

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Andersen, Christina M. et al.

# Working Paper Air Pollution and Cognition in Children: Evidence from National Tests in Denmark

CESifo Working Paper, No. 11434

**Provided in Cooperation with:** Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

*Suggested Citation:* Andersen, Christina M. et al. (2024) : Air Pollution and Cognition in Children: Evidence from National Tests in Denmark, CESifo Working Paper, No. 11434, CESifo GmbH, Munich

This Version is available at: https://hdl.handle.net/10419/307364

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



# WWW.ECONSTOR.EU



# Air Pollution and Cognition in Children: Evidence from National Tests in Denmark

Christina M. Andersen, Jørgen Brandt, Jesper H. Christensen, Lise M. Frohn, Camilla Geels, Timo Hener, Marianne Simonsen, Lars Skipper



### Impressum:

CESifo Working Papers ISSN 2364-1428 (electronic version) Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute Poschingerstr. 5, 81679 Munich, Germany Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de Editor: Clemens Fuest https://www.cesifo.org/en/wp An electronic version of the paper may be downloaded • from the SSRN website: www.SSRN.com

- from the RePEc website: <u>www.RePEc.org</u>
- from the CESifo website: <u>https://www.cesifo.org/en/wp</u>

# Air Pollution and Cognition in Children: Evidence from National Tests in Denmark

## Abstract

This paper examines the effects of daily outdoor air pollution variation on student test scores. Using Danish register data for all elementary and lower secondary students, we link home addresses to a 1 km x 1 km pollution grid to measure test day and lifetime pollution exposure. An increase in fine particles (PM2.5) from a very clean to an average day reduces math scores by 1.8% and reading by 0.9% of a standard deviation. Even at low pollution levels, student performance is harmed, especially in math. We find no evidence of heterogeneity by health, socio-economic status, or lifetime exposure.

JEL-Codes: Q530, I210, I180.

Keywords: air pollution, cognition, test scores.

Christina M. Andersen Department of Economics and Business Economics, Aarhus University / Denmark cmandersen@econ.au.dk

Jesper H. Christensen Department of Environmental Science Aarhus University / Denmark jc@envs.au.dk

Camilla Geels Department of Environmental Science Aarhus University / Denmark cag@envs.au.dk

Marianne Simonsen Department of Economics and Business Economics, Aarhus University / Denmark msimonsen@econ.au.dk Jørgen Brandt Department of Environmental Science Aarhus University / Denmark jbr@envs.au.dk

Lise M.Frohn Department of Environmental Science Aarhus University / Denmark Imf@envs.au.dk

Timo Hener\* Department of Economics and Business Economics, Aarhus University / Denmark thener@econ.au.dk

Lars Skipper Department of Economics and Business Economics, Aarhus University / Denmark lskipper@econ.au.dk

\*corresponding author

October 23, 2024

We thank seminar participants EAERE 2023, Helsinki PhD Workshop on Economics of Education 2023, Aarhus University, AERE 2024, IWAEE 2024, ESPE 2024 The Ragnar Frisch Centre for Economic Research, Statistics Norway, BI Norwegian Business School, CESifo Area Conference on Labor Economics 2024, and Free University Berlin for helpful comments. Hener thanks the Aarhus University Research Foundation for financial support. We thank the Danish National Centre for Register-based Research for granting data access.

#### 1 Introduction

Despite global efforts to reduce air pollution, which have successfully lowered exposure levels in North America and Europe (Sicard *et al.*, 2023), air pollution remains a universal externality, one from which no one is ever entirely exempt. While the health consequences of air pollution are wellestablished, there is a growing body of research focusing on its non-health-related impacts (Aguilar-Gomez *et al.*, 2022; Graff Zivin and Neidell, 2012; Isen *et al.*, 2017), including its effects on cognitive performance (Ebenstein *et al.*, 2016; Carneiro *et al.*, 2021). While much of this research centers on adults, children are particularly vulnerable to air pollution due to their outdoor activities and a potentially heightened sensitivity during critical developmental periods. Understanding the effects of air pollution on children is essential, as childhood is widely recognized as a the formative years for human capital development (Cunha and Heckman, 2007). However, there is limited knowledge about how air pollution impacts children's cognition, particularly at lower pollution levels commonly found in high-income countries, and how these effects might interact with individual predispositions.

To shed light on these questions, our paper studies the effects of daily variation in outdoor air pollution on test scores of primary and lower secondary school students. We use data from Denmark, which allows us to combine information about local levels of ambient air pollution, student test scores, and population-wide registration data. To construct the pollution treatment, we observe the place of residence of each student and match the daily concentration of fine inhalable particles  $(PM_{2.5})$  from an atmospheric model with a resolution of 1 km by 1 km. Crucially for identifying even small treatment effects at low levels of pollution below the WHO guideline<sup>1</sup> limits, our individuallevel administrative data contains information about student performance on 2.6 million repeated national tests in math and reading between grades 2 to 8 from 2010 to 2018. Using the large sample, we are also able to distinguish the impacts along the distribution of test scores and between different test domains, shedding light on the cognitive processes involved. To investigate treatment effect heterogeneity, we access comprehensive health data and are, therefore, the first to speak to the role of underlying health conditions based on pre-existing diagnoses such as asthma, acute, and chronic respiratory diseases. While it is likely that some health conditions may amplify the effects of exposure to pollution, we currently know little about which conditions and the degree to which they matter.

To circumvent a bias from non-random selection of neighborhoods, our main analysis exploits

 $<sup>^1 \</sup>rm{Fine}$  inhalable particles (PM\_{2.5}) below 15  $\mu g/m3.$ 

that students from the same family take several tests across days with varying levels of pollution. We probe the robustness of our results by including fixed effect at the individual instead of the family level, and by instrumenting pollution with local wind directions (Deryugina *et al.*, 2019). Here, the exclusion restriction postulates that no direct or indirect effect of the direction of wind, conditional on weather controls, may affect the outcome other than the effect through air pollution.

We find that a 10  $\mu g/m^3$  increase in fine inhalable particles (PM<sub>2.5</sub>) leads to a reduction in student performance in math by 1.8 percent and in reading by 0.9 percent of a standard deviation. We document that these effects are even present at pollution levels below the recommended safe limit from the WHO and appear to be linear in levels of pollution. Our IV results that exploit municipality-specific wind directions are comparable in magnitude and statistically significant, bolstering the causal interpretation of the effect of air pollution on test scores.

Across the test score distribution, we find that the effects of air pollution increase with the proficiency of the students in math, the effects being approximately 50% larger at the top of the math test score distribution compared to the average impact. No such gradient is visible for the results in reading. To further scrutinize this result, we split the math and reading outcomes into three test domains each: numbers and algebra, geometry, and applications for math, and language comprehension, decoding, and text comprehension for reading. The pattern with increasing effects over the quantiles of the distribution is most pronounced for geometry in math tests that is related to higher order logical reasoning and judgement skills (Spiller *et al.*, 2023). All reading domains, in contrast, have flat profiles.

The impact of air pollution is often assumed to depend on health-related susceptibility and to trigger a multitude of correlated health problems. However, it is still an open question whether the pathways from exposure to cognition are in fact related to other health problems.<sup>2</sup> We observe the entire health histories of all individuals and test for heterogeneity of the effects by prior respiratory diseases including diagnoses of asthma earlier in life. While the health issues are highly predictive of test scores, they do not significantly moderate the effect of air pollution. We further test for habituation effects of air pollution by interacting the same day pollution with cumulative pollution measures over the lifetime and find no heterogeneous effects along this dimension. Moreover, we do not find strong evidence for heterogeneity by mothers' education, but we show that the impact on test scores is twice as large for girls as for boys.

<sup>&</sup>lt;sup>2</sup>The development of mental disorders has also been associated with air pollution exposure (Antonsen *et al.*, 2020; Khan *et al.*, 2019; Thygesen *et al.*, 2020).

We test the robustness of our results across a variety of specifications. For example, our results are robust to different levels of fixed effects, including individual student fixed effects, and to the inclusion of covariates. We also test whether the results can be replicated in a high-stake exam environment, using *Gymnasium* exit exams as the outcome. We also observe large effects of pollution exposure in this setting: a 10-point increase in  $PM_{2.5}$  decreases exam grades in math by 8.7% of a standard deviation and in Danish by 4.6% of a standard deviation.

