The Effect of Air Purifiers in Schools*

Jacopo Bonan[2,3,6,*], Francesco Granella[2,3], Stefania Renna[2,3,4], Luis Sarmiento[1,2,3,5], and Massimo Tavoni[2,3,4]

[1]Department of Management, Technology, and Economics, ETH Zurich, Zurich, Switzerland.
[2]RFF-CMCC European Institute on Economics and the Environment (EIEE), Milan, Italy.
[3]Centro Euro-Mediterraneo Sui Cambiamenti Climatici (CMCC), Lecce, Italy.
[4]Department of Management, Economics, and Industrial Engineering, Politecnico di Milano, Milan, Italy.
[5]Banco de México, Mérida, Mexico.
[6]University of Milan, Milan, Italy.
[*]Correspondence: `jacopo.bonan@unimi.it`

May 8, 2026

Abstract

We randomize the installation of air purifiers across primary school classrooms to reduce children's exposure to air pollution. The intervention reduces indoor fine particulate matter ($PM_{2.5}$) daily concentrations by 33% (18% during school hours) and decreases student absenteeism by 17%. We find larger effects among students with higher pre-treatment absenteeism, and when ambient pollution levels are low enough for purifiers to bring indoor concentrations below health-relevant thresholds. Treated students report fewer respiratory symptoms and exhibit greater awareness of air quality. The intervention has no effect on cognitive skills, mood, and aggressive behavior. Each avoided absence day costs approximately €6.9, yielding a benefit-cost ratio of about 9.5:1, even ignoring potential benefits to academic performance.

Keywords: Indoor air quality, air purifiers, school absences, randomized controlled trial

JEL Classification Numbers: C93, I21, Q53, Q51

---

[*]This study was funded by the European Union - NextGenerationEU, Mission 4, Component 2, in the framework of the GRINS - Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 – CUP C93C22005270001) and from PRIN 2020HKPNPL "DIGITA". The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them. We are thankful to Luca Valentini and Giulio Spaggiari from NETCO for the fruitful collaboration. We thank school directors and the team of teachers who collaborated with us in the project implementation. This project benefited from feedback from Michela Boldrini, Giovanna d'Adda, John Gomez-Mahecha, Wei Guo, Jacopo Lunghi, Beia Spiller, and participants at GLO-SITES 2024, Symposium on Advancing Sustainability Research with Experimental Methods, University of Padua, IAERE 2025, and EIEE Brown Bag seminar. The project was pre-registered on the AEA RCT Registry (AEARCTR-0011960) and received ethical approval from Politecnico di Milano Institutional Review Board (code 18/2023).

1 Introduction

Air pollution is a global health issue contributing to child mortality and morbidity worldwide (Institute for Health Metrics and Evaluation, 2019; Annesi-Maesano et al., 2021). High pollution-related morbidity increases school absences (Currie et al., 2009), negatively impacting learning outcomes and educational activities (Gershenson et al., 2017; Aguilar-Gomez et al., 2022). Although average urban exposure has decreased in recent decades due to different pollution-control measures, levels often exceed WHO guidelines (World Health Organization, 2021), leading to significant health, economic, and welfare losses (Shaddick et al., 2020). Since traditional air pollution control mechanisms, such as low-emission zones or industrial policies, can be costly and complex to implement, adaptive measures are necessary to reduce exposure.

This study evaluates the efficacy and cost-effectiveness of installing portable air purifiers in school classrooms. Our main hypothesis is that air purifiers improve indoor air quality and reduce absenteeism by promoting children's health. Using a *cluster randomized controlled trial* (RCT) across five primary schools in Milan, Italy — an area characterized by poor air quality (EEA, 2022) — we randomly assigned 95 classes to either receive or not receive air purifiers. Additionally, we installed indoor air quality sensors in a subsample of classrooms to collect detailed data on air pollution and environmental conditions.

Air purifiers reduce indoor air pollution on average by 33% and by 18% during school-time. The relative efficacy of the purifiers does not appear to be related to outdoor air pollution levels and remains relatively stable throughout the study period. The treatment decreases school absences by approximately 17%, equivalent to about 2.1 fewer missed days per year. The effect is more pronounced for students with higher baseline absenteeism. Dynamic treatment effects indicate that the reduction in absences primarily occurs in fall and spring, rather than in winter when average pollution levels are significantly higher. Supporting this observation, the treatment effect on absences is approximately halved on days with above-median outdoor $PM_{2.5}$ and becomes statistically indistinguishable from zero in the highest pollution quartile - e.g., when students remain exposed to health-relevant thresholds even in the presence of air purifiers. These insights align with previous findings indicating that the marginal effect of air pollution on health is concave, i.e., improving air quality when pollution is relatively low has larger health effects than when pollution is high (Berkouwer and Dean, 2026; Miller et al., 2024; Weichenthal et al., 2022; Corrigan et al., 2018; Pope III et al., 2015; Aragón et al., 2017). However, our study is not designed to cleanly isolate this mechanism from potential confounders, such as the higher incidence of influenza and other illnesses in winter.

We do not find significant effects on cognitive skills (fluid intelligence), mood, and aggressive behavior; we can rule out effects as big as 0.16-0.2 SD. Our study is underpowered to detect impacts on academic performance. The results on indoor air pollution, absences, and the heterogeneous effects by pre-specified dimensions are robust to multiple hypothesis correction ($q < 0.1$).[1]

Using survey data, we find suggestive evidence that treated students are less likely to report respiratory symptoms over the past week than control students. This result suggests that reduced absences likely arise from improved health. We also observe significant differences in students' perceptions of classroom air quality and their preferences for urban policies related to air quality. However, we cannot rule out experimenter demand and priming effects. To examine the potential confounding role of behavioral changes in our results, we estimate the treatment effect on proxies for opening and closing classroom doors and windows (e.g., sudden shifts in temperature or carbon dioxide, CO_2). We find no evidence that purifiers significantly alter ventilation behavior or classroom occupancy in response to the treatment, as indicated by the lack of differences in classroom CO_2 levels, temperature, and estimated ventilation episodes.

Cost-effectiveness calculations indicate that installing air purifiers leads to a cost per avoided absence of approximately €6.9 and a benefit-cost ratio between 9.5 and 19. This estimate excludes any potential direct or indirect effect of the intervention on human capital accumulation, and is solely based on health and childcare costs.

The analysis closely follows the pre-analysis plan (PAP). Any deviations are explicitly documented and justified in the text, along with additional robustness checks.

Related literature Air pollution is increasingly recognized as a health hazard and a significant barrier to educational success. Even low levels of ambient pollution negatively affect school participation and learning across various contexts (Ebenstein et al., 2016; Carneiro et al., 2021; Sunyer et al., 2017; Roth, 2021; Gilraine and Zheng, 2024; Chiu et al., 2013; Rahai and Evans, 2023; Heyes et al., 2023; Lai et al., 2021; Yao et al., 2023; Palacios et al., 2022). Polluted air harms children's health and cognitive function (La Nauze and Severnini,

---

[1] The FDR correction for the primary outcomes does not include the two test score outcomes, for which the study is underpowered. This is a deviation from the PAP. When including them, the q-value for the main effect on absences is 0.13.

2025; Künn et al., 2019), leading to lower attendance and academic performance (Chen et al., 2018; Currie et al., 2009; Komisarow and Pakhtigian, 2022; Ransom and Pope, 1992; Persico and Venator, 2019; Heissel et al., 2022). This issue is critical, as educational outcomes have long-term implications for human capital formation, productivity, and lifetime earnings (Graff Zivin and Neidell, 2013). The adverse effects are especially pronounced among vulnerable children with pre-existing health issues or higher baseline absenteeism (Liu and Salvo, 2018). Our paper builds on this literature by providing the first experimental evidence of the causal impact of indoor air purifiers on student absenteeism. This design addresses many common confounders in observational studies and expands the literature into a developed country setting with moderate to high ambient pollution levels. In doing so, our work corroborates the negative educational effects of air pollution.

Within education policy, cost-effectiveness is crucial for resource allocation. Policymakers must choose from various interventions to maximize educational gains. Many traditional interventions to improve educational outcomes require substantial investments, yet their cost-effectiveness varies widely (Angrist et al., 2020). Among programs to reduce absenteeism, behavioral interventions informing parents about their child's attendance proved highly cost-effective (Rogers and Feller, 2018; Robinson et al., 2018). However, few studies have assessed the cost-effectiveness of specific interventions to improve the physical learning environment and the quality of indoor air.[2] Some epidemiological studies highlight the benefits of air purifiers on reducing indoor pollution in schools (Carmona et al., 2022; Tong et al., 2020) and associated significant health benefits (Chen et al., 2015; Yang et al., 2021).[3] Closer to our study, Kremer et al. (2025) randomly assign air filters in public high schools in Bogota. They find small increases (+ 3% of a SD) of high-stakes test scores within 4 months of installation, and no effect in the following year. However, the improvements in air pollution induced by the purifiers are 0.47-1.2 μg/m^3 (4-8% of the mean), about a fourth of our experimentally-induced variation. In a similar design, Bharti et al. (2025) find an increase of 0.15 SD in test scores in grade-2 schools in Lahore, Pakistan, shortly after the installation, but no improvements after a prolonged holiday period. In their study, air purifiers reduce indoor $PM_{2.5}$ by 48–76 μg/m^3, or 20%. In a related study, Gilraine (2023) finds that installing air filters in schools leads to a 0.1 to 0.2 standard deviation increase in test scores, utilizing a regression discontinuity design.[4] Several concurrent RCTs examine air purifiers in schools across developing country settings with substantially higher pollution levels: Liu

---

[2] Impact assessments of general school infrastructural investments are provided, for instance, in Cellini et al. (2010). The benefits of air conditioning for learning are studied in Park et al. (2020).

[3] See Cheek et al. (2021) and Xia et al. (2021) for a review.

[4] In an RCT, Gignac et al. (2021) find no short-term effect of purifiers on adolescent attention.

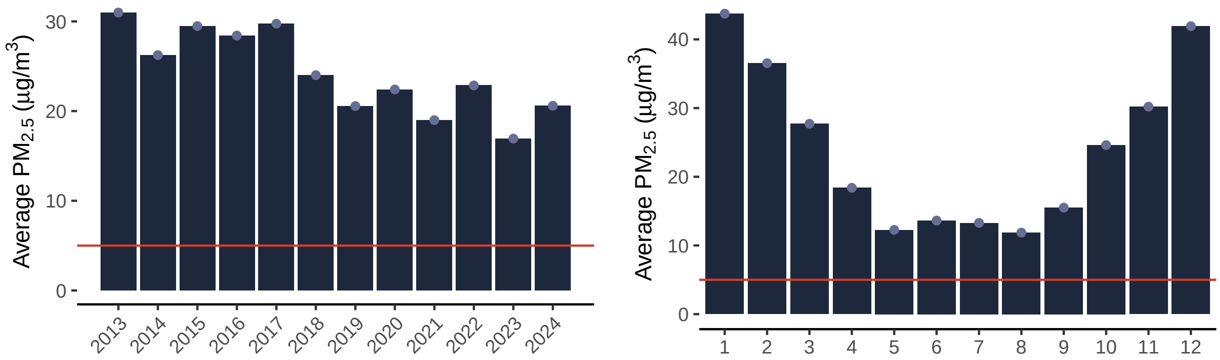
et al. (2024) in Chinese middle schools and Ruiz-Tagle et al. (2024) in middle schools in Delhi. Our study is the first school-based purifier RCT in a developed country with moderate pollution, and the first to focus on primary school absenteeism as the main outcome with direct measurement of indoor air quality. More broadly, our study complements the existing evidence along several dimensions. First, the RCT design eliminates selection concerns that arise in quasi-experimental approaches. Second, we directly measure the first stage — indoor $PM_{2.5}$ reductions — using classroom-level sensors, establishing the causal chain from purifiers to air quality to health proxies. Third, we focus on absenteeism, a distinct margin from test scores with direct implications for household costs and school operations, and provide the first cost-benefit analysis of air purifiers in schools. Fourth, survey data on health symptoms, seasonal heterogeneity in treatment effects, and behavioral adaptation tests provide evidence on mechanisms largely absent from prior work.

Finally, we contribute to the literature on adaptation to environmental stressors and the role of exposure to indoor air pollution (Graff Zivin and Neidell, 2013; Deschênes et al., 2017; Park et al., 2020; Burke et al., 2022; Coury et al., 2024; Barwick et al., 2024). As environmental stressors increasingly challenge public health and productivity, adaptive responses are essential to mitigate the negative effects of environmental hazards and minimize adverse outcomes. Studies indicate that households and firms invest in protective technologies to reduce personal exposure (Deschênes et al., 2017; Ito and Zhang, 2019; Greenstone et al., 2021; Zhang and Mu, 2018; Baylis et al., 2024; Metcalfe and Roth, 2025). While adaptation does not replace pollution control, it serves as a crucial secondary defense mechanism, especially when complete hazard elimination is unfeasible. Nonetheless, research indicates that adaptive measures have limitations; for instance, when outdoor pollution reaches extreme levels, the protective behaviors may not be sufficient to mitigate the health risks (Burke et al., 2022; Barwick et al., 2024). Our study contributes to the adaptation literature by empirically testing an indoor air quality improvement technology as an adaptive strategy in schools. The results show that while the intervention reduces student absences during moderate pollution periods, its effectiveness diminishes under severe outdoor pollution and during periods with overlapping health risks (e.g., respiratory infections, flu, etc.). This finding highlights an important boundary condition for adaptive interventions and suggests that such measures should complement broader pollution abatement policies.

