped. In columns 3 and 4 the 5% or 10% of observations with the lowest variance across items underlying the Big Five in the SOEP are dropped. All regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, controls for parental education, child gender, migration background, age in months and its square, cubic functions of temperature, wind speed and precipitation, plus one lead and lag of the instrument. Standard errors clustered at the county level are reported in parentheses. *p<0.1; **p<0.05; ***p<0.01

Table A.7: Robustness - Alternative IV

	Dependent Variable: Neuroticism	
	(1)	(2)
PM10	.0947*	
in utero	(.0552)	
Inverse PBLH		.0712*
in utero		(.0403)
Observations	4,911	6,548
Counties	122	122
1st Stage F-Stat.	18.5	
1st Stage Effect ($\hat{\alpha}$)	4.48	
Inverse PBLH	Mean	St Dev.
	1.76	.20

Note: Column 1 displays results from 2SLS estimation of the effect of a 1 $\frac{\mu g}{m^3}$ increase in in-utero PM10 exposure on Neuroticism, using the inverse planetary boundary layer height as an instrumental variable. Column 2 reports results from the reduced form model, multiplied by .2 to reflect the effect of a one standard deviation increase in the inverse PBLH. The outcome is age-standardized Neuroticism. Regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, controls for parental education, child gender, migration background, age in months and its square, cubic functions of temperature, wind speed and precipitation, plus one lead and one lag of the instrument. Standard errors clustered at the county level are reported in parentheses. The bottom part shows summary statistics for the instrumental variable. *p<0.1; **p<0.05; ***p<0.01

Panel A: 2- to 3-year-olds



Panel B: 5- to 6- and 9- to 10-year-olds

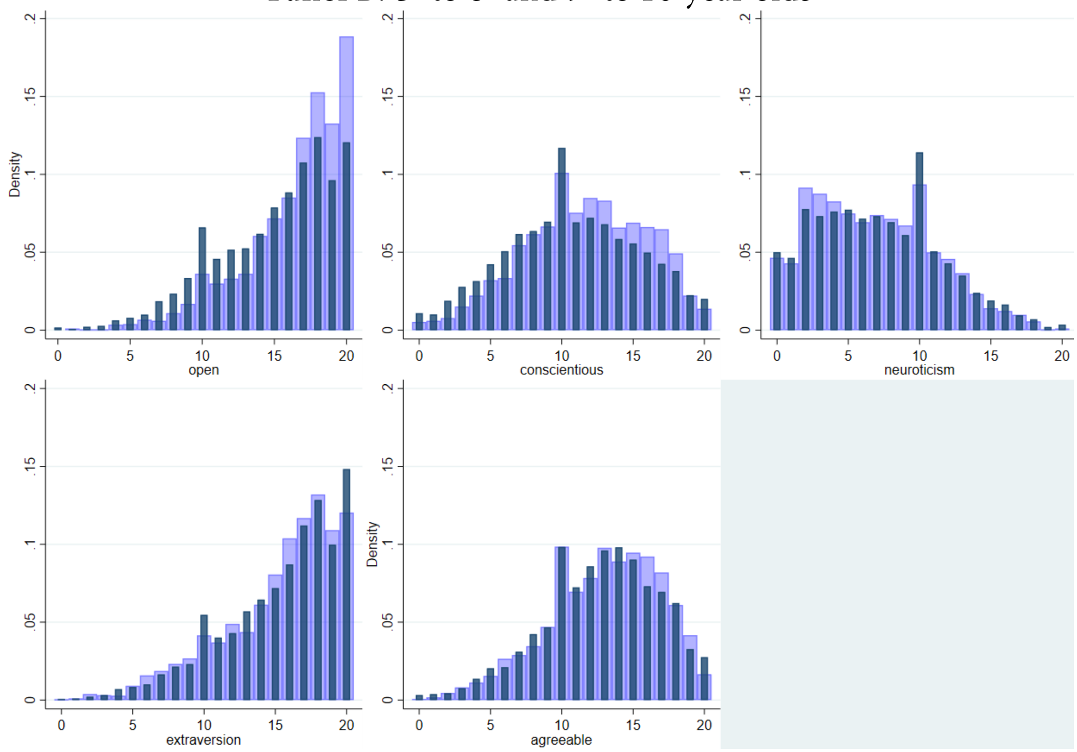


Figure A.1: Distribution of Big Five personality traits by maternal education.

Note: The distributions are based on all available observations in the SOEP. Light blue [Dark blue] bars represent the distribution among children whose mother holds a [holds no] tertiary degree. Panel A: 2,185 answers by highly-educated mothers, 5,803 by less-educated mothers. Panel B: 2,850 answers by highly-educated mothers, 8,518 by less-educated mothers.