Our results contribute directly to a recent and growing literature that connects daily variation in air pollution with performance in exams and assessments in education. In seminal work, Ebenstein et al. (2016) rely on data from Israeli high school exit exams, the Bagrut matriculation, and exploit the fact that the same student takes multiple obligatory tests on days with varying levels of air pollution. The authors document clear and harmful effects of air pollution; a one standard deviation increase in fine inhalable particles (PM<sub>2.5</sub>) leads to a reduction in student performance of 3.9 percent of a standard deviation. In a related paper, Carneiro et al. (2021) combine individuallevel panel data from university entrance exams from Brazil (São Paolo and Rio de Janeiro) with information on local wind conditions to show that a one standard deviation increase in inhalable particles ( $PM_{10}$ ) leads to a reduction in student performance of 8 percent of a standard deviation. Zivin et al. (2020) show that the scores of the college entrance exams in China are significantly decreased due to the downwind air pollution of agricultural fires, with stronger effects in highachieving students. As a rare example of an analysis of indoor air pollution, Roth (2020) finds for London-based university students that particulate matter exposure ( $PM_{10}$ ) reduces their scores in high-stakes exams.

Interestingly in light of the indications from our domain-wise results, a related literature speaks to the cognitive processes involved in explaining the drops in test performance from air pollution exposure. La Nauze and Severnini (2021) use data from brain-training games and show that  $PM_{2.5}$  exposure decreases adult cognitive functioning in the memory and problem solving domains. Stronger results for younger adults are consistent with fluid intelligence, the part that is independent of accumulated knowledge, being more affected by air pollution. Krebs and Luechinger (2021) use data from a brain training game in which players solve arithmetic problems and find that nitrogen oxide pollution  $(NO_x)$  reduces the scores of experienced players. Using laboratory experiments with university students in Brazil, Bedi *et al.* (2021) show that  $PM_{2.5}$  reduces performance in a grammatical reasoning test. The test was designed to involve higher mental processes and relies on fluid intelligence, which predicts educational outcomes and performance in cognitively demanding occupations.

A few recent papers have analyzed the impacts of other measures of air pollution on test scores and with varying lengths of exposure and time horizons. Zhang et al. (2018) use verbal and math test results from a large longitudinal survey in China for respondents of all ages over 9 to estimate the effects of an air pollution index (API), which contains measures of sulfur dioxide  $(SO_2)$ , nitrogen dioxide  $(NO_2)$ , and particulate matter  $(PM_{10})$ . They find that the API affects only verbal tests negatively in the short term, but both tests in the longer term, when regressed on mean API for the previous 90 days through 3 years. In a recent paper, Halliday et al. (2022) estimate the impact of VOG, the air pollutant compounds from volcanic eruptions, on test scores of students in primary and lower secondary school in Hawaii. They find a 1.1 percent of a standard deviation decrease in test scores from a one standard deviation increase in particulate pollution, with the effects concentrated among economically disadvantaged students. To identify causal effects of more long-term pollution exposure, Heissel et al. (2022) exploit the changes in pollution from school transitions in Florida. They find evidence of a medium-term effect of air pollution on test scores of students in middle school and high school. Similarly, Bharadwaj et al. (2017), use sibling comparisons to document negative effects of fetal pollution exposure on 4th grade test scores in Chile. Finally, our work also speaks to a broader literature that is concerned with the effects of air pollution on health (e.g., Currie and Neidell (2005); Schlenker and Walker (2016); Deryugina et al. (2019)), and on labor productivity (e.g., Graff Zivin and Neidell (2012); Isen et al. (2017); Chang et al. (2016, 2019)).

Our results complement the existing evidence and advance it in several ways. First, while most of the existing evidence on the cognition effects of air pollution stems from high-pollution environments (Ebenstein *et al.*, 2016; Carneiro *et al.*, 2021; Zhang *et al.*, 2018), we show that particulate matter exposure is harmful, even at levels below widely used threshold values. In fact, the marginal effects of changes close to zero pollution are hardly any different from impacts at higher levels. This result has important policy implications, as it informs debates about safe levels of air pollution exposure and policies that postulate threshold values, the violation of which invokes reduction efforts. Second, our results indicate that the negative impacts of air pollution exposure on cognition are not confined to adults or situations with high stakes involved. Instead, the impacts extend far into childhood and are visible in cognitive tests of primary school students. Third, there is a lack of evidence on which pathophysiological mechanisms are involved (see Aguilar-Gomez *et al.* (2022) for an overview), an area which we contribute to by showing evidence that the top part of the distribution and, in particular, hard problems in geometry show the largest negative effects. This result is consistent with higher mental processes being directly affected by air pollution, bridging the gap between impacts on test scores and on the evidence on cognitive functioning (Bedi *et al.*, 2021; La Nauze and Severnini, 2021; Krebs and Luechinger, 2021). Fourth, there has been speculation about the susceptibility to air pollution being related to poor health conditions based on the medical literature, while direct evidence has been lacking in the context of cognition impacts. For example, Ebenstein *et al.* (2016) hypothesize that larger effects for boys and low socio-economic status students may be due to higher rates of asthma in these groups, while we do not find differential impacts using data on diagnosed asthma cases. Moreover, even common respiratory diseases are not moderating the impact of air pollution. These results caution against transferring evidence of differences in vulnerability from impacts on health domains (Deryugina *et al.*, 2019) over to impacts on productivity and cognition.

The paper proceeds as follows: Section 2 describes the institutional background for our study. Section 3 presents our data and empirical strategy. Section 4 shows our results, section 5 concludes.

#### 2 Background

#### 2.1 The Danish educational system

The Danish schooling system is characterized by a universal two-tier structure. Primary and lower secondary education, typically spanning ten years from the age of six to sixteen, is conducted in the same elementary school (folkeskole). These schools are typically public, while a private school alternative exists that follows the same structure and receives comparable state funding. The upper secondary education is divided into an academic (*Gymnasium*) and vocational track, where more than 70% of students choose the academic track.

During elementary school, students undergo a number of assessments, the national test (Den Nationale Test), that do not enter their grades. The primary objective of the national test is to collect data on students' proficiency levels in key subjects, including Danish and mathematics. The test enables educators and policymakers to evaluate the overall quality of education and to identify areas of strength and weakness in education and across schools by providing a comprehensive assessment of students' academic skills. The tests are administered to students between the 2nd and 8th grade levels. The tests are mandatory for all students in public primary and lower secondary schools and entire classrooms take the tests at the same time during the school day. The national

tests are carried out between February and April every year.

The national test is a computer-based assessment that incorporates multiple-choice questions and tasks and each test takes at least 45 minutes. The items are aligned with the national curriculum objectives and learning goals. The tests are conducted online in a self-scoring and adaptive program. The adaptivity of the test implies that the program monitors the proficiency of the student during the test and adjusts the difficulty level of the subsequent questions. This procedure ensures a more accurate assessment of proficiency than a linear test. Since the test score is evaluated in the computer-assisted program, teachers are not involved in picking questions or assessing the students' performance, ensuring that the results are comparable among students at the national level.

In practice, the test assesses the student's skills across various subdomains or profile areas. In math, these are numbers and algebra (knowledge about magnitudes and properties of numbers as well as computational strategies); geometry (knowledge about shapes of objects and spatial relationships); and statistics (the concept of uncertainty and analysis of data). In reading, the subdomains are language comprehension (the meaning of words); decoding (word recognition); and reading comprehension (the content of the text).For a detailed description of the test program, see Beuchert and Nandrup (2018).

After completing lower secondary education, students can apply to either a *Gymnasium* or a vocational track. In this study, we focus on grades from *Gymnasium* exit exams along with the test scores of the compulsory tests. Each *Gymnasium* offers a three-year program that provides direct access to university education and other post-secondary programs (conditional upon performance) and has multiple tracks that the students can apply to. All subjects can be taken on three different academic levels, where the student will have the subject for one year for the lowest level, and for three years for the most advanced level. It is mandatory to follow Danish for three years, while the students can choose between following mathematics for either one, two or three years. In the Danish system, it is mandatory for the students to take a written exam in Danish, whereas other exams are drawn at random. All written exams are taken in the end of May or beginning of June each year, and all students are taking the exam at the same time. In contrast to the national test, all students are asked the same questions in the written *Gymnasium* exit exams. The exam in written Danish lasts five hours, whereas the exam in math lasts between three and five hours depending on the level.

#### **3** Data and Estimation strategy

In this section, we first introduce our data for the analysis, which combines the Danish register data, air pollution data, and weather data. We then lay out the main estimation strategy and explain how we instrument air pollution using information on wind conditions.

#### 3.1 Data and Sample

We combine individual-level register data with air pollution and weather data to build a rich data set that allows us to estimate the effect of at-home air pollution on test scores in school.

Our primary outcome data are based on information on compulsory test scores from the Danish National Tests provided by the Danish Agency for IT and Learning (*Styrelsen for IT og Læring*). We use the reading tests in the 2nd, 4th, 6th, and 8th grades and the math tests in the 3rd and 6th grades. Our data comprise the universe of reading and math tests conducted from 2010 to 2018, amounting to 2,598,119 test results in total. In practice, we construct standardized test scores by grade level, year, and subject as our main outcome variable. As a secondary outcome we use grades from written *Gymnasium* exit exams in math and Danish. As for the national test we construct standardized test scores as our outcome.