2 Experimental design

Context Northern Italy features high population density and economic activity, along with poor orographic conditions that impede air circulation. This combination makes it one of the most polluted regions in Europe (European Environmental Agency, 2024). Air pollution levels in Northern Italy consistently exceed WHO guidelines of 15 and 5 $\mu g/m^3$ for the daily and annual concentration of $PM_{2.5}$, respectively. Panel (a) of Figure 2 displays average annual $PM_{2.5}$ concentrations in Milan—the region's largest city—from 2013 to 2024. Although air pollution has declined over this period, levels remain well above WHO thresholds for good air quality. During our study period (the 2023–24 school year), the average annual concentration of $PM_{2.5}$ was 18.5 $\mu g/m^3$.

Annual averages, however, mask substantial seasonal variation, as illustrated in Panel (b). Pollution levels are markedly higher in winter than in summer, driven by factors such as thermal inversions, increased residential heating, and the reduced efficiency of internal combustion engines at lower temperatures. In our study period, average concentrations reached 32.3 $\mu g/m^3$ in winter (December–February) compared to 9.1 $\mu g/m^3$ in spring (April–June), and the WHO daily threshold of 15 $\mu g/m^3$ was exceeded on 43% of days.



(a) Yearly

(b) Monthly

Figure 1: Yearly and monthly time series of average $PM_{2.5}$ in Milan (2013-2024)

Notes: The values come from yearly and monthly averages of daily air pollution measurements across the 2 stations in the metropolitan city of Milan. The red line marks the yearly guideline value set by the WHO (5 $\mu g/m^3$).

To reduce exposure to air pollution, regional and local governments have implemented various measures, including investments in public transportation, upgrades to power plants, promotion of cleaner fuels, improvements in energy efficiency, and public awareness campaigns on air quality (Italian Republic, 2010; Lombardy Region, 2006, 2013). In Milan, vehicle traffic is regulated through a congestion charge introduced in 2012 and a Low Emission Zone (LEZ) established in 2022 (Municipality of Milan, 2022a, 2023).[5] The city also continues to invest in public infrastructure to improve the public transport network and promote cycling as part of its broader Air-Climate Plan (Municipality of Milan, 2022b). However, these efforts have produced only marginal improvements, typically appearing over the medium to long term.

State-owned schools dominate Italy's educational landscape, with approximately 94% of children enrolled in public institutions (Ministero Italiano dell'Istruzione e del Merito, 2023). The average Italian school building is over 50 years old and often does not meet current sustainability standards. Some classrooms exhibit inadequate maintenance, and environmental improvements occur infrequently (Ruggieri et al., 2019). Most schools were constructed before urban development, resulting in their proximity to high-traffic roads, which significantly increases exposure to air pollution.

In the Italian school system, at the beginning of the year, students are assigned to a single classroom where they spend most of their day. They leave daily for lunch at the school canteen and can visit the garden or courtyard during breaks.[6] The choice between spending breaks indoors vs. outdoors depends on the weather and teachers' preferences. Teachers may teach multiple classes and move between different physical classrooms throughout the day.

Intervention and randomization Our intervention installed 42 consumer-grade portable air purifiers in randomly selected classrooms across five schools.[7] We assigned classrooms to treatment and control groups, stratifying by school and grade. All purifiers were installed outside school hours between November 8 and 16, 2023. The intervention did not include targeted information campaigns or communications to teachers or parents. All purifiers operated continuously from November 2023 to June 2024. Because randomization occurs within schools, treated and control students share common spaces such as hallways, gyms, and dining halls. However, purifiers only operate within the assigned classroom, and students spend the large majority of instructional time in their own classroom, limiting the scope for spillover effects. The research team monitored purifier functionality through monthly

---

[5] While many studies find that LEZs significantly improve environmental outcomes (Klauber et al., 2024; Pestel and Wozny, 2021; Gehrsitz, 2017), the local environmental impact of Milan's congestion charge appears to be limited (Percoco, 2013, 2014).

[6] Students spend a few hours per week attending lab sessions in specialized classrooms and exercising in the school gym.

[7] All installations were carried out by professionals from the purifier manufacturer. They ensured correct and safe deployment and also addressed minor electrical issues on-site. This proved to be a crucial factor, as Italian public schools often have outdated electrical infrastructure.

statistical analyses of indoor air pollution data in a subsample of classrooms and bi-monthly on-site visits.

We installed NETCO NIVEUS NV100 air purifiers equipped with U15 Ultra Low Particulate Air (ULPA) filters, capable of capturing up to 99.99% of particles larger than 0.026 microns—the highest efficiency in mechanical filtration technology. These devices are energy-efficient, with power consumption comparable to a 60-watt incandescent bulb, and operate quietly, producing sound levels between 29 and 45 dB(A). Following the manufacturer's guidelines, we selected the model based on the average classroom volume. Purifiers operated at 60% capacity to ensure effective air purification with minimal noise, achieving an average Air Exchange Rate of 1.04. In addition to the purifiers, we randomly installed 31 indoor air quality sensors in a subsample of classrooms. Appendix B provides technical details on the purifiers and monitors. Sensors' installations were stratified by school, treatment status, and grade to ensure balanced representation across treatment and control groups. The sensors measure concentrations of $PM_{2.5}$, PM_{10}, CO_2, carbon monoxide (CO), as well as temperature, humidity, and atmospheric pressure. Once powered and connected to the internet, the sensors transmit data every 30 seconds to an online data platform.

Teachers administered paper-and-pencil surveys to students at two time points: before the intervention in October 2023 (baseline) and during the intervention in April 2024 (endline). Teachers selected survey administration dates within a two-week window based on availability. To improve comprehension, we used capital letters and visual Likert scales with emoticons. First-grade teachers adhered to a dedicated protocol, projecting and reading each question aloud to support student understanding. Participation was optional for first-grade students. The survey took approximately 15 minutes to complete.[8]

2.1 Sample, Data and Outcome Variables

Sample The study sample includes students from the 2023–24 school year, with absence and demographic data obtained from official school ledgers and registries. It comprises 95 classes and 2,050 students across five grades. For a subsample of them (N=1,609), we have absence data for the school-year 2022-23. Data on indoor environmental conditions were collected using 31 air quality sensors; one sensor was excluded from the analysis after data quality checks (see Appendix C for details).

The survey completion rate is approximately 88%, with 1,822 responses in the first wave, 1,815 in the second, and 1,662 in both. Two factors primarily explain this rate. (1) Many

---

[8] The English translation of the survey is available here.

schools chose not to administer the survey to first-grade students in either wave or missed an administration in some classes, and (2) student absences on the survey day. Appendix Tables A.1, A.2, and A.3 show that survey participation rates do not differ significantly between treatment and control groups.

Indoor air quality Indoor air pollution data are available from November 2023 to June 2024. We aggregate monitor readings into daily average concentrations (24-hour averages), as reported in the PAP, and school-time averages, computed using measurements recorded between 8:00 AM and 5:00 PM.

Low-cost air quality sensors can exhibit unit-to-unit variation in readings. If sensors assigned to treated classrooms systematically differed from those in control classrooms, this could bias the estimated treatment effect on indoor pollution. To test for this, we co-located all sensors in a single room for four consecutive days, so that any differences in readings reflect sensor error rather than true air quality differences. We find no evidence that measurement imprecision differs systematically between sensors assigned to treatment and control classrooms (Appendix C).

Absenteeism Schools collected daily absence data digitally and shared it with researchers in anonymized form at the end of the 2023–24 school year. We access data from the school years 2022-23 and 2023-24. The reasons for student absences are neither systematically recorded nor digitized, preventing us from identifying whether health-related issues are the primary cause. Our outcome variable for absenteeism is a binary indicator equal to one if a student is absent on a given day, and zero otherwise.

Academic and cognitive performance Academic performance is measured using scores from the Italian national standardized assessments (INVALSI). These tests are mandatory, anonymous, externally evaluated, and conducted annually. All students in the second and fifth grades take the assessments each spring on a common testing date. The tests evaluate proficiency in Mathematics and Italian. Test scores are constructed using a Rasch Item Response Theory model, consistent with the approach adopted in international assessments such as the OECD Programme for International Student Assessment (PISA). The main outcomes of our analysis are students' test scores in Mathematics and Italian.

The cognitive skill assessment is realized through a Raven test included in the student survey. It consists of a series of visual patterns with a missing piece, where test-takers must choose the correct piece from multiple options. We selected the suitable version for children aged 5 to 12 years from the first wave of the Mexican Family Life Survey (Rubalcava and Teruel, 2006). We summed all correct answers to create a score ranging from 0 to 18 and standardized this score to have a mean of zero and a variance of one for each grade, using the mean and standard deviation from the control group.[9] The Raven test is widely used as a measure of fluid intelligence, capturing individuals' ability to reason abstractly, identify patterns, and solve novel problems independently of acquired knowledge or language skills.

Mood and aggressive behavior We assess students' mood over the previous week using a survey question based on a Likert scale (very positive, positive, negative, very negative), which we convert into a 1-4 index.[10] To proxy aggressive episodes, we create a dummy variable that equals one if students report arguing or quarreling with any classmate during the past week (sometimes, often, very often) and zero otherwise (never).

Subjective health symptoms We use students' self-reported symptoms as a proxy for health conditions. Children reported the frequency of various respiratory and non-respiratory symptoms experienced over the past week using a four-point Likert scale: never, sometimes, often, and every day.[11] The symptoms include: runny nose, blocked nose, sneezing, cough, shortness of breath, tiredness, headache, and stomachache. We classify the first five symptoms as respiratory, the last two as non-respiratory, and the remaining as general. For each symptom, we create a binary indicator equal to one if the student reported experiencing it at least "sometimes," and zero if they selected "never."

Perceptions, beliefs, and behavioral responses The intervention did not include explicit communication or awareness campaigns about indoor air quality or environmental issues; however, it may have implicitly raised environmental awareness. We assess perceptions of air quality across different settings (overall, city, classroom, and courtyard) using a four-point Likert scale: very bad, bad, good, and very good. Responses received a score from 1 to 4. We also evaluated children's views on the importance of addressing urban challenges such as street garbage, lack of green areas or playgrounds, insufficient sports facilities, air pollution, and road traffic. These were rated on a four-point Likert scale: to a great extent,

---

[9] We attribute a missing value in the absence of all responses to the 18 tests, while missing on a subset of answers counts as zero in the final score.

[10] This outcome operationalization departs from what we promised in the PAP in that it does not use the question asking how students felt over the last week (very well, generally well, sometimes well, not very well), as the question was discarded in the second wave after we received negative feedback on its formulation from some schools.

[11] These questions are adapted from a validated survey on acute respiratory illnesses for children aged 4 to 10, developed by Schmit et al. (2021).

to some extent, to a limited extent, and not at all, and similarly coded as a score from 1 to 4.[12]

To measure teachers' behavioral responses to the purifiers, we monitored indoor environmental conditions—specifically CO_2 levels, temperature, and the frequency of window openings. These parameters are influenced by classroom occupancy and ventilation, but are not directly affected by the purifiers. We calculated minute-level averages of CO_2 and temperature using data from the air quality sensors. We identified ventilation episodes with sharp drops in CO_2 levels alongside changes in indoor temperature. Further details are provided in Appendix D.

Controls and dimensions of heterogeneity We obtained students' socio-demographic information from school administrative records, focusing specifically on gender and nationality. We created a binary variable, assigning a value of one for female students and zero for male students, along with a binary indicator for non-Italian citizenship.[13]

Outdoor air pollution We used outdoor pollution data from the European Environmental Agency database (EEA, 2024). We calculated daily average $PM_{2.5}$ levels for each school by applying inverse distance weighting from the two nearest background air quality monitoring stations.[14]

3 Descriptive Statistics

3.1 Sample characteristics and balance

Table 1 presents descriptive statistics for the student sample. Overall, 47% of students are female, and 37% have foreign citizenship. The distribution across the five grades is relatively balanced. In the pre-treatment period, i.e., over school-year 2022-23 and the period preceding treatment deployment in the school-year 2023-24 (September and October 2023), students

---

[12] This block of questions was inspired by the scales used in Cori et al. (2020).

[13] Under Italian law, children born in Italy to non-Italian parents acquire Italian citizenship at the age of 18.

[14] The two nearest stations are both located inside the city of Milan. The third closest monitoring station is far enough to have a negligible contribution to the monitored average using inverse-distance weighting, and is therefore excluded. A measurement error on the outdoor pollution assigned to schools may occur if there are major sources of pollution between the school and the monitoring stations, such as industrial plants or major road arteries, and the wind blows such that the school is up(down)wind and the monitor is down(up)wind. Any such measurement error would have no effect on the estimates of the effects of air purifiers on absences nor on indoor pollution. However, it might bias the ratio of indoor-to-outdoor concentrations and the heterogeneity of effects by outdoor air pollution.

were absent for approximately 5.7% of school days, averaging about 12 missed days in a standard 200-day school year.