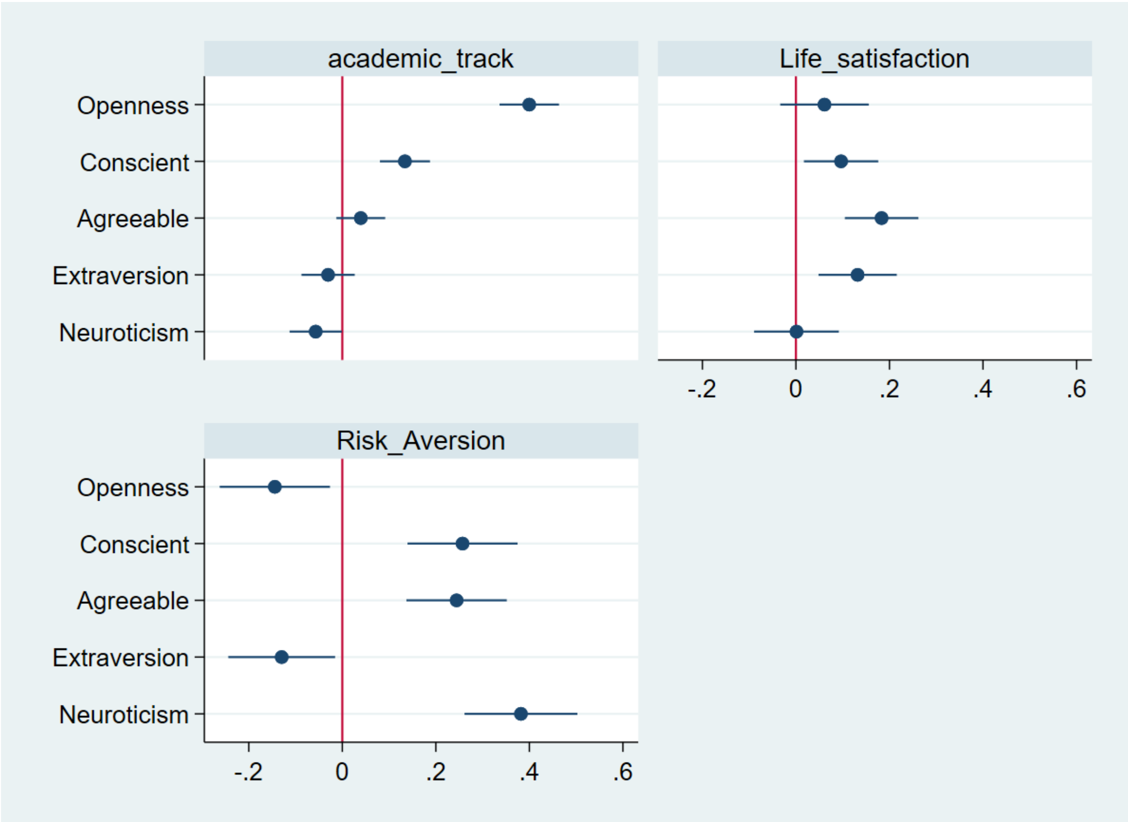


Figure A.2: Non-cognitive abilities and child outcomes

Note: The figures show partial correlations between mother-assessed Big Five and other child outcomes measured in the SOEP, after controlling for parental education, a single-parent household dummy, child gender and migration background. Panel A: Correlations between standardized mother-reported Big Five when the child is aged 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 to 8 years old. Sample size: 3,842. Panels B and C depict correlations between standardized mother-reported Big Five when the child is aged 9-10 years and child-reported life satisfaction and risk aversion at age 11-12, respectively. Both are measured on an 11-point scale. Sample Sizes: 2,210 (Panel B) and 2,190 (Panel C). 95%-Confidence Intervalls are based on robust standard errors. Based on data from SOEP, version 35.