From the national test register, we have access to unique (encrypted) individual identifiers making it possible for us to link all (100 %) test-takers to administrative register from Statistics Denmark. This step allows us to add a rich set of demographic, socio-economic, and health information to the students. First, we include basic demographics of the students including age and place of living, and we identify the students' siblings and parents via mother and father identifiers. Second, from earnings and education registers, we add information about parental income and highest educational attainment to infer the socio-economic status of the students. Third, we use the National Patient Register to describe the medical history of the students by adding all medical diagnoses from hospitals. As air pollution leads to respiratory (Dockery and Pope, 1994; Ward, 2015) and cardiovascular disease (Brook *et al.*, 2009; Dominici *et al.*, 2006; Langrish *et al.*, 2013), disease onsets are potentially important mediators and moderators. The contemporaneous disease may carry a direct health impact on cognitive performance, while previous medical conditions may increase the susceptibility of students to contemporaneous pollution exposure. We use the diagnoses to construct measures of early respiratory disease, including asthma.

Our measure of pollution exposure is the daily concentration of fine particulate matter  $PM_{2.5}$ .

Pollution concentrations are provided by a high-resolution modeling system at 1 km x 1 km resolution from 1980 to 2018. A hemispheric model (DEHM) and an urban background model (UBM) feed the modeling system that uses inputs from emissions and weather and calculates atmospheric transport, chemical transformation and deposition of 80 chemical species, where  $PM_{2.5}$  consists of primary emitted particles, secondary inorganic aerosols, secondary organic aerosols, and sea salt (Brandt *et al.*, 2001, 2003, 2012). We match the register information on addresses to the 1 km x 1 km grid cells and thus obtain daily pollution concentrations at the current home address. For each student, we observe all addresses from birth and onwards and use these to compute the cumulative pollution exposure. Auxiliary pollutants — sulfur dioxide (SO<sub>2</sub>), carbon monoxide (CO), and nitrogen dioxide (NO<sub>2</sub>) — are matched in the same manner. Figure 1 show the distribution average  $PM_{2.5}$  levels across Danish municipalities.





*Notes*: The figure shows the distribution of  $PM_{2.5}$  in our estimation sample as average deviation from the national mean for municipalities in Denmark.

Daily weather measurements come from Denmark's Meteorological Institute (DMI). We use data from 102 monitoring stations covering all major mainland and islands. Municipalities without available data are assigned the weather from the neighboring municipalities or the average within a grouping of municipalities that reflect common maritime influences (see Figure A1 for the grouping of municipalities). Precipitation, temperature, humidity, and cloud cover act as control variables. Temperature is measured in degrees Celsius, precipitation in millimeters, and humidity and cloud cover are measured as percentage points. Wind directions are used for the construction of the instrumental variables.

The summary statistics for our sample are presented in Table 1. Panel A reports sample means for our test dates for both the compulsory tests and *Gymnasium* exit exams, while Panel B reports sample means at the student-level. For the compulsory test, our sample consists of 2,598,731 tests, while our sample for *Gymnasium* exit exams consists of 457,245 exams. The reason behind the large difference in these numbers is that not everyone in a cohort enrolls in *Gymnasium* and that *Gymnasium* exit exams are only taken once per student, while the compulsory tests are taken at multiple grade levels. For each test and exam, we observe the pollution on the day of the test/exam (PM<sub>2.5</sub>, NO<sub>2</sub>, SO<sub>2</sub>, and CO), and the average temperature, precipitation, humidity, and percentage cloud cover. Panel A shows that, on average, children face levels of PM<sub>2.5</sub> of 9.9  $\mu g/m^3$ for the compulsory test and 7.9  $\mu g/m^3$  for the *Gymnasium* exit exams. It also indicates that the temperatures are generally higher on the days when students participate in *Gymnasium* exit exams compared to the compulsory test.

In Panel B of Table 1, we report student-level means, which include demographic information. Half of included students in the compulsory tests are females, whereas the share of females is higher in *Gymnasium* exit exams, because of their higher enrollment rate. Additionally, for both compulsory school tests and *Gymnasium* exit exams, the share of native Danes is just above 90%. Finally, most students live with both parents during primary and lower secondary education.

#### 3.2 Empirical Strategy

The goal of our paper is to estimate the effect of daily outdoor air pollution on student test scores and our baseline empirical model is the following:

$$TS_{it} = \delta P M_{2.5,it} + X'_{it}\beta + \phi_t + \nu_f + u_{it},\tag{1}$$

where  $TS_{it}$  is the standardized test score for math or reading from a test that individual *i* of family f takes on date *t*. The variable of interest is contemporaneous pollution  $PM_{2.5,it}$ , which is measured in  $\mu g/m^3$ . The coefficient  $\delta$  shows the linear effect of air pollution on test scores in percent of a standard deviation of the outcome.

As there is potential self-selection into high or low pollution areas that correlates with test

	Compul	sorv test	High sch	lool exam
	Mean	Std	Mean	Std
Panel A. Environmental data				
Pollution Exposure at test day				
$PM_{2.5}$	9.9	(6.1)	7.9	(2.8)
$NO_2$	10.8	(7.3)	10.8	(5.2)
$SO_2$	1.9	(2.1)	1.3	(1.2)
CO	157.9	(37.2)	127.1	(14.1)
Previous exposure				
$PM_{2.5}$ in-utero	11.9	(1.7)	14.3	(2.2)
$PM_{2.5}$ in first three years of life	11.8	(1.6)	13.8	(2.1)
$PM_{2.5}$ before test	10.2	(1.9)	12.5	(1.6)
Climate controls at test day				
Temperature (degrees Celsius)	6.9	(4.4)	14.1	(3.6)
Precipitation (cm)	0.1	(0.2)	0.0	(0.1)
Humidity (percentages)	80.6	(9.8)	75.0	(10.8)
Cloud cover (percentages)	50.4	(30.2)	38.3	(31.1)
Panel B. Student-level data				
Female	0.49	(0.50)	0.58	(0.49)
Number of children within the family	2.16	(0.81)	2.37	(20.88)
Ethnicity		. ,		· /
Native Dane	0.89	(0.31)	0.92	(0.27)
Non-Dane	0.11	(0.31)	0.08	(0.27)
Living with parents		. ,		· · · ·
Both parents	0.67	(0.47)	0.62	(0.49)
Father	0.04	(0.20)	0.06	(0.23)
Mother	0.27	(0.44)	0.24	(0.42)
Observations (Tests)	2,59	8,731	457	,245
Number of families	393	3,516	227	,414
Number of children	731	,960	276	5,700
Years	2010	-2018	2010	-2018

Table 1: 1	Descriptive	Statistics
------------	-------------	------------

*Notes:* The table reports average characteristics of the environmental data as daily averages in Panel A and averages as the student-level per student in the data in Panel B. Standard deviations are shown in parentheses.

scores, we include family fixed effects  $\nu_f$ . They account for neighborhood effects of the families' addresses and for unobserved time-constant family-specific factors like the human capital of parents, joint investments in siblings' education, or pollution avoidance investments. We further control for the student's sex and grade level in  $X_{it}$  to account for potential correlations with performance in tests. Air pollution follows predictable time patterns across years, days of the week, and time of day (early morning, morning, later) that we control for in  $\phi_t$  to account for potential correlations with performance in tests. Finally, we include a range of weather controls that potentially affect air pollution and performance at the test simultaneously. Weather controls in  $X_{it}$  are temperature, precipitation, humidity, and cloud cover. Lastly, we cluster the standard errors at the school level.

The identifying assumption for pollution in this first model is that after controlling for family and time fixed effects, demographics, and weather measures, air pollution on the day of the test is as good as random. We are thus exploiting the within-family variation in test scores between siblings and within-student variation between test iterations. The treatment variation in pollution comes from unanticipated variation between the dates when schools take the tests that any student from the same family is participating in. In general, differences in pollution come primarily from variation across dates rather than between addresses as shown in Figure 2. The left graph shows the distribution of PM<sub>2.5</sub> across all tests in our sample. The vertical line indicates the WHO limit value of 15  $\mu g/m^3$ ,<sup>3</sup> which is exceeded in rare instances. The right graph depicts the variation of air pollution within families. Even after controlling for the mean level in the family that absorbs the differences between locations, tests taken on different days show almost as much variability as the raw data.

Figure 2: Distribution of air pollution



*Notes*: The figure plots the distribution of  $PM_{2,5}$  in our estimation sample as (a) raw data, (b) residualized with family fixed effects, and (c) residualized with individual fixed effects. The vertical line in (a) depicts the café limit of the WHO.