Table 1: Descriptive statistics and balance

	(1)	(2)	(3)	(4)	(5)
	Obs	Control		Difference	
Variable		Mean	(SD)	Diff.	(SE)
Demographics and administrative data					
Female	2051	0.469	0.499	0.002	0.016
Grade 1	2051	0.184	0.388	0.016	0.081
Grade 2	2051	0.215	0.411	-0.020	0.086
Grade 3	2051	0.209	0.407	-0.017	0.085
Grade 4	2051	0.195	0.396	-0.012	0.081
Grade 5	2051	0.197	0.398	0.033	0.087
Non-Italian	2051	0.369	0.483	0.025	0.020
Class size	95	21.585	2.575	0.197	0.414
Pre-treatment absences	2051	0.057	0.054	0.001	0.003
Responded to baseline survey	2051	0.888	0.315	0.001	0.014
Responded to endline survey	2051	0.876	0.330	0.020	0.014
Cognitive and behavioral (baseline survey)					
Raven score (standardized)	1693	-0.000	0.998	-0.150*	0.080
Mood scale	1502	3.432	0.743	0.011	0.057
Aggressive episodes	1630	0.449	0.498	-0.033	0.037
Health symptoms (baseline survey)					
Runny nose	1451	0.456	0.498	-0.008	0.040
Blocked nose	1437	0.541	0.499	-0.001	0.039
Sneezing	1430	0.620	0.486	0.007	0.038
Cough	1433	0.556	0.497	0.030	0.038
Short of breath	1375	0.249	0.433	-0.004	0.033
Tiredness	1397	0.615	0.487	-0.002	0.039
Headache	1405	0.482	0.500	-0.039	0.035
Stomach ache	1389	0.424	0.494	-0.057	0.037
Air quality perceptions (baseline survey)					
Overall	1557	3.091	0.850	-0.036	0.092
City	1554	2.716	0.954	-0.142	0.089
Classroom	1553	3.230	0.747	-0.036	0.059
Schoolyard	1551	3.513	0.680	-0.112*	0.059
Urban policy priorities (baseline survey)					
City cleaning	1559	3.467	0.900	-0.014	0.067
Green areas and playgrounds	1540	3.278	0.982	-0.004	0.083
Sport infrastructure	1525	3.098	1.087	-0.105	0.099
Air quality	1536	3.486	0.955	-0.061	0.068
Less traffic	1513	3.169	0.997	-0.061	0.081

Notes: The table presents the mean and standard deviation in the control group (columns 2-3) for student socio-demographic characteristics and survey variables at the baseline, and the difference in means between treatment and control group, with its standard error (columns 4-5). Pre-treatment absences are calculated over the school year 2022-23 and the period preceding treatment deployment in the school-year 2023-24 (September and October 2023). Mood over the previous week is assessed on a scale from 1 (very negative) to 4 (very positive). Aggressive behavior is a dummy equal to one if students report arguing or quarreling with any classmate during the past week, zero if this never happened. Air Quality (AQ) perception indexes are expressed on a scale from 1 (very bad) to 4 (very good). Priority policy scores are expressed on a scale from 1 (not important at all) to 4 (very important). Significance levels: *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

As for the survey outcomes, the average Raven score across all students is 13.45 out of 18 (SD = 2.96). The score is subsequently normalized by grade. On average, students report a mood level of 3.4 out of 4, and 45% report having had an argument or quarrel with peers in the week preceding the survey. At baseline, between 46% and 62% of students reported experiencing a runny nose, blocked nose, sneezing, cough, or shortness of breath in the previous week, 61% reported tiredness, 48% headaches, and 42% stomachaches. Perceptions of air quality averaged 3.1 on a 1-to-4 scale, with lower ratings for city air quality (2.7) and higher ratings for classrooms (3.2). Students regarded outdoor school spaces, such as courtyards, as safer in terms of air quality (3.5) compared to the broader urban environment (2.7). Regarding urban priorities, students emphasized city cleanliness and air quality (both around 3.5), followed by green areas, playgrounds, traffic (3.2), and sports infrastructure (3.1). Administrative and survey-based measures show no significant differences between treatment and control groups, except the Raven score (significant at 10%) (see Columns 4–5).

Survey measures are affected by missing data at both baseline and endline. Appendix Tables A.1, A.2 and A.3 present the extent of missingness and test for correlations with treatment status in the first, second, and across both waves, respectively. Approximately one-quarter of students—and 13–15% at endline—did not respond to health-related questions. Non-response rates for items on air quality perceptions and policy preferences ranged from 15% to 18% at baseline and from 4% to 6% at endline. When considering both waves jointly, missing data reach up to 30–35% for reported symptoms, reducing the sample to approximately 1,100-1,200 observations. We find no evidence of systematic differential missingness by treatment status at either time point.

Appendix Table A.4 compares the demographic composition and academic performance of schools in the study sample with corresponding averages at the provincial (Milan), regional (Lombardy), and national levels, providing an assessment of the sample's external validity. Overall, the sample appears broadly comparable in terms of gender composition, with the share of female students closely aligned with all reference levels. However, notable differences emerge along other dimensions. First, students in the sample perform systematically worse in standardized INVALSI tests in both Italian and mathematics, across grades 2 and 5. These lower average scores are accompanied by substantially higher dispersion, particularly in Italian test results at grade 2, indicating greater heterogeneity in student performance within the sample. Second, the sample is characterized by larger class sizes, with an average of 21.6 students per class compared to 17 at the national level. Third, and most importantly, the share of non-Italian students is markedly higher in the sample (38%) than in the province (26%), region (21%), and country (14%). Taken together, these patterns suggest that the sample over-represents more diverse and potentially more challenging school environments, which likely contributes to both the lower average performance and the higher variability in test scores. In this sense, while the sample remains broadly comparable along some dimensions, it is skewed toward schools with greater demographic complexity and larger class sizes, factors that are well known to be associated with educational outcomes.

4 Empirical strategy and results

Our empirical strategy leverages the random assignment of air purifiers across classrooms. Identification relies on the assumption that treated and control students are comparable in both observable and unobservable characteristics. We also assume that control students do not receive indirect benefits from the treatment or change their behavior due to the absence of purifiers. We discuss potential threats to these assumptions in Section 4.4. Most analyses were pre-registered, and deviations from the PAP are detailed in the text and in Appendix E.

4.1 Impact on indoor air quality

We assess the impact of air purifiers on indoor air quality using Equation 1, where Y_{ct} represents the indoor air quality measure in classroom c on day t; $AirPurifier_c$ indicates whether the classroom received a purifier.[15] We include calendar-day fixed effects (λ_t), strata-by-weekday fixed effects ($\lambda_{s(c),w(t)}$), and strata-by-month fixed effects ($\lambda_{s(c),m(t)}$), where $s(c)$ denotes the randomization stratum (school × grade) of classroom c, $w(t)$ the day of the week, and $m(t)$ the calendar month. Calendar-day fixed effects capture daily pollution shocks that affect all schools. Strata-by-weekday and strata-by-month fixed effects control for location- and grade-specific seasonal factors that may influence air quality. We cluster standard errors at the classroom level.

$$Y_{ct} = \alpha + \beta \, AirPurifier_c + \lambda_t + \lambda_{s(c),w(t)} + \lambda_{s(c),m(t)} + \varepsilon_{ct} \tag{1}$$

Table 2 presents the treatment effects on $PM_{2.5}$ and PM_{10} (Panel A, Columns 1-2). Purifiers significantly reduce concentrations by approximately 4.6 $\mu g/m^3$, which corresponds to a 33% reduction compared to control classes. The result is robust to controlling for indoor temperature and humidity (Appendix Table A.5), different inference methods (Appendix

---

[15] We do not have pre-treatment measurements, as air quality monitors and purifiers were installed simultaneously.

Table 2: Average and dynamic treatment effects on indoor air quality and environmental variables

Panel A: Average treatment effects on indoor environmental variables

	(1)	(2)	(3)	(4)	(5)	(6)
	Indoor air quality			Other environmental variables		
	$PM_{2.5}$	PM_{10}	CO	CO_2	Temp.	N. Ventilation episodes
Estimate	-4.615***	-4.715***	-0.195	31.017	-0.120	0.290
	(0.3791)	(0.4028)	(0.2168)	(54.3764)	(0.1316)	(0.2484)
N.Obs	3,456	3,456	3,456	3,456	3,456	3,461
N.Clusters	30	30	30	30	30	30
Control Mean	14.15	14.85	1.31	805.9	20.97	3.46
Rel. Change %	32.60	31.76	14.96	3.85	0.57	8.38

Panel B: Dynamic treatment effects on indoor $PM_{2.5}$

	(1) 11-2023	(2) 12-2023	(3) 01-2024	(4) 02-2024	(5) 03-2024	(6) 04-2024	(7) 05-2024	(8) 06-2024
Estimate	-4.512***	-6.190***	-7.171***	-8.038***	-3.561***	-1.850***	-2.420***	-2.274***
	(0.4476)	(0.8896)	(0.7922)	(0.7728)	(0.3847)	(0.2204)	(0.2799)	(0.3407)
N.Obs	390	356	503	483	528	475	587	129
N.Clusters	30	30	30	28	28	28	27	27
Control Mean	13.78	21.74	24.47	22.93	9.37	5.88	6.59	6.48
Treated Mean	9.25	15.00	17.17	15.36	6.05	4.13	4.26	4.30
Outdoor Mean	22.48	32.99	44.26	32.88	20.64	11.94	9.77	11.00
Rel. Change %	-32.91	-31.03	-29.82	-32.99	-35.45	-29.70	-35.32	-33.73
I/O Ratio	0.613	0.659	0.553	0.697	0.454	0.492	0.675	0.589

Notes: Panel A reports the average treatment effects (ATE) on indoor air quality measures ($PM_{2.5}$, PM_{10}, CO) alongside effects on other environmental variables (CO_2 and temperature) and the number of ventilation episodes. The sample is restricted to school days. Panel B presents the dynamic treatment effects on indoor $PM_{2.5}$ by calendar month. Models in Panel A include calendar date, strata-by-weekday, and strata-by-month fixed effects (where strata are school × grade, the randomization stratification). Models in Panel B include calendar date, strata-by-weekday, and strata-by-month fixed effects. In Panel B, the number of clusters varies across columns because three monitors went offline during the study window (one in November–December 2023, two from February 2024 onward, with two additional brief interruptions in May and June). Panel-level missingness is uncorrelated with treatment status: 5.0% of control and 6.7% of treated class-months have no $PM_{2.5}$ data, and a linear probability model of monthly missingness on treatment with month fixed effects yields a coefficient of 0.017 ($p = 0.79$). In both panels, Rel. Change % is computed as the regression estimate divided by the control mean; Panel A reports absolute values and Panel B retains the sign to reflect the direction of the effect. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.6)[16], and to multiple hypothesis correction.[17] As expected, we observe no effects on CO since air purifiers do not target this pollutant.

Exploiting the high-frequency sensor data, Panel A of Appendix Table A.7 compares treatment effects during school hours (8am–5pm) and non-school hours. The reduction in $PM_{2.5}$ is more than twice as large outside school hours ($-5.9\ \mu g/m^3$, or 41%) compared to school

---

[16] Conventional clustered standard errors may lead to over-rejecting the null, when the number of clusters is relatively small. We perform wild cluster bootstrap (Cameron et al., 2008) and randomization inference (Young, 2019). Results are similar.

[17] Following the PAP, we control for multiple hypothesis testing for all main pre-registered outcomes (List et al., 2019). In particular, we calculate and report the FDR-adjusted q-values (Benjamini et al., 2006). These are shown in Table 3.

hours (-2.5 $\mu g/m^3$, or 18%). This difference is highly statistically significant ($p < 0.001$) and is consistent with ventilation and pollution-generating activities, e.g., dust movement and blackboard use, during occupied hours, partially offsetting the purifiers' effectiveness. CO_2 concentrations confirm classroom occupancy (1,159 ppm during school hours vs. 595 ppm otherwise). In Panel B, we compare the effects on 24-hour averages in school days vs. non-school days. The absolute treatment effect is larger on school days (-4.6 vs. -3.6 $\mu g/m^3$, $p = 0.0002$), reflecting the higher baseline indoor $\text{PM}_{2.5}$ on occupied days (control mean of 14.1 vs. 10.2 $\mu g/m^3$). Conditional on day-type fixed effects, the proportional reduction is larger on non-school days (34.7% vs. 32.6%, $p = 0.006$), consistent with the same occupancy-and-ventilation interference channel operating on school days.

Panel B of Table 2 presents the dynamic treatment effects, with each coefficient capturing the monthly impact of the intervention. The absolute difference in $PM_{2.5}$ concentrations between treated and control classrooms is largest during the winter months, at 7.6 in February and 7.3 $\mu g/m^3$ in January. In contrast, the absolute differences in April and June are smaller, at 1.8 and 2.3 $\mu g/m^3$, respectively. Although the absolute differences are greater in winter, the relative reduction does not seem to correlate with outdoor air pollution. Sustained reductions between 30% and 35% are found throughout the study period, with no clear seasonal pattern in relative terms.