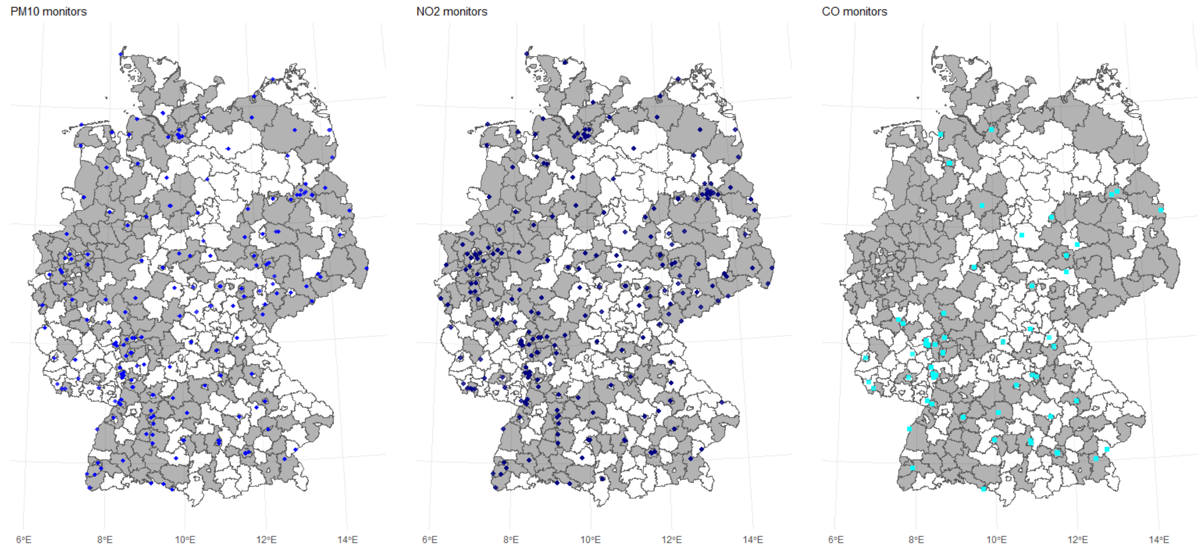


Figure A.3: Sample Counties and pollution monitors

Note: Grey shaded counties are those included in *Sample I*. Based on data from UBA and SOEP, version 35. Dots represent pollution monitors which were active throughout the sample period and are used in the analysis. Blue dots on the left map represent PM10 monitors, dark blue dots in the middle map represent NO2 monitors, light blue dots on the right map represent CO monitors.

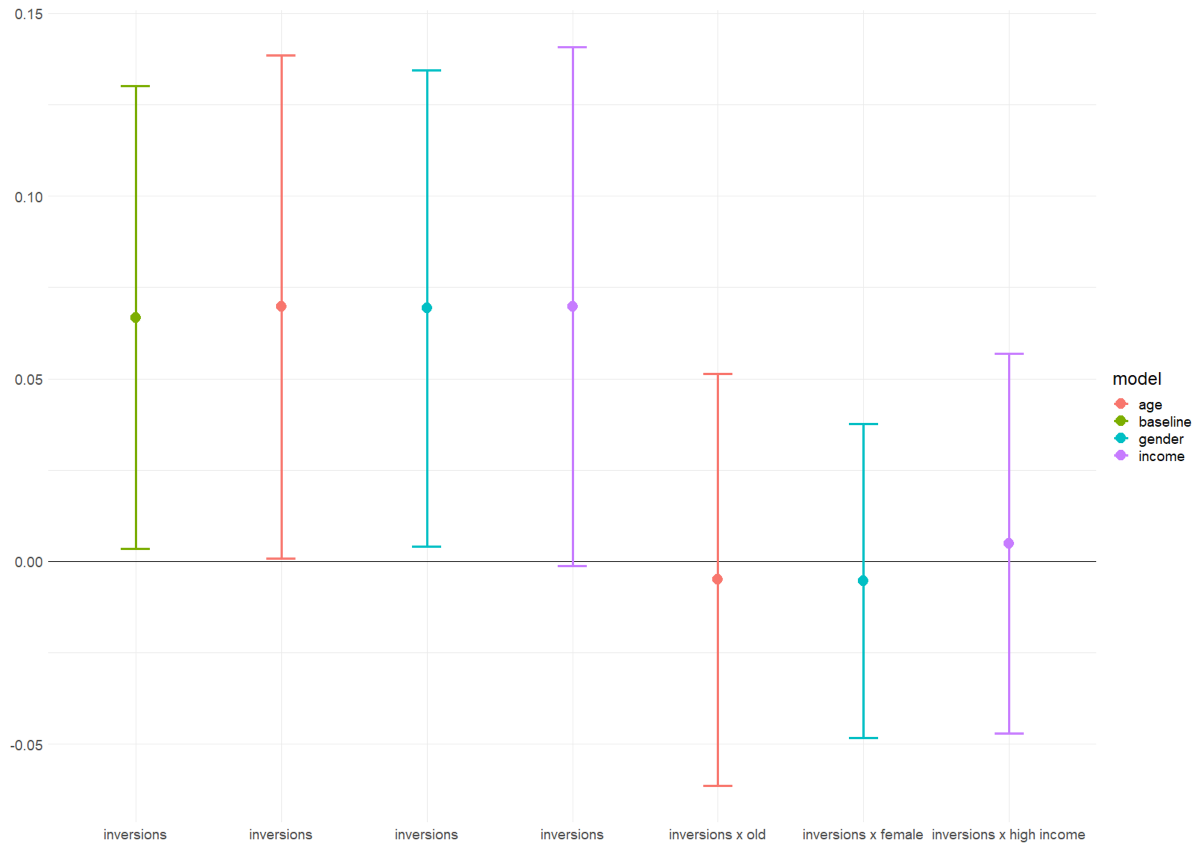


Figure A.4: Effect of thermal inversions on Neuroticism: Heterogeneity

Note: The figure depicts estimated coefficients from OLS regressions of Neuroticism on the share of days with a nighttime inversion during the in-utero period and interactions between inversions and child characteristics. Colors represent different regression models. Green: Baseline model (cf. table 4, Panel B, column 5). Red: Baseline model + interaction btw. inversions and dummy indicating age group 9-10. Blue: Baseline model + interaction btw. inversions and dummy indicating female child. Purple: Baseline model + indicator for above median income + interaction btw. inversions and high income indicator. All regressions include year-, month- and county-of-birth fixed effects, age group fixed effects, controls for parental education, child gender, migration background, age and age squared, cubic functions of temperature, wind speed and precipitation, plus one lead and one lag of the instrument. Bars represent 95%-confidence intervals, based on standard errors clustered at the county level.