<sup>&</sup>lt;sup>3</sup>24-hour average exposure limit.

The fixed effects specification may still suffer from date specific unobserved factors that affect pollution and test scores simultaneously. For example, a city wide festivity may increase traffic and thereby air pollution, but also distract students and suppress test performance. To avoid such biases, we propose a second estimation strategy, where we instrument air pollution with local wind directions similar to the procedure in Deryugina *et al.* (2019). The idea is that the local wind direction determines which pollutants are transported to a location. For example, the pollution produced on a highway can be transported to eastern or western neighborhoods depending on the direction of the wind (Anderson, 2020). The exclusion restriction for the instrument requires that the wind direction, after controlling for fixed effects and all other weather influences, only affects test scores via air pollution. We implement the 2SLS estimation with the first stage equation

$$PM_{2.5,it} = \sum_{m=1}^{98} \sum_{d=1}^{4} \gamma^{md} w d_{it}^{md} + X'_{it}\beta + \phi_t + \nu_f + e_{it}, \qquad (2)$$

where the endogenous air pollution is the dependent variable. The instruments are the municipalityspecific wind directions. The wind direction in municipality m on date t is recorded in four indicator variables:  $wd_{it}^{m1}$  for 1-90 degree wind,  $wd_{it}^{m2}$  for 91-180 degree wind,  $wd_{it}^{m3}$  for 181-270 degree wind,  $wd_{it}^{m4}$  for 271-360 degree wind. The indicator  $wd_{it}^{md}$  is equal to one if the wind comes from interval din any hour of the day, all are zero when there is no wind. For each wind direction and municipality combination, we identify a first stage parameter  $\gamma^{md}$ . The intuition of the instrument is that wind directions affect every location differently. Our instrument extracts the variation in air pollution that is only due to wind directions on the test date and the municipality specific response to changes in wind directions. We combine the instrumental variables strategy with our fixed effects framework from above by keeping all control variables and adding a wind speed control.

One caveat of our estimation strategy and of most other pollution studies is that the concentrations of different types of air pollutants strongly correlate with each other. It implies that the effect we estimate for fine particulate matter ( $PM_{2.5}$ ) may be overestimated as it includes the impact of all other correlated air pollutants. A relative increase in our measure should thus be understood as a relative increase of correlated pollutants (see Table B1). To address the issue, we also estimate a model for each of the criteria air pollutants as a robustness check.

#### 4 Results

This section describes all results, beginning with the main estimates of air pollution impacts on test scores. We then explore non-linearities in the effects at different levels of air pollution and investigate how the impacts vary across the distribution of test scores. Next, we will explore heterogeneous impacts of air pollution across several dimensions, including life-time pollution exposure and prior health conditions. We end with results on an alternative outcome from Gymnasium exit exams and probe the robustness of our main results.

#### 4.1 Main Findings of Pollution Effects on Test Scores

We present a visual preview of our results in Figure 3, where we show residualized test scores and residualized air pollution, using the set of controls and fixed effects described above. Each circle represents the average test scores of 1  $\mu g/m^3$  PM<sub>2.5</sub> bins. The Lowess bandsmoother estimate of the non-parametric relationship between air pollution and test scores suggests a negative effect. The reduction in test scores is fairly linear across the distribution of air pollution.



Figure 3: Scatter plot of residual air pollution and test scores

*Notes:* The figure plots the relationship between residual (a) math and (b) reading test scores and residual  $PM_{2,5}$  estimated by Lowess bandsmoother. Each observation in the plot shows residual test scores averaged over bins of width three units of residual  $PM_{2,5}$ . Residual test scores and residual  $PM_{2,5}$  are generated by regressing each variable on all fixed effects. The lowest 1% and highest 1% values for residual  $PM_{2,5}$  are left out.

We present our main regression results for the effect of  $PM_{2.5}$  on standardized test scores in math and reading in Table 2. Regressions in all columns include the full set of control variables and family fixed effects. In columns 1 and 2, we show results for the effects of contemporaneous air pollution concentrations on the day of the test. The result in column 1 suggests that a 10  $\mu g/m^3$  increase in air pollution (about the mean in our sample) decreases math test scores by 1.8% of a standard deviation. Reading test scores decrease by 0.9% of a standard deviation, and both results are statistically significant at the 1% level. To put these estimates into perspective, remember that the within-family variation in pollution levels, i.e., differences in concentrations that affect family members by pure coincidence on different days, range from -9 to +14 points. Thus, a change in air pollution by 10  $\mu g/m^3$  merely corresponds to the difference between days with good and poor air quality.<sup>4</sup>

	Fixed	Fixed effects		IV	
	Math (1)	Reading (2)	Math (3)	Reading (4)	
$PM_{2.5} * 10$	$-0.018^{***}$ (0.002)	$-0.009^{***}$ (0.003)	-0.020** (0.010)	-0.008 (0.006)	
Average $PM_{2.5}$ Observations	9.89 846 342	9.97 1 752 389	9.89 846 342	9.97 1 752 389	
First stage F	Family	Family	46.79 Family	120.6 Family	

Table 2: Impact of  $PM_{2.5}$  on test scores

Notes: The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Results in columns 1 and 2 are from fixed effects regressions, results in columns 3 and 4 are from 2SLS regressions. The instrument is constructed following Deryugina et al (2019). Wind dummies are equal to 1 if there is wind in any hour during the test date within 90 degree intervals, with no wind as the omitted category. The dummies are interacted with municipality identifiers. IV regressions additionally control for average wind speed. Across all specifications we are using family fixed effects and standard errors clustered at the school level in parentheses: \*\*p < 0.01, \*\*p < 0.05, \*p < 0.1.

Treatment effects of 1.8% and 0.9% of a standard deviation in test scores are non-negligible, but still smaller than what has been found for school exit exams and university entry exams in other settings. Ebenstein *et al.* (2016) find for Israeli school exit exams, that a comparable change in PM<sub>2.5</sub> air pollution decreases school exit exam scores by about 4.7% of a standard deviation.<sup>5</sup>

<sup>&</sup>lt;sup>4</sup>A possible concern could be that parents are keeping their children at home when pollution levels are high. However, we do not find any indication that pollution levels affect the probability of participating in the test.

<sup>&</sup>lt;sup>5</sup>The student fixed effects estimates suggest a -0.040 effect per AQI unit, which corresponds roughly to 2.8  $\mu g/m^3$  and the standard deviation of the test points is 23.7.

Carneiro *et al.* (2021) find that a  $10\mu g/m^3$  change in PM<sub>10</sub> air pollution reduces university entry exam scores in Brazil by 8% of a standard deviation. Remember that our results are taken in an environment with much lower air pollution levels (9.9  $\mu g/m^3$  PM<sub>2.5</sub> in Denmark vs. 21.1  $\mu g/m^3$ PM<sub>2.5</sub> in Israel (Ebenstein *et al.*, 2016) and 17.5  $\mu g/m^3$  PM<sub>10</sub> in Brazil (Carneiro *et al.*, 2021)) and with national tests at much younger ages in elementary and lower secondary school.

	(1)	(2)	(3)
Math domain:	Algebra	Geometry	Applied
$PM_{-}(2.5)*10$	$-0.015^{***}$ (0.002)	$-0.016^{***}$ (0.002)	$-0.016^{***}$ (0.002)
Observations R-squared	$846,\!342$ 0.56	$846,\!342 \\ 0.56$	$846,342 \\ 0.58$
Reading domain:	Language Comprehension	Decoding	Text Comprehension
$PM_{-}(2.5)*10$	-0.009*** (0.002)	$-0.008^{***}$ (0.002)	$-0.007^{***}$ (0.002)
Observations R-squared	$1,752,389 \\ 0.49$	$1,752,389 \\ 0.59$	$1,752,389 \\ 0.58$
Level of fixed effects	Family	Family	Family

Table 3: Impact of  $PM_{2.5}$  on test scores by domain

*Notes:* The dependent variables in all regressions are standardized test scores from the national compulsory tests split into three domains for both math and reading. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level in parentheses: \* \* \* p < 0.01, \* \* p < 0.05, \* p < 0.1.

While our main specification using family fixed effects illustrates a negative relationship between the level of air pollution and students' academic performance, there might well be date specific unobserved factors that affect pollution and test scores simultaneously. Besides endogeneity concerns, instrumenting air pollution can alleviate measurement error issues in the estimation. The instrument only exploits municipality variation in air pollution that is driven by wind direction and, thus, affects everyone in the municipality similarly.<sup>6</sup> The IV results are presented in columns 3 and 4 of Table 2. Our findings suggests that a 10-point increase in air pollution decreases test scores in math with 2.0% of a standard deviation, and decreases test scores in reading with 0.8% of a standard deviation. These estimates are very close to the main estimates, only much less precisely estimated. The effect on math test scores is still statistically significant at the 5% level. Overall, these results suggest that our FE estimates in columns 1 and 2 are not subject to endogeneity or measurement error concerns.