Indoor air pollution levels appear to be influenced by several factors: outdoor pollution levels, the indoor/outdoor (I/O) ratio, which indicates the extent of $PM_{2.5}$ penetration, and the presence of air purifiers. The I/O ratio ranged from 45.4% to 69.7% (with a weighted average of 59%), with no clear seasonal pattern. Outdoor $PM_{2.5}$ concentrations exceeded the WHO daily limit of 15 $\mu g/m^3$ on 48% of days during the study period. Consequently, indoor $PM_{2.5}$ levels in control classrooms exceeded the WHO threshold on 28.2% of days. This percentage drops to 18.7% in treated classrooms, reflecting the purifiers' mitigating effect.

4.2 Impact on absences

We test the impact of air purifiers on absences using a difference-in-differences design that pools the treatment year (2023–24) with the preceding school year (2022–23):

$$Absent_{ict} = \beta_1 AirPurifier_c \times Post_t + \delta_{s(c),\,g(c)} \cdot t + \lambda_i + \lambda_t + \varepsilon_{ict} \qquad (2)$$

where $Absent_{ict}$ is a binary variable equal to one if student i in class c is absent on date t, and zero otherwise. $AirPurifier_c$ indicates whether class c received an air purifier, while $Post_t$ equals one for dates after the first purifiers were installed on November 8, 2023. The term $\delta_{s(c),\,g(c)} \cdot t$ denotes strata-by-treatment-group linear time trends, where $s(c)$ is the randomization stratum (school × grade) of classroom c and $g(c) \in \{0, 1\}$ is its treatment-group assignment.[18] The model includes student fixed effects (λ_i) to control for time-invariant individual characteristics. Time fixed effects (λ_t) include calendar date, strata-by-month, and strata-by-year, where strata are defined by the randomization stratification (school × grade). These absorb strata-specific seasonal patterns and year-level baseline differences. We cluster standard errors at the classroom level, following Abadie et al. (2023). We estimate Equation 2 using a Linear Probability Model (LPM).[19]

The strata-by-treatment-group linear time trends $\delta_{s(c),\,g(c)} \cdot t$ are a correction not included in the PAP but motivated by the data: when pooling both school years without adjustment, the treatment effect attenuates substantially (-0.002, $p = 0.52$ in the LPM). An examination of pre-treatment absence data from 2022–23 reveals that treated and control classrooms followed different trajectories within randomization strata: a joint F-test across all 95 clusters rejects the null of no differential pre-trends ($p < 0.001$), with 7 of 20 testable strata showing nominally significant differences in stratum-level tests. Because randomization was stratified by school × grade with only 2–6 classrooms per cell, such chance imbalances in classroom-level dynamics are not unexpected. The strata-by-treatment trends absorb these differential trajectories. Appendix F documents the diagnostic evidence, validates the correction through subsample analysis, and shows that a single aggregate treatment trend recovers most of the effect.

Results are shown in Column 2 of Table 3. The LPM estimate is -0.0103, statistically significant at the 5% level, suggesting a reduction of approximately 1.0 percentage point,

---

[18] Defining the treatment period from the date of the first installation is a conservative approach, as classes that had not yet received the purifiers are still assigned treatment status for the period between November 8 and their actual installation date.

[19] Air purifiers reduce concentrations of various airborne substances that can affect respiratory health, such as pollen and viruses. Thus, we cannot use the random installation of purifiers as an instrumental variable for $PM_{2.5}$, as it likely violates the exclusion restriction. Also note that purifiers may introduce unintended effects beyond reducing the concentration of airborne particles, such as noise or light emissions. To mitigate these risks, we turned off all indicator lights and operated the purifiers at reduced speeds, maintaining noise levels within the WHO-recommended threshold for classrooms (35 dB(A)).

about 17% relative to the control mean of 6.0%, or equivalently 2.1 fewer missed days in a 200-day school year. The result is robust across estimators: Probit, Logit, and a Poisson model at the classroom level all yield consistent estimates, significant at the 5% level, with implied effects ranging from 17 to 19.3% (Appendix Table A.8). The results are also robust to alternative inference procedures, including wild bootstrap cluster standard errors and randomization inference (Appendix Table A.9).

The ex-ante MDE from our power analysis (with 5% significance and 80% power) was 0.3 percentage points, 4.5% of the control mean, while the ex-post MDE is 1.4 percentage points for the LPM. The misalignment between ex-ante and ex-post MDE is attributable to higher noise in the outcome, compared to that used in the design phase, where data from only two schools were used.

The result on absences is robust to multiple hypotheses correction ($q = 0.093$). This correction represents a deviation from what was reported in the PAP in that it excludes two primary outcomes, test scores, for which our study was underpowered since the start. Once we include them, the result is not robust to the correction anymore ($q = 0.131$).

Table 3: Average treatment effect on primary outcomes

	(1) $\text{PM}_{2.5}$	(2) Daily Absence (LPM)	(3) Italian Score	(4) Math Score	(5) Stz Raven Score	(6) Mood Scale	(7) Aggressive Episodes
Estimate	-4.615***	-0.0103**	-0.826	0.560	-0.028	0.048	0.029
	(0.379)	(0.0049)	(6.264)	(4.479)	(0.063)	(0.045)	(0.034)
p-value	<0.001	0.037	0.896	0.901	0.660	0.292	0.397
q-value	[<0.001]	[0.093]			[0.660]	[0.487]	[0.497]
q-value (PAP)	[<0.001]	[0.131]	[0.901]	[0.901]	[0.901]	[0.682]	[0.696]
N.Obs	3,456	620,145	700	710	1,808	1,651	1,770
N.Clusters	30	95	37	37	92	92	92
Control Mean	14.15	0.060	190.2	187.1	0.000	3.424	0.395
Control SD	12.25	0.238	49.69	38.20	0.998	0.772	0.489
ex-ante MDE	0.920	0.0030	12.920	9.930	0.690		
ex-ante MDE %SD	0.075	0.013	0.260	0.260	0.130		
ex-post MDE	1.061	0.014	17.539	12.540	0.176	0.126	0.097
ex-post MDE %SD	0.087	0.057	0.353	0.328	0.176	0.164	0.197

Notes: Column 1 is estimated with OLS on the subsample of monitored classrooms. Column 2 pools both school years (2022–23 and 2023–24) and includes student, calendar date, strata-by-month, and strata-by-year fixed effects (where strata are school × grade, the randomization stratification), with strata-by-treatment-group linear time trends to correct for differential pre-treatment dynamics within randomization strata (see Appendix F); estimated with a Linear Probability Model (LPM). Columns 3–4 (INVALSI) include gender, grade, and above-median pre-treatment absences as controls with school fixed effects, as privacy restrictions prevent richer specifications. Columns 5–7 are estimated with OLS including strata and survey-wave fixed effects. Standard errors are clustered at the classroom level. FDR-adjusted q-values (Benjamini et al., 2006) are reported in brackets. Ex-ante MDE values are from the pre-analysis plan power calculations, where available. Ex-post MDE is computed as 2.8 × SE, corresponding to 80% power at the 5% significance level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4: Heterogeneous effects on school absences

	(1) Pre-treatment absences	(2) Outdoor $PM_{2.5}$	(3) Seasonality
Treat × Post	-0.0035	-0.0149^{***}	-0.0135^{**}
	(0.0045)	(0.0054)	(0.0053)
Treat × Post × Above median	-0.0151^{*}	0.0062^{*}	
	(0.0080)	(0.0032)	
Treat × Post × Winter			0.0068^{*}
			(0.0035)
N.Obs	620,145	605,985	620,145
N.Clusters	95	95	95
Control Mean ($M = 0$)	0.039	0.061	0.065
Control Mean ($M = 1$)	0.081	0.056	0.050
Total TE ($M = 0$)	-0.0035	-0.0149^{***}	-0.0135^{**}
	(0.0045)	(0.0054)	(0.0053)
Total TE ($M = 1$)	-0.0186^{**}	-0.0087^{*}	-0.0068
	(0.0079)	(0.0050)	(0.0050)
q-value	[0.062]	[0.062]	[0.062]

Notes: The dependent variable is a binary indicator for daily absence. All columns are estimated with a Linear Probability Model pooling school years 2022–23 and 2023–24, with student, calendar date, strata-by-month, and strata-by-year fixed effects (where strata are school × grade), and strata-by-treatment-group linear time trends. Column 1 interacts the treatment effect with an indicator for above-median pre-treatment absence rates. Column 2 interacts the treatment effect with an indicator for above-median outdoor $PM_{2.5}$ (seven-day lagged rolling average), and includes a post × above-median pollution interaction to allow for differential post-treatment shifts by pollution level. Column 3 interacts the treatment with a winter indicator (December 21 to March 20), and includes a post × winter interaction. All three heterogeneity dimensions follow the pre-analysis plan. Bottom rows report the total treatment effect (TE) for each subgroup, with delta-method standard errors. FDR q-values (Benjamini et al., 2006) are reported in brackets, computed across the three triple-interaction coefficients (the three pre-specified heterogeneity dimensions). Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4 examines heterogeneous treatment effects along three dimensions pre-specified in the PAP: pre-treatment absence levels (above vs. below the median), outdoor air pollution (above vs. below the preceding 7-day rolling average), and seasonality (Winter vs. Spring and Fall months). To do so, we augment Equation 2 by interacting the treatment with a binary moderator M:

$$\begin{aligned} Absent_{ict} &= \beta_1 \left(AirPurifier_c \times Post_t\right) + \beta_2 \left(AirPurifier_c \times Post_t \times M_i\right) \\ &\quad + \left(\delta_{s(c),\,g(c)} + \gamma_{s(c),\,g(c)}\, M_i\right) \cdot t + \lambda_i + \lambda_{t,\,M_i} + \varepsilon_{ict} \end{aligned} \tag{3}$$

where β_1 captures the treatment effect for the reference group ($M_i = 0$) and β_2 the differential treatment effect for the group of interest ($M_i = 1$). The strata-by-treatment-group time trends are fully interacted with M_i (slope $\gamma_{s(c),g(c)}$), and the subgroup-specific time fixed effects λ_{t,M_i}—which include date, strata-by-month, and strata-by-year fixed effects estimated separately for each M_i group—absorb all differential time-varying patterns across subgroups, including the post-period level shift $Post_t \times M_i$. This fully interacted specification applies to Column 1, where the moderator is time-invariant. For the time-varying moderators in Columns 2 and 3, the specification retains a $Post_t \times M_t$ lower-order term but does not interact fixed effects with M_t.[20]

Column 1 of Table 4 interacts the treatment with an indicator for above-median pre-treatment absence rates, using the fully interacted specification of Equation 3. The treatment effect is concentrated among high-absence students: the interaction term is -0.0151, significant at the 10% level ($q = 0.062$). The implied total treatment effect for above-median students is -1.9 percentage points ($p = 0.019$), equivalent to 23% of their control mean of 8.1%. The treatment effect for below-median students is -0.4 percentage points, statistically indistinguishable from zero ($p = 0.44$).

Appendix Table A.10 presents robustness using finer partitions. Under the split-sample specification, the effect is concentrated in the highest-absence quartile (-2.8 percentage points, $p = 0.026$), with near-zero effects in the lower three quartiles. A continuous specification yields a slope of -0.187 on Treatment × PreAbs ($p < 0.001$), implying that a 10 percentage point increase in a student's pre-treatment absence rate corresponds to an additional 1.9 percentage point reduction in their treatment effect.

We also examine heterogeneity by student characteristics (Appendix Table A.11 and A.12). There is no evidence of differential effects by gender and grade. However, we find suggestive evidence that the treatment is more effective for students with foreign citizenship, who tend to have higher pre-treatment absence rates. The average number of pre-treatment absences among Italian students is approximately 25% lower than that of foreign students.

The top panel of Figure 2 presents bimonthly dynamic treatment effects on absences. The treatment effect is negative and statistically significant at the 5% level in November–December 2023 and March–June 2024, but not during the peak winter months (January–February 2024), when outdoor pollution levels are highest. Importantly, there is no evidence of fade-out: the spring 2024 coefficients are at least as large as those in November–December 2023, suggesting that purifier effectiveness in reducing absences is sustained over the school

---

[20] The split-sample equivalence motivating Column 1 requires time-invariant group membership; with a time-varying M_t, the same student belongs to different groups on different days. In Column 3, the winter indicator is a deterministic function of the calendar date, so λ_{t,M_t} reduces mechanically to λ_t. In Column 2, the $PM_{2.5}$ indicator is school-specific; interacting date fixed effects with it would absorb the cross-school pollution variation that identifies the heterogeneous effect.