We round off this section by exploring the effects on the subdomains of the national tests. Table 3 shows that the average effects are very similar across all subdomains of both the math and the reading tests. It suggests to us that our conclusions are robust and not sensitive to the precise type of skill tested within broader categories: air pollution affects cognitive functioning in general, though it is most severe for the skills applied in math regardless of the domain.

	(1)	(2)	(3)
	EU: $20\mu g/m^3$	WHO: $15\mu g/m^3$	WHO: $5\mu g/m^3$
Panel A: Math			
Above limit	-0.020***	-0.023***	-0.017***
	(0.007)	(0.005)	(0.005)
Observations	846,342	846,342	846,342
R-squared	0.63	0.63	0.63
Panel B: Reading			
Above limit	-0.012***	-0.013***	-0.008***
	(0.004)	(0.003)	(0.003)
Observations	1,752,389	1,752,389	1,752,389
R-squared	0.63	0.63	0.63

#### 4.2 Non-linearities Across Pollution Levels

Table 4: Effects of  $PM_{2.5}$  limit violations

Notes: The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. For WHO the safe limit is 15 and the low limit is 5, for EU the safe limit is 20. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level in parentheses: \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1.

<sup>&</sup>lt;sup>6</sup>Local measurements of air pollution may, conversely, be poor approximations of individual exposure, particularly if the measurement locations are not representative of the location of the students.

Whether or not there is a safe level of pollution has long been debated. Air quality guidelines that posit limits to air pollution seemed to suggest that such safe levels of pollution exist. In its latest guidelines, however, the World Health Organization emphasizes that there is no evidence for a safe level of air pollution (WHO, 2021). To test whether exceeding low air pollution limits yields consistent effects on test scores, we construct dummies for a 20  $\mu g/m^3$  limit violation (yearly average limit from European Union (2008)), for a 15  $\mu g/m^3$  limit violation (daily value that should not be exceeded at the 99th percentile of days throughout a year by WHO (2021)), and a 5  $\mu g/m^3$ limit violation (strict yearly average from WHO (2021)). The results are shown in Table 4. We find similar effect sizes across limits for both math and reading, where the estimates are significant for the 15 and 5  $\mu g/m^3$  limits. Importantly, even exceeding the low 5  $\mu g/m^3$  limit shows consistent and statistically significant negative effects on both test scores. This result supports the notion by the WHO (2021) that pollution at any level matters.

As a next step, we exploit the full distribution of pollution levels to test for non-linearities in the effects on test scores. Figure 4 shows results with discretized levels of pollution in 5  $\mu g/m^3$ intervals with a reference category of below 5  $\mu g/m^3$ . The effects on math and reading test scores are increasing almost linearly from 5-10  $\mu g/m^3$ , 10-15  $\mu g/m^3$ , and 15-20  $\mu g/m^3$ . Above 20  $\mu g/m^3$ , the effect start to level off, but also become less precisely estimated. This pattern of treatment effect sizes with increasing levels of pollution suggests two important interpretations. First, there is a significant negative impact on test scores even at levels far below the ones suggested by current and past limit guidelines. Second, the impact of pollution seems largely to be linear in levels.

#### 4.3 Are the Top and Bottom Performers Equally Affected?

How exactly air pollution affects cognition and what pathophysiological processes are involved is still largely unknown (Aguilar-Gomez *et al.*, 2022). What we do know is that brain function can be hampered by both a direct entrance of pollutant particles through the olfactory bulb and indirectly through inflammatory processes in the body, as shown in experiments with mice (Underwood, 2017). Whether these impacts manifest themselves in lower cognitive function across the board or in specific brain tasks is a priori unclear. It is, however, important to understand whether tasks that require complex brain function are more or less affected by exposure to air pollution. We attempt to explore the mechanisms indirectly by asking where in the distribution of test scores air pollution has the largest impact. In our setting, the tests are adaptive, such that higher test scores imply harder problems to be solved by the students. Heterogeneous effects across the difficulty of



Figure 4: Effects across pollution distribution

Notes: The figure plots the effects of  $5 - \mu g/m^3$ -interval dummies of air pollution relative to less than  $5 \ \mu g/m^3$  on standardized (a) math and (b) reading test scores. The regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level, and confidence intervals at 95% level are depicted.

the questions can help us understand better what type of mental processes are most susceptible to air pollution.

To test for heterogeneity across the test score distribution, we estimate unconditional quantile treatment effects (Firpo, 2007; Firpo *et al.*, 2009), using estimators developed by Borgen *et al.*  $(2021)^7$  that allow for high-dimensional fixed effects and continuous treatments. Using this method, we obtain quantile treatment effects that control for confounding variation as in our main estimation and are interpretable as the impact of air pollution at the q-th quantile of the unconditional test score distribution.

Figure 5 shows quantile treatment effects of air pollution on test scores in math in panel (a) and reading in panel (b). At the lowest percentiles in the math test score distribution, air pollution has a very small impact, less than half of our baseline estimate. The treatment effects increase steadily with the quantiles of the distribution. At the top, treatment affects exceed 3% of a standard deviation for a 10  $\mu g/m^3$  increase in air pollution. Although there is indication of larger impacts on higher test scores in math, the profile for reading is relatively flat. While the difficult parts in math tests are arguably more heavily relying on fluid intelligence, much of the reading test relies on crystallized intelligence and accumulated knowledge that is less prone to being influenced by outside factors. This result is consistent with the effects on problem solving and arithmetic problems in

<sup>&</sup>lt;sup>7</sup>The residualized quantile regression is a two-step procedure that first purges confounding variation from the treatment variable and then regresses the raw outcome variable on the residualized treatment.

#### Figure 5: Quantile treatment effects



Notes: The graphs show unconditional quantile treatment effects (Borgen et al., 2021) for every fifth percentile of the outcome distribution. Panel (a) shows results for math, panel (b) results for reading. Solid lines are quantile treatment effects, grey areas are 95% confidence bands, the dark dashed line is the linear OLS estimate with 95% confidence bands as light dashed lines. Pollution is measured at the home address. The estimation controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level.

brain-training games (La Nauze and Severnini, 2021; Krebs and Luechinger, 2021). Results for an air pollution effect on a language test for university students in a laboratory environment that was designed to involve higher mental processes point in a similar direction (Bedi *et al.*, 2021).

To further investigate this result, we obtain quantile treatment effects for the six test domains separately in Figure 6. The test in numbers and algebra builds strongly on accumulated knowledge of math principles and calculation rules, and we see that the results in Panel (a) are the least steep. The differences across the distribution in the geometry test in Panel (b) are much stronger. This pattern of results is consistent with fluid intelligence playing a larger role for problems that involve visualization and spatial reasoning. The results for applications in math in Panel (c) that are derived from a mixture of problems across math domains are less distinct than for geometry. All three reading domains show no pattern across the test score distribution, which is in line with the quantile regressions results for the summary measure.

#### 4.4 Who is Most Susceptible to Air Pollution?

The existing literature has struggled with linking children's test performance and exposure to pollution to indicators of their prior health, yet precisely health capital has been proposed as a potential moderator. Table 5 explores the extent to which our findings vary with children's early respiratory health. For both math and reading tests, we do not find any evidence that having a



Figure 6: Quantile treatment effects by domain

*Notes*: The graphs show unconditional quantile treatment effects (Borgen *et al.*, 2021) for every fifth percentile of the outcome distribution. Panel (a)-(b) shows results for the three math domains, panel (d)-(f) results for the three reading domains. Solid lines are quantile treatment effects, grey areas are 95% confidence bands, the dark dashed line is the linear OLS estimate with 95% confidence bands as light dashed lines. Pollution is measured at the home address. The estimation controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level.

respiratory health condition affects how daily levels of air pollution affect test scores. This finding goes against the intuition that air pollution only operates through diminished health and that individuals with prior health conditions are more heavily affected (Beatty and Shimshack, 2014; Schlenker and Walker, 2016; Deryugina *et al.*, 2019; Achakulwisut *et al.*, 2019). In contrast, our findings suggest that air pollution affects cognition directly and without health deterioration as the mediator.

An alternative explanation could be that we can only identify asthma diagnoses related to inand outpatient hospitalization, and not from primary health care physicians. In Denmark, it is estimated that around 10% of children have asthma,<sup>8</sup> however, we only identify a rate of 5%.

It is important to note that our results do not provide information about how air pollution affects children's risk of developing a respiratory health conditions, or whether children with respiratory health conditions are generally more susceptible to air pollution impacts in other domains.

In Table 6 we explore the extent to which our findings vary with exposure to air pollution early

<sup>&</sup>lt;sup>8</sup>Asthma-Allergy Denmark

	(1)	(2)	(3)
	Asthma	Respiratory	Chronic respiratory
Math			
$PM_{2.5}$	-0.0018***	-0.0018***	-0.0018***
	(0.0005)	(0.0005)	(0.0005)
$PM_{2.5}$ * 1(Health)	0.0019	0.0007	0.0017
	(0.0013)	(0.0007)	(0.0013)
1(Health)	-0.0287*	-0.0249***	-0.0272*
	(0.0166)	(0.0092)	(0.0162)
Observations	$583,\!026$	$583,\!026$	583,026
Reading			
$PM_{2.5}$	-0.0008***	-0.0008***	-0.0008***
	(0.0002)	(0.0002)	(0.0002)
$PM_{2.5}$ * 1(Health)	0.0003	0.0002	0.0002
	(0.0006)	(0.0003)	(0.0006)
1(Health)	-0.0223**	-0.0224***	-0.0194**
	(0.0100)	(0.0054)	(0.0098)
Observations	$1,\!341,\!116$	$1,\!341,\!116$	$1,\!341,\!116$
Mean condition	0.05	0.18	0.05

Table 5: Heterogeneity by prior health conditions in first 5 life-years

Notes: The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level in parentheses: \* \* \*p < 0.01, \* \* p < 0.05, \*p < 0.1.

in life. Evidence provided by Bharadwaj *et al.* (2017); Sanders (2012); Isen *et al.* (2017) shows that in-utero and early-life exposure to air pollution affects children's test scores and later life outcomes negatively. Therefore, we may expect children who have been exposed the most early in life, to be more affected by higher levels of air pollution. We use the three different measures of lifetime air pollution from Table 1, and construct an indicator variable to capture whether the student belongs to the 25% of children with the highest average exposure in the specific period of life. For both math and reading tests, our main results do not vary across average exposure to air pollution from birth until the year before first test (column 1), during the first three years of life (column 2), or in-utero exposure to air pollution (column 3). This suggests that current levels of air pollution affect all children equally, independently of their previous exposure.

Again, it is important to note that these results do not provide information about how early in life or lifetime exposure affects children's performance, but on the interaction between previous average exposure and current levels of air pollution.

Additionally, we explore the extent to which our results vary with the socioeconomic status of the family (Table B2), measured as top and bottom income quartiles. We do not find any heterogeneity across this dimension. Lastly, we explore heterogeneity by sex and find that although both boys and girls are significantly affected by air pollution, girls appear more susceptible with point estimates twice as large as those for boys (Table B3).

#### 4.5 Alternative Outcome with High Stakes: *Gymnasium* Exit Exams

To test whether our results on test scores carry over to high-stake situations, we use exit exam grades from upper secondary education, the *Gymnasium*, as an alternative outcome. All students in the same cohort sit the same exam, which is graded by an external censor. Roughly half of a cohort completes this educational track, which gives access to higher education. Grades from the *Gymnasium* exit exams contribute to the GPA that students use to apply for university and post-secondary education.

We present results on the effect of  $PM_{2.5}$  on standardized *Gymnasium* exit exam grades in Table 7. The regressions include a full set of control variables and school fixed effects. We rely on school fixed effects for this analysis, as there are fewer families with two or more siblings attending *Gymnasium* exit exams, leading to a dramatic drop in sample size when using family fixed effects. <sup>9</sup> The result in column 1 indicates that a 10-point increase in air pollution decreases the exam

 $<sup>^{9}</sup>$ We note that our main results with the national tests in the previous section also hold with school fixed effects,

	(1)	(2)	(3)
	Before test	First three years	In-utero
Math			
$PM_{2.5}$	-0.0018***	-0.0018***	-0.0017***
	(0.0003)	(0.0003)	(0.0003)
$PM_{2.5}*1(Top exposure)$	-0.0014	-0.0003	-0.0007
· /	(0.0023)	(0.0012)	(0.0008)
1(Top exposure)	0.0474	-0.0692***	0.0202
	(0.0352)	(0.0226)	(0.0128)
Observations	846,342	846,342	846,342
Reading			
$PM_{2.5}$	-0.0009***	-0.0009***	-0.0009***
	(0.0002)	(0.0002)	(0.0002)
$PM_{2.5}*1(Top exposure)$	-0.0005	0.0012	0.0005
	(0.0010)	(0.0008)	(0.0005)
1(Top exposure)	0.0106	$-0.1567^{***}$	-0.0068
	(0.0198)	(0.0157)	(0.0080)
Observations	1,752,389	1,752,389	1,752,389

Table 6: Heterogeneity by life-time pollution exposure

Notes: The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level in parentheses: \* \* \* p < 0.01, \* \* p < 0.05, \* p < 0.1.

	(1)	(2)
	Math	Danish
$PM_{2,5} * 10$	$-0.087^{***}$ (0.023)	-0.046** (0.023)
Average level of pollution Observations Level of fixed effects	8.4 191,584 School	7.5 265,661 School

Table 7: Impact of  $PM_{2.5}$  on *Gymnasium* exit exams

Notes: The dependent variables in all regressions are standardized exam grades. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, day-of-the-week, year, and grade level. Robust standard errors clustered by school level in parentheses: \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1.

grades in math by 8.7% of a standard deviation. Column 2 shows the effect is smaller but also significant for grades in Danish with a 4.6 of a standard deviation drop. Overall, these results for exit exams are larger in magnitude compared to those using test scores from compulsory tests and are much more similar to the effect sizes found in Ebenstein *et al.* (2016) and Carneiro *et al.* (2021). Note that exit exams have a number of distinct features that make them different from those of our main analysis: A smaller fraction of students takes these exams, the outcomes are grades given by censors, they are high-stake tests, students only do them once, and the tests are not adaptive.

#### 4.6 Robustness Checks

For our main results, we use a simple specification with air pollution on the day of the test to construct our measure of exposure. It is conceivable however, that the pollution level of the previous day still has an impact on test performance one day later. Moreover, air pollution may be correlated across consecutive days. We test for time lags in the treatment mechanism using a more flexible lead and lag specification in Figure 7. The graphs show coefficients for air pollution three days before the test and three days after the test. For math in Panel (a), we see that the negative effect accrues mostly on the test day and to a lesser extent from one day before. The results for reading in Panel (b) show a shift of the impact to the day before the test. With the tests being held mostly in the morning, results for the test date and the day before are consistent with a transitory pathway to cognitive performance.

see Table B5



Figure 7: Impact of air pollution on test scores in the days before and after the test

Notes: The figure plots the coefficients from a regression of test scores on  $PM_{2.5}$  for the test date (day=0) as well as the three days prior and post the test, estimated in a single regression. The regression includes controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors are clustered at school level. The confidence interval is for the 95% level.

We also see results for the days after the test in Figure 7. For both math and reading, none of the coefficients are statistically significant thus supporting our main specification. To further assess the robustness of our main results, we continue to investigate how the inclusion of different covariates affects the estimates. These findings are presented in Appendix Table B4. For both math and reading, leaving out weather information, including family background information, and controlling for birth order, does not change any of the coefficients appreciatively.

Additionally, we investigate how the results vary with the inclusion of different levels of fixed effects. These findings are presented in Appendix Table B5. For both math and reading tests, the OLS estimation (Column 1) yields smaller magnitudes than our preferred estimation after controlling for family fixed effects (Column 3): -1.2% vs -1.8% in math and +0.3% vs -0.9% in reading, with the former result for reading being statistically insignificant. When we include school fixed effects instead (column 2), the estimates come closer to our preferred results at -1.4% for math and -0.2% for reading. As an additional specification check, we control for individual fixed effects (column 4). The estimates are essentially the same as in our preferred model, but somewhat less precisely estimated.

We also change the specification of the included weather variables: temperature, precipitation, relative humidity, and percentage cloud cover (Table B6). For both math and reading tests, the results do not change when including a linear and quadratic term of temperature, precipitation,

	(1)	(2)	(3)	(4)
Pollutant:	$\mathbf{PM}_{2.5}$	$\mathbf{NO}_2$	$\mathbf{SO}_2$	CO
Math				
Pollutant*10	-0.018***	-0.008***	-0.034***	-0.004***
	(0.002)	(0.002)	(0.005)	(0.000)
Average level of pollution	9.89	10.69	1.80	156.02
Observations	846,342	$846,\!342$	$846,\!342$	846,342
Reading				
Pollutant*10	-0.009***	-0.003***	-0.019***	-0.002***
	(0.001)	(0.001)	(0.003)	(0.000)
Average level of pollution	9.97	10.87	2.91	158.81
Observations	1,752,389	1,752,389	1,752,389	1,752,389
Level of fixed effects	Family	Family	Family	Family

Table 8: Relationship to other air pollutants

Notes: The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level in parentheses: \* \* \* p < 0.01, \* \* p < 0.05, \* p < 0.1.

relative humidity, and percentage cloud cover (column 2), or when including all two-way interactions between weather variables (e.g. precipitation by temperature) (column 3).