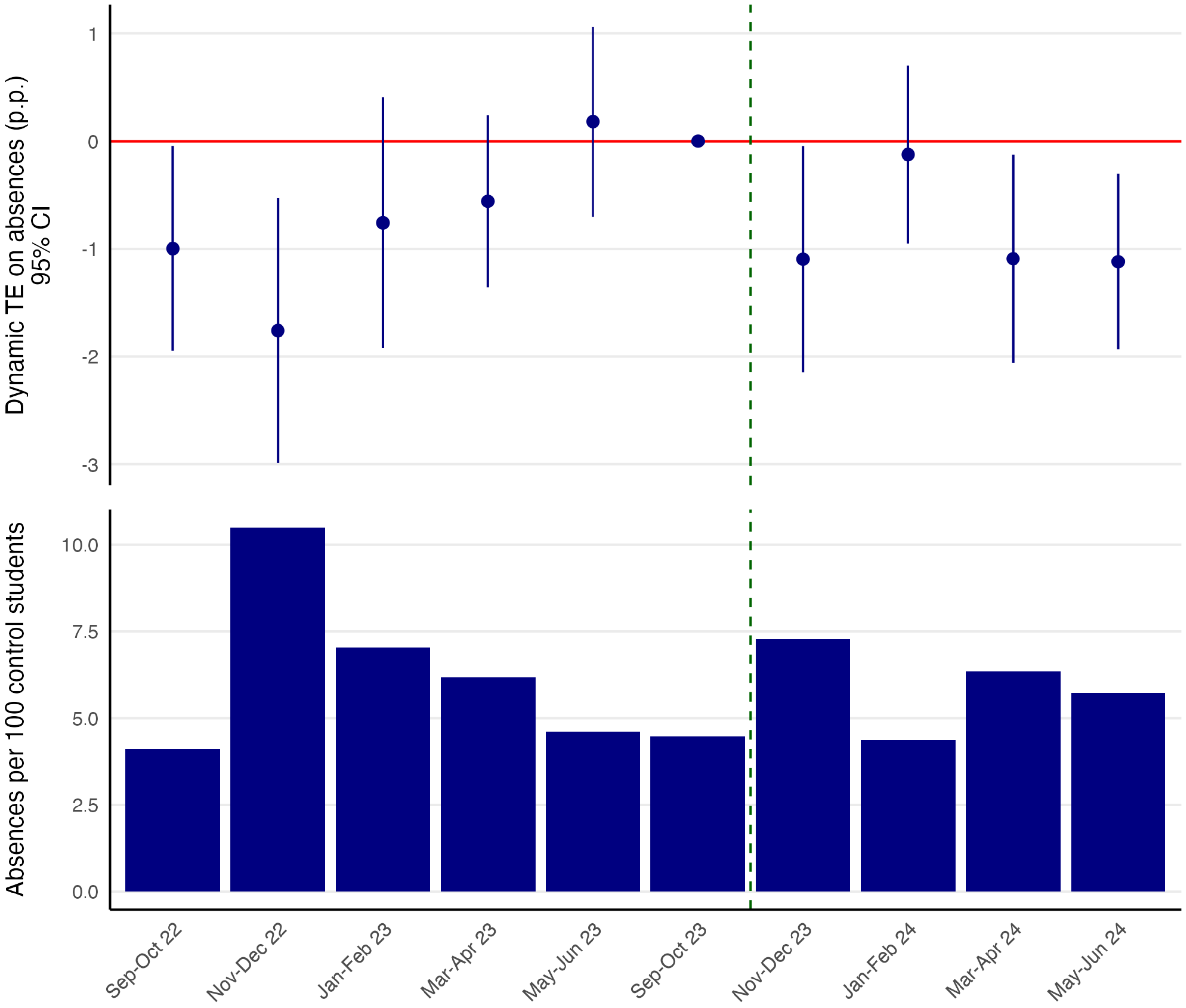


Figure 2: Dynamic effects on absences

Notes: The top panel shows bimonthly dynamic treatment effects on daily absences with 95% confidence intervals, estimated with a two-year LPM including student, calendar date, strata-by-month, and strata-by-year fixed effects. Strata-by-treatment-group time trends are excluded from the event study to avoid collinearity between unit-specific trends and event-study leads (Borusyak et al., 2024). The reference period is September–October 2023, and the dashed vertical line marks the start of treatment (November 8, 2023). The bottom panel shows the seasonality of absences in the control group over the same bimonthly bins, expressed as average absences per 100 control students. Standard errors are clustered at the classroom level.

year. The temporal variation in treatment effects instead tracks seasonal pollution patterns, with the largest reductions occurring when outdoor concentrations are moderate and purifiers can bring indoor levels closer to health-relevant thresholds. The bottom panel displays the seasonality of absences in the control group, aggregated to the same bimonthly bins. The seasonal pattern of treatment effects does not appear to result from seasonal trends in absenteeism. The pre-treatment coefficients exhibit a slight upward drift, consistent with the differential pre-treatment trends within randomization strata documented in Appendix

F. This pattern motivates the strata-by-treatment-group time trends included in the level estimates of Tables 3 and 4.[21]

Column 3 of Table 4 formalizes this pattern by interacting the treatment with a winter indicator (December 21 to March 20), as pre-specified in the PAP. The treatment effect outside winter is -1.4 percentage points, statistically significant at the 5% level, and is approximately halved during winter months, leading to a non-significant effect (implied effect: -0.7 percentage points, $p = 0.173$). The winter interaction is marginally significant at the 10% level.

Column 2 of Table 4 interacts the treatment with an indicator for above-median outdoor $PM_{2.5}$ levels, measured as the seven-day lagged rolling average, as pre-specified in the PAP. Air purifiers significantly reduce absences on low-pollution days, with an implied effect of -1.5 percentage points ($p = 0.0029$). The interaction term is positive and statistically significant at the 10% level, indicating that the treatment effect is approximately halved on high-pollution days (-0.87 percentage points, $p = 0.082$). Appendix Table A.13 presents a finer quartile decomposition using both rolling $PM_{2.5}$ averages and WHO daily exceedance counts, confirming a monotonic attenuation: the treatment effect declines steadily from the cleanest to the most polluted days, becoming statistically indistinguishable from zero in the highest pollution quartile.

This pattern can reflect a concave dose-response relationship between air pollution and health. While the purifiers reduce indoor $PM_{2.5}$ by a stable 29–38% throughout the year (Panel B of Table 2), the health gains from a given percentage reduction depend on where students sit on the dose-response curve. On low-pollution days (outdoor $PM_{2.5} \approx 9\ \mu g/m^3$), the purifier brings indoor concentrations well below WHO guidelines, effectively eliminating harmful exposure during school hours. On high-pollution days (outdoor $PM_{2.5} \approx 32\ \mu g/m^3$), the purifier reduces indoor levels from approximately 22 to 15 $\mu g/m^3$, but students remain exposed to concentrations near or above health-relevant thresholds for the full school day, and are further exposed to unfiltered outdoor air during commutes and, potentially, at home.

A complementary explanation is that winter absences are disproportionately driven by influenza and other viral infections transmitted through close interpersonal contact, a channel that air filtration does not address. Appendix Figure A.1 shows that control-group absence

---

[21] Strata-by-treatment-group time trends are excluded from the event study to avoid the collinearity between unit-specific trends and event-study leads identified by Borusyak et al. (2024). The dynamic treatment effects depicted represent lower bounds, as they are not detrended.

peaks coincide with influenza-like illness incidence in Lombardy.[22] Because flu season and high-pollution season largely overlap in Milan's climate (Figure 2), the two mechanisms are difficult to disentangle.

A third possibility is that seasonal allergens explain the stronger spring effects. To test this, we estimate heterogeneous treatment effects based on pollen concentration using a 10-day rolling average. We find no evidence that purifier effectiveness varies with pollen levels (Appendix Table A.14), reducing the likelihood that this channel drives the seasonal pattern.

4.3 Impact on cognitive performance and behavior

We test the intervention's impact on cognitive performance and behavior using:

$$Y_{ich} = \alpha + \beta AirPurifier_{ch} + \gamma X_{ich} + \lambda_h + \varepsilon_{ich} \tag{4}$$

Each endline outcome Y for student i in class c and school h is regressed on a treatment indicator and a vector of student characteristics X, including gender, citizenship, and grade, along with school fixed effects (λ_h).[23] We cluster standard errors at the classroom level.[24]

Models of test scores are estimated using data provided by INVALSI. In accordance with institutional privacy rules, these data were further anonymized and could not be matched to our full original dataset. Specifically, INVALSI data could only be linked to a dataset in which individual units cannot be identified.[25] As a result, the set of controls in the specification is severely limited to gender, grade, and an indicator for above-median pre-treatment absences.

Columns 3–7 of Table 3 report treatment effects on Italian and mathematics standardized test scores (INVALSI), cognitive skills (Raven), mood, and aggressive episodes. None of these outcomes show statistically significant effects. For test scores, both ex-ante (0.26 SD) and ex-post MDEs (0.33–0.35 SD) are substantially larger than the 0.1–0.2 SD effects

---

[22] We use regional monthly data on average influenza-like-illness (ILI) among children aged 5-14 (Branda, 2024). The lack of more granular data, both in time and space, impedes the possibility to analyse heterogeneous effects, similarly to what we do below for pollens.

[23] The school index is denoted by h to avoid collision with s, which we reserve for the randomization stratum (school × grade) introduced in equations 2 and 3.

[24] This specification represents a deviation from PAP, where we committed to using ANOVA for outcomes measured in the survey. Our decision is driven by the very high attrition and non-responses over the two waves which lead to sample reductions of 33–45% for some outcomes. While ex-post MDEs are similar in the two specifications, we prefer reporting results on a broader sample, while presenting the ANOVA models in Appendix Table A.15. Results are similar.

[25] This requires that each combination of categories appear at least three times in the dataset to be linked.

typically found in the literature (Gilraine, 2023; Bharti et al., 2025). The study is therefore underpowered and uninformative with respect to test score impacts. The imprecision of the test score estimates reflects the limited INVALSI sample: only second- and fifth-grade students take the national exam, reducing the effective sample to approximately 700 students in 37 classrooms.[26]

For the Raven score, both ex-ante (0.13 SD) and ex-post MDEs (0.18 SD) fall within the range of effects reported in the literature (e.g., Bharti et al. (2025)), indicating that the study is sufficiently powered to rule out effects of that magnitude.

Similarly, mood and aggressive episodes exhibit relatively small MDEs (0.16–0.2 SD), yet no statistically significant effects are detected. This suggests that moderate treatment effects on these outcomes can be ruled out.

4.4 Impact on self-reported health symptoms, perceptions, preferences, and adaptive behaviors

We assess the intervention's impact on self-reported health symptoms, perceptions of air quality in different environments, and preferences for various urban policies using the same model in 4 estimated via OLS.[27]

Panel A of Table 5 reports the estimated effects on self-reported health symptoms. We observe negative treatment effects for respiratory-related symptoms, with statistically significant reductions in the incidence of runny nose and blocked nose at the 10% level. The effect sizes for these symptoms range from 4.6 to 5.9 p.p., indicating decreases of approximately 9% and 11% relative to the control group mean. Columns 6 to 8 show no significant effects for general or non-respiratory symptoms.[28] These results provide suggestive evidence that the treatment operates through improvements in respiratory health.

---

[26] Grade retention is also uninformative in this context. Italian primary schools have near-universal promotion: legislation requires that students be admitted to the next class even when learning levels are only partially achieved, with retention permitted only in exceptional and unanimously agreed cases (Italian Republic, 2017). Italy's primary completion rate exceeds 98% (World Bank, 2023).

[27] Once again, the model represents a deviation from the PAP where we committed to use ANOVA. ANOVA results controlling for baseline outcomes are shown in Appendix Table A.16. Findings are qualitatively similar, despite a 33–45% reduction in sample size due to cross-wave attrition and non-response. In particular, the estimated effect on the "runny nose" symptom is consistent in both magnitude and precision, the effect on "blocked nose" declines and becomes statistically insignificant, while the effect on "stomach ache" increases and becomes statistically significant.

[28] Prior work has documented associations between $PM_{2.5}$ exposure and headache incidence, particularly for migraine (Portt et al., 2023). The null result in our sample may reflect limited statistical power for individual symptom-level effects.

Table 5: Impact on Self-Reported Health Symptoms, Perceptions, and Preferences

Panel A: Impact on Self-Reported Health

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Symptoms							
	Runny nose	Blocked nose	Sneezing	Cough	Short breath	Tiredness	Headache	Stomachache
Estimate	-0.059*	-0.046*	-0.035	-0.023	0.026	0.010	0.002	0.013
	(0.031)	(0.026)	(0.026)	(0.033)	(0.031)	(0.029)	(0.026)	(0.026)
N.Obs	1,589	1,597	1,587	1,621	1,555	1,570	1,571	1,574
N.Clusters	89	89	89	89	89	89	89	89
Control Mean	0.532	0.569	0.664	0.576	0.259	0.646	0.452	0.403
Control SD	0.499	0.496	0.473	0.494	0.438	0.479	0.498	0.491
Ex-post MDE	0.087	0.073	0.072	0.093	0.087	0.081	0.074	0.073

Panel B: Impact on Perceptions and Preferences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Air Quality Perception				Urban Policy Priorities				
	General	City	Class	Schoolyard	City cleaning	Green playgrounds	Sport infrastructure	Air quality	Less traffic
Estimate	0.102	-0.036	0.154**	-0.015	0.065	0.100	0.007	0.132**	-0.010
	(0.080)	(0.082)	(0.070)	(0.075)	(0.046)	(0.069)	(0.071)	(0.053)	(0.059)
N.Obs	1,715	1,742	1,733	1,725	1,729	1,720	1,704	1,716	1,695
N.Clusters	92	92	92	92	92	92	92	92	92
Control Mean	2.978	2.559	3.109	3.360	3.509	3.280	3.088	3.456	3.158
Control SD	0.880	0.983	0.793	0.808	0.871	0.980	1.074	0.977	1.014
Ex-post MDE	0.223	0.229	0.196	0.209	0.128	0.192	0.200	0.148	0.165

Notes: Panel A reports OLS estimates of the treatment effect on self-reported health symptoms. The dependent variables equal one if the student reported the symptom at least some time over the previous week and zero otherwise. Panel B presents estimates of the impact on perceptions and policy preferences using OLS. The first four columns report an air quality perception index (scored from 1 (very bad) to 4 (very good)), and the next five columns report a policy priority score on urban issues (scored from 1 (not important at all) to 4 (very important)). Models control for gender and foreign nationality and include grade and school fixed effects; standard errors are clustered at the treatment (classroom) level. Sample sizes vary across symptom outcomes due to item-level non-response in the endline survey. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1.

We investigate how classroom air purifiers affect students' perceptions of air quality in different environments. Panel B of Table 5 shows that students in treated classrooms report significantly higher perceptions of air quality than students in control classrooms. In contrast, we find no significant differences in perceptions of overall city air quality or the schoolyard (Columns 1–4). Moreover, the treatment significantly increased priority scores for green policies, including city cleaning, green playgrounds, and air quality. However, the effect is statistically significant only for air quality (Columns 5-9).

However, one limitation is the potential for experimenter demand effects—treated students may feel pressured to report symptoms, perceptions, and preferences that align with the perceived goals of the intervention. Additionally, the physical presence of purifiers may prime students' responses. To support the validity of self-reported health measures, we examine their correlation with pre-treatment absences. We find a positive and borderline statistically significant (at 10% level) relationship between baseline reports of runny nose and pre-treatment absences (Appendix Table A.20). Notably, this symptom coincides with that for which we observe significant treatment effects in Table 5, lending credibility to the self-reported outcome. This finding also has policy implications: students who report respiratory symptoms at baseline are both more likely to be absent and more likely to benefit from improved air quality. Air purifiers may therefore be especially valuable in schools or classrooms with high prevalence of respiratory conditions, offering a targeted rationale for deployment in areas with elevated pollution or allergy burden.

Parents were not informed about the intervention and did not receive the survey, which was administered to students in the classroom. We cannot rule out that children in treated classrooms communicated about the purifiers to their families, potentially triggering home air quality investments. To the extent that such responses occurred, our estimates would capture the total effect of the school-based intervention, including any induced private behavioral changes.

Changes in classroom air quality perception may prompt behavioral adaptations. Students and teachers in treated classrooms might modify ventilation practices, such as reducing the frequency of window openings, or alter decisions about classroom occupancy. These behavioral responses could compromise the study's identification strategy if they correlate with treatment status, leading to conflated treatment effect estimates that mix the direct impact of purifiers with the effects of behavioral adaptation. To assess potential behavioral adaptation, we compare average daily levels of indoor CO_2, temperature, and estimated window-opening events between treatment and control classrooms, using a specification similar to Equation 1. These variables depend on student density, class duration, and ventilation frequency. As shown in Panel A of Table 2 (Columns 4–6), we observe no significant differences in these environmental measures between the treatment and control groups. Since window-opening behavior may vary seasonally, we also test whether treatment affects ventilation patterns across seasons. We find no significant relationship between treatment status and window-opening frequency, either overall or by season (Appendix D).

5 Discussion

Cost-effectiveness We assess the cost-effectiveness of our intervention by calculating the cost per avoided student absence. The intervention's total cost is approximately €2,910 per purifier over a 10-year lifespan, which includes purchase, installation, maintenance, and energy use.[29] In a typical classroom with 21.6 students, this results in an annual cost of about €14.18 per student.[30] Since the intervention reduces absences by approximately 2.06 days per student per year, the estimated cost per avoided absence day is €6.88. Assuming a shorter purifier's lifetime of 5 years, the figure increases to €13.77. The manufacturer provides a three-year warranty and recommends filter replacement every three years. All 42 purifiers installed in November 2023 remain fully operational as of 2026. Chowdhury et al. (2026), Ruiz-Tagle et al. (2024), and Kremer et al. (2025) use cheaper air purifiers with HEPA filters and lower CADR, bringing down the yearly cost per student to €2.86.

In high-income contexts, interventions designed to reduce absenteeism among at-risk students often include mentorship programs and behaviorally informed attendance reports sent to parents (Rogers and Feller, 2018; Heppen et al., 2018; Robinson et al., 2018; Bergman, 2021; Guryan et al., 2021). In lower-income settings, deworming is considered one of the most cost-effective strategies for improving both attendance and academic outcomes (Miguel and Kremer, 2004). The cost per additional day of attendance varies widely across interventions—from approximately €7.10 for deworming to €9–15 for behavioral interventions that reduce information frictions, and over €580 for hiring dedicated support staff.[31] Our intervention falls at the lower end of this cost range. Additionally, unlike information-based

---

[29] This includes a bulk purchase price of €2,000, three replacement filters at €250 each (every 3 years), and approximately €314 in electricity costs. The latter is based on manufacturer data: assuming 14 hours of operation per school day at speed 3 (56W), over 200 school days per year and 10 years, total consumption is 1,568 kWh. At an electricity price of €0.20/kWh, this corresponds to a total cost of about €314.

[30] The average class size in primary education in OECD countries is 20.6 (OECD, 2025).

[31] Original cost values were converted to euros and adjusted for inflation to ensure comparability in 2024. It is expected that traditional interventions, although more expensive, provide other benefits aside from just avoided absences.

treatments, air purifiers are likely to produce lasting effects without behavioral mean reversion or externalities.

We calculate the benefits of avoided health-related school absences, considering only two immediate costs: healthcare and childcare expenditure. To be conservative, we do not consider the potential direct and indirect benefits of improved air quality and avoided absences on academic performance and long-term human capital (Liu et al., 2021; Cattan et al., 2022; Goodman, 2014). Federici et al. (2018) estimate the healthcare cost of a single school day missed due to influenza-related illness in Italy at approximately €73. Assuming 8 hours of childcare at the nationally contracted minimum salary for childcare of €7.10 (Ministero del Lavoro e delle Politiche Sociali, 2025), the total daily cost, ignoring lost parental productivity, is €130. Based on this conservative estimate, the benefit-cost ratio ranges between 9.5 and 19 for purifier lifetimes of 5 and 10 years, respectively, when considering only the immediate economic costs of absenteeism.

This estimate should be interpreted as an upper bound on the benefit-cost ratio. The per-absence benefit of €130 is calibrated to influenza-related illness, which typically involves physician visits and multi-day recovery. The marginal absence prevented by air purifiers may involve milder respiratory symptoms with lower healthcare utilization. Moreover, students who attend school thanks to improved air quality may still experience residual symptoms that reduce their engagement, an effect our binary attendance measure cannot capture. That said, the cost-effectiveness results remain robust to substantial downward adjustments. Even when considering only childcare costs and excluding healthcare benefits, the benefit–cost ratio remains between 4.1 and 8.3 (for 5- and 10-year purifier lifetimes, respectively).

Results discussion Air purifiers reduce indoor air pollution by 33% on average and by 18% during school hours. These effects are large relative to the existing literature. Comparable studies report reductions of 6–8% in settings with similar outdoor pollution levels (Kremer et al., 2025) and around 20% in more polluted environments (Bharti et al., 2025). The relatively large effect in our context is likely driven by the higher filtration capacity of the devices installed.

We find that purifiers reduce absenteeism by approximately 17%, corresponding to 2.1 fewer missed school days per student per year. This effect is economically meaningful and compares favorably with prior evidence exploiting variation in outdoor pollution. For example, Komisarow and Pakhtigian (2022) document a reduction of 0.66 days following coal plant closures in Chicago. Chen et al. (2018) show that a 10-unit increase in the Air Quality

Index increases absence rates by 2.31% of the daily mean in China. In Texas, Currie et al. (2009) find that high levels of carbon monoxide reduce school attendance. Persico and Venator (2019) and Heissel et al. (2022) report reductions of 0.6 and 0.82 annual absences, respectively, following changes in local pollution exposure.

Treatment effects are concentrated among students with higher pre-treatment absence rates, consistent with evidence that the burden of air pollution is heterogeneous (Liu and Salvo, 2018). This pattern may reflect larger health gains among more vulnerable students or differential behavioral responses by households. For instance, families with more health-sensitive children may be more likely to respond to symptoms by keeping them at home. While our design does not allow us to disentangle these mechanisms, the evidence on the reduction in reported respiratory symptoms suggests that improved health is a primary channel (Currie et al., 2009; Mendoza et al., 2020).

We find no statistically significant effects on cognitive outcomes or behavior. Although the study is underpowered to detect effects on academic performance, it is sufficiently powered to rule out moderate effects on fluid intelligence and aggressive behavior. In particular, we can rule out effect sizes comparable to those reported in Bharti et al. (2025). Nevertheless, the absenteeism channel alone carries substantial economic relevance, as missed school days generate immediate costs and have long-term consequences for human capital formation (Liu et al., 2021; Cattan et al., 2022; Goodman, 2014).

Finally, treatment effects vary with ambient pollution levels. The impact on absenteeism is larger when outdoor pollution is relatively low and attenuates as pollution intensifies. This pattern is consistent with concave concentration–response relationships between $PM_{2.5}$ exposure and health (Pope III et al., 2015; Nasari et al., 2016; Corrigan et al., 2018; Weichenthal et al., 2022), and with experimental evidence showing limited health gains when pollution reductions occur at very high baseline exposure levels (Berkouwer and Dean, 2026; Miller et al., 2024). In high-pollution environments, the reductions achieved by purifiers may be insufficient to bring exposure below harmful thresholds, limiting their effectiveness.

6 Conclusion

In this paper, we evaluate the impact of installing portable air purifiers in schools. The study occurs in a developed region characterized by moderate ambient pollution and standard school infrastructure. Our findings indicate that air purifiers effectively lower indoor pollution and decrease student absenteeism, especially among the most vulnerable students. The intervention is cost-effective and straightforward to deploy.

The effects on absences are most pronounced during periods of moderate outdoor pollution. During high-pollution episodes, the reduction in $PM_{2.5}$ achieved by purifiers could be insufficient, as students remain exposed to elevated pollution levels both indoors and outdoors. Although our study is not well-suited to fully test this mechanism, the results suggest that in areas with severe air pollution, purifiers alone may have limited impacts on health outcomes and absenteeism. Improving indoor air quality in such settings may require reducing infiltration through better building infrastructure or increasing the capacity and operating intensity of purifiers. More broadly, air purifiers should complement—not substitute for—policies aimed at reducing emissions, raising awareness, and encouraging adaptive behaviors.

References

Abadie A., Athey S., Imbens G. W., and Wooldridge J. M. (2023) When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, 138(1): 1–35. 17

Aguilar-Gomez S., Dwyer H., Graff Zivin J., and Neidell M. (2022) This Is Air: The "Non-health" Effects of Air Pollution. *Annual Review of Resource Economics*, 14(Volume 14, 2022): 403–425. 2

Angrist J. et al. (2020) *How to Improve Education Outcomes Most Efficiently?: A Comparison of 150 Interventions Using the New Learning-adjusted Years of Schooling Metric*: World Bank Group. 4

Annesi-Maesano I., Maesano C. N., Biagioni B., D'Amato G., and Cecchi L. (2021) Call to action: Air pollution, asthma, and allergy in the exposome era. *Journal of Allergy and Clinical Immunology*, 148(1): 70–72. 2

Aragón F. M., Miranda J. J., and Oliva P. (2017) Particulate matter and labor supply: The role of caregiving and non-linearities. *Journal of Environmental Economics and Management*, 86: 295–309, Special issue on environmental economics in developing countries. 2

Barwick P. J., Li S., Lin L., and Zou E. Y. (2024) From Fog to Smog: The Value of Pollution Information. *American Economic Review*, 114(5): 1338–81. 5

Baylis P., Greenstone M., Lee K., and Sahai H. (2024) Is the Demand for Clean Air Too Low? Experimental Evidence from Delhi. 5

Benjamini Y., Krieger A. M., and Yekutieli D. (2006) Adaptive linear step-up procedures that control the false discovery rate. *Biometrika*, 93(3): 491–507. 15, 18, 19, 53, 55, 56, 57

Bergman P. (2021) Parent-child information frictions and human capital investment: Evidence from a field experiment. *Journal of political economy*, 129(1): 286–322. 27

Berkouwer S. B. and Dean J. T. (2026) Cooking, Health, and Daily Exposure to Pollution Spikes. *American Economic Journal: Economic Policy*, 18(2): 511–47. 2, 29

Bharti N. K., Blakeslee D., Figueiredo Walter T., Malik S., and Naeem T. (2025) Clean Air in the Classroom: Environmental Inputs and Human Capital Formation. 4, 24, 28, 29

Borusyak K., Jaravel X., and Spiess J. (2024) Revisiting Event-Study Designs: Robust and Efficient Estimation. *Review of Economic Studies*, 91(6): 3253–3285. 21, 22

Branda F. (2024) InfluNet: Italian Influenza Surveillance Data. Data from the Italian National Institute of Health (ISS) RespiVirNet surveillance system. 23, 59

Burke M., Heft-Neal S., Li J., Driscoll A., Baylis P., Stigler M., Weill J. A., Burney J. A., Wen J., Childs M. L., and Gould C. F. (2022) Exposures and behavioural responses to wildfire smoke. *Nature Human Behaviour*, 6(10): 1351–1361, Publisher: Nature Publishing Group. 5