For our main result, we estimate the effect of  $PM_{2.5}$  on test scores, acknowledging that other pollutants are positively correlated (Table B1) and could explain parts of the impact. Therefore, we interpret the effect of  $PM_{2.5}$  as the impact of compounded air pollution. As a final robustness test, we estimate separate models with other criteria air pollutants, nitrogen dioxide (NO<sub>2</sub>), sulfur dioxide (SO<sub>2</sub>), and carbon monoxide (CO), as the treatment variables. The results in Table 8 indicate that air pollution has a statistically significant and negative impact regardless of which pollutant is used. The NO<sub>2</sub> and SO<sub>2</sub> estimates are smaller and the CO estimate is larger in magnitude relative to the treatment mean.

#### 5 Conclusion

We have shown that an increase in fine particulate matter on the day of the test decreases the performance of students in math and reading in elementary and lower secondary school. The effect is most pronounced at the higher end of the test score distribution in math, suggesting that more complex cognitive functions are particularly vulnerable to air pollution. We did not find significant heterogeneity by previous health conditions, socioeconomic status, or long-term exposure to pollution, even though we analyzed extensive register data.

Our findings also indicate that even low air pollution targets do not ensure the absence of adverse effects on cognitive performance. Even at low levels, the marginal effect of a unit of air pollution has a comparably negative impact. This result is relevant for many places in the world. Although air pollution levels in Denmark are generally low, they are not very different from large population centers. Air pollution levels in Denmark are virtually the same as in the United Kingdom; concentration levels in the biggest cities, Copenhagen and Aarhus, place them between Los Angeles, Miami, and Seattle in the United States.<sup>10</sup>

In general, the negative effect of short-term air pollution on children's cognition is worrisome. If air pollution universally diminishes performance, human capital accumulation and productivity are hampered by the repeated exposure to every-day air pollution anywhere and at any time. A more positive reading of the results implies that the transitory effect on performance is reversible and vanishes as soon as air pollution is reduced, which should encourage policy makers to consider

 $<sup>^{10}\</sup>mathrm{Data}$  based on 2023 averages for  $\mathrm{PM}_{2.5}$  from www.iqair.com.

these potential benefits even at low prevailing levels of air pollution.

#### References

- ACHAKULWISUT, P., BRAUER, M., HYSTAD, P. and ANENBERG, S. C. (2019). Global, national, and urban burdens of paediatric asthma incidence attributable to ambient no2 pollution: estimates from global datasets. *The Lancet Planetary Health*, **3** (4), e166–e178.
- AGUILAR-GOMEZ, S., DWYER, H., GRAFF ZIVIN, J. and NEIDELL, M. (2022). This is air: The "nonhealth" effects of air pollution. *Annual Review of Resource Economics*, **14**, 403–425.
- ANDERSON, M. L. (2020). As the wind blows: The effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association*, **18** (4), 1886–1927.
- ANTONSEN, S., MOK, P. L., WEBB, R. T., MORTENSEN, P. B., MCGRATH, J. J., AGERBO, E., BRANDT, J., GEELS, C., CHRISTENSEN, J. H. and PEDERSEN, C. B. (2020). Exposure to air pollution during childhood and risk of developing schizophrenia: a national cohort study. *The Lancet Planetary Health*, 4 (2), e64–e73.
- BEATTY, T. K. and SHIMSHACK, J. P. (2014). Air pollution and children's respiratory health: A cohort analysis. *Journal of Environmental Economics and Management*, **67** (1), 39–57.
- BEDI, A. S., NAKAGUMA, M. Y., RESTREPO, B. J. and RIEGER, M. (2021). Particle pollution and cognition: Evidence from sensitive cognitive tests in brazil. *Journal of the Association of Environmental and Resource Economists*, 8 (3), 443–474.
- BEUCHERT, L. V. and NANDRUP, A. B. (2018). The danish national tests at a glance. Nationaløkonomisk Tidsskrift, p. 2.
- BHARADWAJ, P., GIBSON, M., ZIVIN, J. G. and NEILSON, C. (2017). Gray matters: Fetal pollution exposure and human capital formation. *Journal of the Association of Environmental and Resource Economists*, 4 (2), 505–542.
- BORGEN, N. T., HAUPT, A. and WIBORG, Ø. N. (2021). A new framework for estimation of unconditional quantile treatment effects: The residualized quantile regression (rqr) model.
- BRANDT, J., CHRISTENSEN, J., FROHN, L. and BERKOWICZ, R. (2003). Air pollution forecasting from regional to urban street scale—implementation and validation for two cities in Denmark. *Physics and Chemistry of the Earth, Parts A/B/C*, 28 (8), 335–344.

- —, CHRISTENSEN, J. H., FROHN, L. M., PALMGREN, F., BERKOWICZ, R. and ZLATEV, Z. (2001). Operational air pollution forecasts from european to local scale. *Atmospheric Environment*, **35**, S91–S98.
- —, SILVER, J., FROHN, L., GEELS, C., GROSS, A., HANSEN, A., HANSEN, K., HEDEGAARD, G., SKJØTH, C., VILLADSEN, H. et al. (2012). An integrated model study for Europe and North America using the Danish Eulerian Hemispheric Model with focus on intercontinental transport of air pollution. Atmospheric Environment, 53, 156–176.
- BROOK, R. D., URCH, B., DVONCH, J. T., BARD, R. L., SPECK, M., KEELER, G., MORISHITA, M., MARSIK, F. J., KAMAL, A. S., KACIROTI, N. *et al.* (2009). Insights into the mechanisms and mediators of the effects of air pollution exposure on blood pressure and vascular function in healthy humans. *Hypertension*, **54** (3), 659–667.
- CARNEIRO, J., COLE, M. A. and STROBL, E. (2021). The effects of air pollution on students' cognitive performance: Evidence from brazilian university entrance tests. *Journal of the Association* of Environmental and Resource Economists, 8 (6), 1051–1077.
- CHANG, T., GRAFF ZIVIN, J., GROSS, T. and NEIDELL, M. (2016). Particulate pollution and the productivity of pear packers. *American Economic Journal: Economic Policy*, 8 (3), 141–69.
- CHANG, T. Y., GRAFF ZIVIN, J., GROSS, T. and NEIDELL, M. (2019). The effect of pollution on worker productivity: evidence from call center workers in china. *American Economic Journal: Applied Economics*, **11** (1), 151–172.
- CUNHA, F. and HECKMAN, J. (2007). The technology of skill formation. *American economic review*, **97** (2), 31–47.
- CURRIE, J. and NEIDELL, M. (2005). Air pollution and infant health: what can we learn from california's recent experience? *The Quarterly Journal of Economics*, **120** (3), 1003–1030.
- DERYUGINA, T., HEUTEL, G., MILLER, N. H., MOLITOR, D. and REIF, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, **109** (12), 4178–4219.
- DOCKERY, D. W. and POPE, C. A. (1994). Acute respiratory effects of particulate air pollution. Annual review of public health, 15 (1), 107–132.

- DOMINICI, F., PENG, R. D., BELL, M. L., PHAM, L., MCDERMOTT, A., ZEGER, S. L. and SAMET, J. M. (2006). Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *Jama*, **295** (10), 1127–1134.
- EBENSTEIN, A., LAVY, V. and ROTH, S. (2016). The long-run economic consequences of highstakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, 8 (4), 36–65.
- EUROPEAN UNION (2008). Directive 2008/50/ec of the european parliament and of the council of 21 may 2008 on ambient air quality and cleaner air for europe. Official Journal of the European Union.
- FIRPO, S. (2007). Efficient semiparametric estimation of quantile treatment effects. *Econometrica*, 75 (1), 259–276.
- -, FORTIN, N. M. and LEMIEUX, T. (2009). Unconditional quantile regressions. *Econometrica*, **77** (3), 953–973.
- GRAFF ZIVIN, J. and NEIDELL, M. (2012). The impact of pollution on worker productivity. American Economic Review, 102 (7), 3652–73.
- HALLIDAY, T. J., LUSHER, L., INAFUKU, R. and DE PAULA, A. (2022). Vog: Using volcanic eruptions to estimate the impact of air pollution on student learning outcomes. *IZA Discussion Paper Series*, **15696**.
- HEISSEL, J. A., PERSICO, C. and SIMON, D. (2022). Does pollution drive achievement? the effect of traffic pollution on academic performance. *Journal of Human Resources*, **57** (3), 747–776.
- ISEN, A., ROSSIN-SLATER, M. and WALKER, W. R. (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.
- KHAN, A., PLANA-RIPOLL, O., ANTONSEN, S., BRANDT, J., GEELS, C., LANDECKER, H., SULLI-VAN, P. F., PEDERSEN, C. B. and RZHETSKY, A. (2019). Environmental pollution is associated with increased risk of psychiatric disorders in the us and denmark. *PLoS biology*, **17** (8), e3000353.
- KREBS, B. and LUECHINGER, S. (2021). Air pollution, cognitive performance, and the role of task proficiency. *Cognitive Performance, and the Role of Task Proficiency (October 21, 2021)*.