Cameron A. C., Gelbach J. B., and Miller D. L. (2008) Bootstrap-Based Improvements for Inference with Clustered Errors. *The Review of Economics and Statistics*, 90(3): 414–427. 15, 44, 47

Carmona N., Seto E., Gould T. R., Rasyid E., Shirai J. H., Cummings B. J., Hayward L., Larson T. V., and Austin E. (2022) Indoor Air Quality Intervention in Schools: Effectiveness of a Portable HEPA Filter Deployment in Five Schools Impacted by Roadway and Aircraft Pollution Sources. *Atmosphere*, 13(10): 1623, Number: 10 Publisher: Multidisciplinary Digital Publishing Institute. 4

Carneiro J., Cole M., 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*. 3

Cattan S., Kamhöfer D. A., Karlsson M., and Nilsson T. (2022) The Long-Term Effects of Student Absence: Evidence from Sweden. *The Economic Journal*, 133(650): 888–903. 28, 29

Cellini S. R., Ferreira F., and Rothstein J. (2010) The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design*. *The Quarterly Journal of Economics*, 125(1): 215–261. 4

Cheek E., Guercio V., Shrubsole C., and Dimitroulopoulou S. (2021) Portable air purification: Review of impacts on indoor air quality and health. *The Science of the Total Environment*, 766: 142585. 4

Chen R., Zhao A., Chen H., Zhao Z., Cai J., Wang C., Yang C., Li H., Xu X., Ha S. et al. (2015) Cardiopulmonary benefits of reducing indoor particles of outdoor origin: a randomized, double-blind crossover trial of air purifiers. *Journal of the American College of Cardiology*, 65(21): 2279–2287. 4

Chen S., Guo C., and Huang X. (2018) Air Pollution, Student Health, and School Absences: Evidence from China. *Journal of Environmental Economics and Management*, 92: 465–497. 4, 28

Chiu Y.-H. M., Bellinger D. C., Coull B. A., Anderson S., Barber R., Wright R. O., and Wright R. J. (2013) Associations between Traffic-Related Black Carbon Exposure and Attention in a Prospective Birth Cohort of Urban Children. *Environmental Health Perspectives*, 121(7): 859–864. 3

Chowdhury A., Garg T., Jagnani M., and Mattsson M. (2026) Residential Air Purifiers and the Value of Clean Air. 27

Cori L., Donzelli G., Gorini F., Bianchi F., and Curzio O. (2020) Risk Perception of Air Pollution: A Systematic Review Focused on Particulate Matter Exposure. *International*

Journal of Environmental Research and Public Health, 17(17): 6424. 11

Corrigan A. E., Becker M. M., Neas L. M., Cascio W. E., and Rappold A. G. (2018) Fine particulate matters: The impact of air quality standards on cardiovascular mortality. *Environmental research*, 161: 364–369. 2, 29

Coury M., Falconer L., and La Nauze A. (2024) Wildfire smoke and private provision of public air-quality monitoring. *Journal of Environmental Economics and Management*, 127: 103036. 5

Currie J., Hanushek E. A., Kahn E. M., Neidell M., and Rivkin S. G. (2009) Does pollution increase school absences? *The Review of Economics and Statistics*, 91(4): 682–694. 2, 4, 29

Deschênes O., Greenstone M., and Shapiro J. S. (2017) Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program. *American Economic Review*, 107(10): 2958–2989. 5

Ebenstein A., Lavy V., and Roth S. (2016) The Long-Run Economic Consequences of High-Stakes Examinations: Evidence from Transitory Variation in Pollution. *American Economic Journal: Applied Economics*, 8(4): 36–65. 3

EEA (2022) Air Quality e-Reporting: Air quality annual statistics calculated by the EEA. 2

——— (2024) `https://www.eea.europa.eu/en`. 11

ENVIRA (2024a) `https://envira.global/nanoenvi-iaq/`. 61

——— (2024b) `https://envira.global/wp-content/uploads/2022/03/EN_IAQ_LORA_Data_Sheet_V1.3.3.pdf`. 61

European Environmental Agency (2024) Europe's air quality status 2024.Technical report. 6

Federici C., Cavazza M., Costa F., and Jommi C. (2018) Health care costs of influenza-related episodes in high income countries: A systematic review. *PLOS ONE*, 13(9): e0202787, Publisher: Public Library of Science. 28

Gehrsitz M. (2017) The effect of low emission zones on air pollution and infant health. *Journal of Environmental Economics and Management*, 83: 121–144. 7

Gershenson S., Jacknowitz A., and Brannegan A. (2017) Are student absences worth the worry in US primary schools? *Education Finance and Policy*, 12(2): 137–165. 2

Gignac F., Barrera-Gómez J., Persavento C., Solé C., Tena È., López-Vicente M., Foraster M., Amato F., Alastuey A., Querol X. et al. (2021) Short-term effect of air pollution on attention function in adolescents (ATENC! Ó): A randomized controlled trial in high schools in Barcelona, Spain. *Environment international*, 156: 106614. 4

Gilraine M. (2023) Air Filters, Pollution, and Student Achievement. *Journal of Human Resources*. 4, 24

Gilraine M. and Zheng A. (2024) JUE insight: Air pollution and student performance in the U.S.. *Journal of Urban Economics*, 143: 103686. 3

Goodman J. (2014) Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time. Working Paper 20221. 28, 29

Graff Zivin J. and Neidell M. (2013) Environment, Health, and Human Capital. *Journal of Economic Literature*, 51(3): 689–730. 4, 5

Greenstone M., Lee K., and Sahai H. (2021) Indoor Air Quality, Information, and Socioeconomic Status: Evidence from Delhi. *AEA Papers and Proceedings*, 111: 420–424. 5

Guryan J., Christenson S., Cureton A., Lai I., Ludwig J., Schwarz C., Shirey E., and Turner M. C. (2021) The Effect of Mentoring on School Attendance and Academic Outcomes: A Randomized Evaluation of the Check & Connect Program. *Journal of Policy Analysis and Management*, 40(3): 841–882. 27

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. 4, 29

Heppen J. B., Zeiser K., Holtzman D. J., O'Cummings M., Christenson S., and Pohl A. (2018) Efficacy of the Check & Connect mentoring program for at-risk general education high school students. *Journal of Research on Educational Effectiveness*, 11(1): 56–82. 27

Heyes A., Cook N., and Rivers N. (2023) Clean Air and Cognitive Productivity: Effect and Adaptation. *Journal of the Association of Environmental and Resource Economists*. 3

Institute for Health Metrics and Evaluation (2019) Global Health Data Exchange - Global Burden of Disease Study 2019 (GBD 2019). 2

Italian Republic (2010) DECRETO LEGISLATIVO 13 agosto 2010, n. 155, "Attuazione della direttiva 2008/50/CE relativa alla qualita' dell'aria ambiente e per un'aria piu' pulita in Europa." (10G0177), in G. U. 15/09/2010, n. 217/L). 6

——— (2017) Decreto Legislativo 13 aprile 2017, n. 62: Norme in materia di valutazione e certificazione delle competenze nel primo ciclo ed esami di Stato. Art. 3: Ammissione alla classe successiva nella scuola primaria. 24

Ito K. and Zhang S. (2019) Willingness to Pay for Clean Air: Evidence from Air Purifier Markets in China. *Journal of Political Economy*: 000–000. 5

Klauber H., Holub F., Koch N., Pestel N., Ritter N., and Rohlf A. (2024) Killing Prescriptions Softly: Low Emission Zones and Child Health from Birth to School. *American Economic Journal: Economic Policy*, 16(2): 220–48. 7

Komisarow S. and Pakhtigian E. L. (2022) Are power plant closures a breath of fresh air? Local air quality and school absences. *Journal of Environmental Economics and Management*, 112: 102569. 4, 28

Kremer M., Romero M., and Saavedra S. (2025) Do Air Filters Improve Student Learning? Experimental Evidence from Colombia. Mimeo. 4, 27, 28

Künn S., Palacios J., and Pestel N. (2019) Indoor Air Quality and Cognitive Performance. IZA Discussion Papers 12632. 4

La Nauze A. and Severnini E. (2025) Air Pollution and Adult Cognition: Evidence from Brain Training. *Journal of the Association of Environmental and Resource Economists*, 12(1): 221–255. 3

Lai W., Li S., Li Y., and Tian X. (2021) Air Pollution and Cognitive Functions: Evidence from Straw Burning in China. *American Journal of Agricultural Economics*, n/a(n/a). 3

List J. A., Shaikh A. M., and Xu Y. (2019) Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4): 773–793. 15

Liu H. and Salvo A. (2018) Severe air pollution and child absences when schools and parents respond. *Journal of Environmental Economics and Management*, 92: 300–330. 4, 29

Liu J., Lee M., and Gershenson S. (2021) The short- and long-run impacts of secondary school absences. *Journal of Public Economics*, 199: 104441. 28, 29

Liu T., Yan W., and Peng K. (2024) A large-scale examination of the efficacy of air quality intervention in Chinese middle schools. AEA RCT Registry. 4

Lombardy Region (2006) LEGGE REGIONALE 11 dicembre 2006, n. 24, "Norme per la prevenzione e la riduzione delle emissioni in atmosfera a tutela della salute e dell'ambiente.". 6

——— (2013) DECRETO REGIONALE 11/09/2013, n. 593, "Approvazione del Piano Regionale degli interventi per la qualità dell'aria e dei relativi documenti previsti dalla procedura di valutazione ambientale strategica - (atto da trasmettere al consiglio regionale". 6

Mendoza D. L., Pirozzi C. S., Crosman E. T., Liou T. G., Zhang Y., Cleeves J. J., Bannister S. C., Anderegg W. R., and Robert P. I. (2020) Impact of low-level fine particulate matter and ozone exposure on absences in K-12 students and economic consequences. *Environmental Research Letters*, 15(11): 114052. 29

Metcalfe R. D. and Roth S. (2025) Making the Invisible Visible: The Impact of Revealing Indoor Air Pollution on Behavior and Welfare. Working Paper 33510, National Bureau of Economic Research. 5

Miguel E. and Kremer M. (2004) Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. *Econometrica*, 72(1): 159–217. 27

Miller N. H., Molitor D., and Zou E. (2024) The Nonlinear Effects of Air Pollution on Health: Evidence from Wildfire Smoke. Working Paper 32924, National Bureau of Economic Research. 2, 29

Ministero del Lavoro e delle Politiche Sociali (2025) Lavoro domestico - Tabelle dei minimi retributivi. 28

Ministero Italiano dell'Istruzione e del Merito (2023) Portale Unico dei Dati della Scuola. 7

Municipality of Milan (2022a) Atto n. DD 8014 del 30/09/2022 - Approvazione del "Programma di Monitoraggio del Piano Aria e Clima" del Comune di Milano, 30/09/2022. 7

——— (2022b) Atto n. DD 8519 del 12/10/2022 - Approvazione del "Programma di Monitoraggio del Piano Aria e Clima" del Comune di Milano, 12/10/2022. 7

——— (2023) Atto n. DD 6121 del 13/07/2023, "Approvazione del disciplinare dei sistemi di pagamento, di comunicazione e di registrazione della targa per l'accesso dei veicoli nella ZTL Cerchia dei Bastioni – Area C.". 7

Nasari M. M., Szyszkowicz M., Chen H., Crouse D., Turner M. C., Jerrett M., Pope C. A., Hubbell B., Fann N., Cohen A., Gapstur S. M., Diver W. R., Stieb D., Forouzanfar M. H., Kim S.-Y., Olives C., Krewski D., and Burnett R. T. (2016) A class of non-linear exposure-response models suitable for health impact assessment applicable to large cohort studies of ambient air pollution. *Air Quality, Atmosphere & Health*, 9(8): 961–972. 29

OECD (2025) Average class size. OECD Data Explorer, Accessed: 2025-10-31. 27

Palacios J., Eichholtz P., Kok N., and Duran N. (2022) Indoor Air Quality and Learning: Evidence from a Large Field Study in Primary Schools. SSRN Scholarly Paper No. 4296077. 3

Park R. J., Goodman J., Hurwitz M., and Smith J. (2020) Heat and learning. *American Economic Journal: Economic Policy*, 12(2): 306–39. 4, 5

Percoco M. (2013) Is road pricing effective in abating pollution? Evidence from Milan. *Transportation Research Part D: Transport and Environment*, 25: 112–118. 7

——— (2014) The effect of road pricing on traffic composition: Evidence from a natural experiment in Milan, Italy. *Transport Policy*, 31: 55–60. 7

Persico C. L. and Venator J. (2019) The Effects of Local Industrial Pollution on Students and Schools. *Journal of Human Resources*: 0518–9511R2. 4, 29

Pestel N. and Wozny F. (2021) Health effects of Low Emission Zones: Evidence from German hospitals. *Journal of Environmental Economics and Management*, 109: 102512. 7

Pope III C. A., Cropper M., Coggins J., and Cohen A. (2015) Health benefits of air pollution abatement policy: Role of the shape of the concentration–response function. *Journal of the Air & Waste Management Association*, 65(5): 516–522. 2, 29