- LA NAUZE, A. and SEVERNINI, E. R. (2021). Air pollution and adult cognition: evidence from brain training. Tech. rep., National Bureau of Economic Research.
- LANGRISH, J. P., UNOSSON, J., BOSSON, J., BARATH, S., MUALA, A., BLACKWELL, S., SÖDERBERG, S., POURAZAR, J., MEGSON, I. L., TREWEEKE, A. et al. (2013). Altered nitric oxide bioavailability contributes to diesel exhaust inhalation-induced cardiovascular dysfunction in man. Journal of the American Heart Association, 2 (1), e004309.
- ROTH, S. (2020). The effect of indoor air pollution on cognitive performance: Evidence from the uk. Manuscript, London School of Economics https://personal. lse. ac. uk/roths/JMP. pdf.
- 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.
- SCHLENKER, W. and WALKER, W. R. (2016). Airports, air pollution, and contemporaneous health. The Review of Economic Studies, 83 (2), 768–809.
- SICARD, P., AGATHOKLEOUS, E., ANENBERG, S. C., DE MARCO, A., PAOLETTI, E. and CA-LATAYUD, V. (2023). Trends in urban air pollution over the last two decades: A global perspective. Science of The Total Environment, 858, 160064.
- SPILLER, J., CLAYTON, S., CRAGG, L., JOHNSON, S., SIMMS, V. and GILMORE, C. (2023). Higher level domain specific skills in mathematics; the relationship between algebra, geometry, executive function skills and mathematics achievement. *Plos one*, **18** (11), e0291796.
- THYGESEN, M., HOLST, G. J., HANSEN, B., GEELS, C., KALKBRENNER, A., SCHENDEL, D., BRANDT, J., PEDERSEN, C. B. and DALSGAARD, S. (2020). Exposure to air pollution in early childhood and the association with attention-deficit hyperactivity disorder. *Environmental re*search, 183, 108930.
- UNDERWOOD, E. (2017). The polluted brain. Science, 355 (6323), 342–345.
- WARD, C. J. (2015). It's an ill wind: the effect of fine particulate air pollution on respiratory hospitalizations. *Canadian Journal of Economics/Revue canadienne d'économique*, 48 (5), 1694– 1732.
- WHO (2021). WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide. World Health Organization.

- ZHANG, X., CHEN, X. and ZHANG, X. (2018). The impact of exposure to air pollution on cognitive performance. *Proceedings of the National Academy of Sciences*, **115** (37), 9193–9197.
- ZIVIN, J. G., LIU, T., SONG, Y., TANG, Q. and ZHANG, P. (2020). The unintended impacts of agricultural fires: Human capital in china. *Journal of Development Economics*, **147**, 102560.

## **Online Appendix**

"Air Pollution and Cognition in Children: Evidence from National Tests in

Denmark"

## **Online Appendix A: Additional Figures**

Figure A1: Grouping of municipalities



*Notes*: Guldborgsund and Lolland are one group, Samsø is a part of east-mid-Jytland, Bornholm is grouped for itself and Læsø is a part of east-north Jutland.

# **Online Appendix B: Additional Tables**

	(1)	(2)	(3)	(4)
	$NO_2$	$SO_2$	CO	$PM_{2.5}$
(1) $NO_2$	1.000			
(2) $SO_2$	0.693	1.000		
(3) CO	0.753	0.678	1.000	
(4) $PM_{2.5}$	0.620	0.543	0.697	1.000

Table B1: Correlation between difference types of pollutants

	(1)	(2)
Math		
$PM_{2.5}$	-0.0017***	-0.0016***
	(0.0004)	(0.0004)
$PM_{2.5}^*$ low SES	-0.0001	-0.0000
	(0.0005)	(0.0005)
$PM_{2.5}$ * high SES	-0.0001	-0.0005
	(0.0005)	(0.0005)
Observations	$846,\!342$	846,342
Reading		
$PM_{2.5}$	-0.0007***	-0.0008***
	(0.0002)	(0.0002)
$PM_{2.5} * low SES$	-0.0003	0.0000
	(0.0003)	(0.0003)
$PM_{2.5}$ * high SES	-0.0005*	-0.0002
	(0.0003)	(0.0003)
Observations	1,752,389	1,752,389
Definition of low SES	25% lowest income	25% lowest income
Definition of high SES	25% highest income	25% highest income

Table B2: Heterogeneity by parental SES

Notes: The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level in parentheses: \* \* \* p < 0.01, \* \* p < 0.05, \* p < 0.1.

	(1)	(2)
	Math	Reading
$PM_{2.5}$	-0.0012***	-0.0006***
	(0.0004)	(0.0002)
$PM_{2.5}*Female$	-0.0011***	-0.0006***
	(0.0004)	(0.0002)
Female	-0.0086	$0.1924^{***}$
	(0.0056)	(0.0038)
Observations	846,342	1,752,389

Table B3: Heterogeneity by sex of student

Notes: The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level in parentheses: \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1.

	(1)	(2)	(3)	(4)	(5)
Math					
$PM_{2,5} * 10$	-0.018***	-0.018***	-0.016***	-0.016***	-0.017***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Average level of pollution	9.89	9.89	9.89	9.89	9.89
Observations	$846,\!342$	$846,\!342$	$809,\!431$	$809,\!431$	$809,\!431$
R-squared	0.63	0.63	0.62	0.63	0.63
Reading					
$PM_{2,5} * 10$	-0.012***	-0.009***	-0.010***	-0.008***	-0.009***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Average level of pollution	9.97	9.97	9.97	9.97	9.97
Observations	1,751,846	1,751,846	1,673,564	1,673,564	1,673,564
R-squared	0.64	0.64	0.64	0.64	0.64
Weather control	No	Yes	No	Yes	Yes
Background control	No	No	Yes	Yes	Yes
Birth order	No	No	No	No	Yes
Level of fixed effects	Family	Family	Family	Family	Family

Table B4: Robustness to including covariates

*Notes:* The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level in parentheses: \* \* \* p < 0.01, \* \* p < 0.05, \* p < 0.1.

	(1)	(2)	(3)	(4)
Math				
$PM_{2,5} * 10$	-0.012**	-0.014***	-0.018***	-0.016***
	(0.004)	(0.002)	(0.003)	(0.005)
Average level of pollution	9.89	9.89	9.89	9.89
Observations	$846,\!342$	$846,\!342$	$846,\!342$	$846,\!342$
R-squared	0.00	0.08	0.63	0.86
Reading				
$PM_{2,5} * 10$	0.003	-0.002*	-0.009***	-0.009***
	(0.003)	(0.001)	(0.002)	(0.002)
Average level of pollution	9.97	9.97	9.97	9.97
Observations	1,752,389	1,752,389	1,752,389	1,752,389
R-squared	0.01	0.09	0.63	0.83
Level of fixed effects	No	School	Family	Individual

Table B5: Robustness to levels of fixed effects

Notes: The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. Standard errors clustered at the school level in parentheses: \* \* \* p < 0.01, \* \* p < 0.05, \* p < 0.1.

	(1)	(2)	(3)
Math			
$PM_{2,5} * 10$	-0.018***	-0.013***	-0.016***
	(0.003)	(0.003)	(0.003)
Observations	$846,\!342$	$846,\!342$	$846,\!342$
Reading			
$PM_{2,5} * 10$	-0.009***	-0.006***	-0.008***
	(0.002)	(0.002)	(0.002)
Observations	1,752,389	$1,\!752,\!389$	1,752,389
Weather squared	No	Yes	No
Weather interaction	No	No	Yes

Table B6: Impact of  $PM_{2.5}$  on test scores – different weather specifications

Notes: The dependent variables in all regressions are standardized test scores from the national compulsory tests. Pollution is measured at the home address. All regressions include controls for temperature, precipitation, humidity, and cloud cover, as well as fixed effects for sex, family, time-of-date (early morning, morning, later), day-of-the-week, year, and grade level. In weather squared, we include a squared term of temperature, precipitation, relative humidity, and percentage cloud cover. In weather interaction, we include all two-way interactions between weather variables (e.g., temperature by precipitation). Standard errors clustered at the school level in parentheses: \*\*\*p < 0.01, \*\*p < 0.05, \*p < 0.1.