Portt A., Engel L., Engel R., Bhatt D. K., Goadsby P. J., and Bhogal S. K. (2023) Migraine and air pollution: A systematic review. *Headache: The Journal of Head and Face Pain*, 63(10): 1313–1338. 24

Rahai R. and Evans G. W. (2023) Cumulative Neurotoxicological Air Pollution Exposure Is Associated with Lower Reading Improvement and Diminished Benefits of Literacy Interventions for Urban Elementary Students of Color. *Journal of Urban Health*, 100(3): 493–503. 3

Ransom M. R. and Pope C. A. (1992) Elementary school absences and PM10 pollution in Utah Valley. *Environmental Research*, 58(1): 204–219. 4

Robinson C. D., Lee M. G., Dearing E., and Rogers T. (2018) Reducing student absenteeism in the early grades by targeting parental beliefs. *American educational research journal*, 55(6): 1163–1192. 4, 27

Rogers T. and Feller A. (2018) Reducing student absences at scale by targeting parents' misbeliefs. *Nature Human Behaviour*, 2(5): 335–342. 4, 27

Roth S. (2021) The Effect of Indoor Air Pollution on Cognitive Performance: Evidence from the UK.Technical report. 3

Rubalcava L. and Teruel G. (2006) Mexican Family Life Survey, First Wave. 10

Ruggieri S., Longo V., Perrino C., Canepari S., Drago G., L'Abbate L., Balzan M., Cuttitta G., Scaccianoce G., Minardi R. et al. (2019) Indoor air quality in schools of a highly polluted south Mediterranean area. *Indoor Air*, 29(2): 276–290. 7

Ruiz-Tagle J. C., Sangwan N., Jaime M. M., Salazar C., Kacker K., and Kumar P. (2024) Technological and Behavioral Strategies to Mitigate the Effects of Air Pollution on Children in Delhi's Schools. Pre-Registered report accepted at the Journal of Development Economics. 5, 27

Schmit K. M., Brown R., Hayer S., Checovich M. M., Gern J. E., Wald E. R., and Barrett B. (2021) Wisconsin Upper Respiratory Symptom Survey for Kids: Validation of an Illness-specific Quality of Life Instrument. *Pediatric Research*, 90(6): 1207–1214. 10

Shaddick G., Thomas M. L., Mudu P., Ruggeri G., and Gumy S. (2020) Half the world's population are exposed to increasing air pollution. *NPJ Climate and Atmospheric Science*, 3(1): 23. 2

Sunyer J., Suades-González E., García-Esteban R., Rivas I., Pujol J., Alvarez-Pedrerol M., Forns J., Querol X., and Basagaña X. (2017) Traffic-related Air Pollution and Attention in Primary School Children. *Epidemiology*, 28(2): 181–189. 3

Tong Z., Li Y., Westerdahl D., and Freeman R. B. (2020) The impact of air filtration units on primary school students' indoor exposure to particulate matter in China. *Environmental Pollution (Barking, Essex: 1987)*, 266(Pt 2): 115107. 4

Weichenthal S., Pinault L., Christidis T., Burnett R. T., Brook J. R., Chu Y., Crouse D. L., Erickson A. C., Hystad P., Li C. et al. (2022) How low can you go? Air pollution affects mortality at very low levels. *Science Advances*, 8(39): eabo3381. 2, 29

World Bank (2023) Primary Completion Rate, Total (% of Relevant Age Group) – Italy. World Development Indicators. 24

World Health Organization (2021) *WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*, Genève, Switzerland: World Health Organization. 2

Xia X., Chan K. H., Lam K. B. H., Qiu H., Li Z., Yim S. H. L., and Ho K.-F. (2021) Effectiveness of indoor air purification intervention in improving cardiovascular health: A systematic review and meta-analysis of randomized controlled trials. *The Science of the total environment*, 789: 147882. 4

Yang X., Wang Q., Han F., Dong B., Wen B., Li L., Ruan H., Zhang S., Kong J., Zhi H., Wang C., Wang J., Zhang M., and Xu D. (2021) Pulmonary Benefits of Intervention with HEPA Air Purifier in Schoolchildren: A Double-Blind Crossover Study. February. 4

Yao Z., Zhang W., Ji X., and Weng W. (2023) Short-Term Exposure to Air Pollution and Cognitive Performance: New Evidence from China's College English Test. *Environmental and Resource Economics*, 85(1): 211–237. 3

Young A. (2019) Channeling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results*. *The Quarterly Journal of Economics*, 134(2): 557–598, _eprint: https://academic.oup.com/qje/article-pdf/134/2/557/28289325/qjy029.pdf. 15

Zhang J. and Mu Q. (2018) Air pollution and defensive expenditures: Evidence from particulate-filtering facemasks. *Journal of Environmental Economics and Management*, 92: 517–536. 5

Online Appendix

The Effect of Air Purifiers in Schools

Jacopo Bonan, Francesco Granella, Stefania Renna, Luis Sarmiento, Massimo Tavoni

A Additional tables and figures

Table A.1: Survey attrition and missing values at the baseline

	(1) Obs	(2) Control	(3)	(4) Difference	(5)
Variable		Mean	(SD)	Diff.	(SE)
Surveyed in wave 1	2,051	0.888	0.315	−0.006	0.057
Missing values conditional on participation					
Cognitive and behavioral					
Raven score (standardized)	1,822	0.069	0.253	0.006	0.015
Mood scale	1,822	0.191	0.393	−0.034	0.026
Aggressive episodes	1,822	0.108	0.311	−0.007	0.020
Health symptoms					
Runny nose	1,822	0.223	0.417	−0.040	0.038
Blocked nose	1,822	0.228	0.420	−0.035	0.041
Sneezing	1,822	0.225	0.418	−0.019	0.041
Cough	1,822	0.232	0.422	−0.039	0.040
Short of breath	1,822	0.265	0.441	−0.039	0.043
Tiredness	1,822	0.257	0.437	−0.049	0.044
Headache	1,822	0.244	0.430	−0.030	0.043
Stomach ache	1,822	0.259	0.438	−0.043	0.043
Air quality perceptions					
Overall	1,822	0.149	0.356	−0.007	0.026
City	1,822	0.145	0.352	0.006	0.025
Classroom	1,822	0.150	0.357	−0.004	0.027
Schoolyard	1,822	0.153	0.360	−0.008	0.027
Urban policy priorities					
City cleaning	1,822	0.153	0.360	−0.020	0.027
Green areas and playgrounds	1,822	0.161	0.368	−0.016	0.026
Sport infrastructure	1,822	0.169	0.375	−0.016	0.026
Air quality	1,822	0.159	0.366	−0.007	0.026
Less traffic	1,822	0.177	0.382	−0.020	0.030

Notes: The table presents the probability of participating in the baseline survey (first row) and the probability of non-response for each question conditional on participation. Columns 2–3 report the mean and standard deviation in the control group. Columns 4–5 report the difference between treatment and control with school fixed effects; standard errors are clustered at the classroom level. Significance levels: *** $p<0.01$, ** $p<0.05$, and * $p<0.1$.

Table A.2: Survey attrition and missing values at the endline

Variable	(1) Obs	(2) Control Mean	(3) (SD)	(4) Difference Diff.	(5) (SE)
Surveyed in wave 2	2,051	0.876	0.330	0.024	0.034
Missing values conditional on participation					
Academic performance					
Italian test score	858	0.202	0.402	−0.043	0.043
Math test score	858	0.193	0.395	−0.043	0.044
Cognitive and behavioral					
Raven score (standardized)	1,815	0.003	0.055	0.002	0.003
Mood scale	1,815	0.090	0.286	0.001	0.019
Aggressive episodes	1,815	0.025	0.156	0.000	0.007
Health symptoms					
Runny nose	1,815	0.132	0.338	−0.018	0.034
Blocked nose	1,815	0.128	0.334	−0.020	0.035
Sneezing	1,815	0.139	0.346	−0.032	0.034
Cough	1,815	0.124	0.329	−0.042	0.034
Short of breath	1,815	0.155	0.362	−0.028	0.036
Tiredness	1,815	0.152	0.359	−0.041	0.036
Headache	1,815	0.155	0.362	−0.048	0.037
Stomach ache	1,815	0.151	0.358	−0.043	0.036
Air quality perceptions					
Overall	1,815	0.052	0.222	0.008	0.016
City	1,815	0.040	0.196	0.003	0.013
Classroom	1,815	0.042	0.200	0.009	0.015
Schoolyard	1,815	0.049	0.216	0.002	0.015
Urban policy priorities					
City cleaning	1,815	0.048	0.214	−0.002	0.012
Green areas and playgrounds	1,815	0.049	0.216	0.007	0.015
Sport infrastructure	1,815	0.065	0.246	−0.009	0.015
Air quality	1,815	0.056	0.230	−0.004	0.014
Less traffic	1,815	0.063	0.243	0.007	0.018

Notes: The table presents the probability of participating in the endline survey (first row) and the probability of non-response for each question conditional on participation. Columns 2–3 report the mean and standard deviation in the control group. Columns 4–5 report the difference between treatment and control with school fixed effects; standard errors are clustered at the classroom level. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1.

Table A.3: Survey attrition and missing values at both waves

	(1) Obs	(2) Control	(3)	(4) Difference	(5)
Variable		Mean	(SD)	Diff.	(SE)
Surveyed in both waves	2,051	0.798	0.402	0.023	0.057
Missing values conditional on participation					
Cognitive and behavioral					
Raven score (standardized)	1,662	0.068	0.252	−0.002	0.016
Mood scale	1,662	0.250	0.433	−0.036	0.035
Aggressive episodes	1,662	0.126	0.332	−0.015	0.022
Health symptoms					
Runny nose	1,662	0.312	0.464	−0.052	0.043
Blocked nose	1,662	0.321	0.467	−0.064	0.047
Sneezing	1,662	0.326	0.469	−0.063	0.046
Cough	1,662	0.320	0.467	−0.083*	0.045
Short of breath	1,662	0.364	0.481	−0.064	0.047
Tiredness	1,662	0.356	0.479	−0.086*	0.049
Headache	1,662	0.344	0.475	−0.071	0.049
Stomach ache	1,662	0.352	0.478	−0.075	0.048
Air quality perceptions					
Overall	1,662	0.183	0.387	−0.011	0.031
City	1,662	0.169	0.375	−0.001	0.028
Classroom	1,662	0.176	0.381	−0.012	0.031
Schoolyard	1,662	0.176	0.381	−0.013	0.030
Urban policy priorities					
City cleaning	1,662	0.181	0.385	−0.026	0.029
Green areas and playgrounds	1,662	0.193	0.395	−0.016	0.029
Sport infrastructure	1,662	0.210	0.408	−0.028	0.030
Air quality	1,662	0.197	0.398	−0.017	0.030
Less traffic	1,662	0.212	0.409	−0.019	0.034

Notes: The table presents the probability of participating in both survey waves (first row) and the probability of non-response for each question in either wave, conditional on participation. Columns 2–3 report the mean and standard deviation in the control group. Columns 4–5 report the difference between treatment and control with school fixed effects; standard errors are clustered at the classroom level. Significance levels: *** p<0.01, ** p<0.05, and * p<0.1.

Table A.4: School demographics and student performance indicators

Variable	(1) Sample	(2) Province	(3) Region	(4) Country
Academic performance (INVALSI)				
Italian test (grade 2)	187.4	194.0	194.6	194.7
	(55.8)	(47.8)	(47.6)	(46.0)
Math test (grade 2)	186.0	195.7	195.7	193.8
	(41.0)	(42.9)	(42.4)	(44.6)
Italian test (grade 5)	190.7	198.1	197.9	194.9
	(41.6)	(38.0)	(37.6)	(39.8)
Math test (grade 5)	187.7	196.6	196.1	193.9
	(35.3)	(37.1)	(37.1)	(39.1)
Demographics				
Students per class	21.6	19.8	18.7	17.0
	(2.5)	(1.9)	(3.2)	(3.8)
Share female	0.47	0.48	0.48	0.48
	(0.50)	(0.04)	(0.05)	(0.06)
Share non-Italian	0.38	0.26	0.21	0.14
	(0.49)	(0.18)	(0.15)	(0.13)

Notes: Mean and standard deviation (in parentheses) of school-level demographics and student performance indicators for the 2023–24 school year. Column 1 reports sample averages across the five study schools. Columns 2–4 report averages for the province of Milan, the Lombardy region, and Italy. Student performance is measured using INVALSI standardized test scores in Italian and mathematics for grades 2 and 5. Demographic indicators include average class size and school-level shares of female and non-Italian students. Standard deviations in Column 1 are computed across students (demographics) or test-takers (INVALSI); in Columns 2–4 they are computed across schools.

Table A.5: Average treatment effects on indoor air pollution, controlling for temperature and humidity

	(1) $PM_{2.5}$	(2) PM_{10}	(3) CO
Estimate	-4.550^{***}	-4.706^{***}	-0.201
	(0.337)	(0.360)	(0.204)
Temperature	-1.084^{***}	-1.017^{***}	0.003
	(0.193)	(0.200)	(0.085)
Rel. Humidity	-0.188	-0.126	0.006
	(0.111)	(0.118)	(0.043)
N.Obs	3,456	3,456	3,456
Control Mean	14.15	14.85	1.31

Notes: This table replicates the pollution results from Panel A of Table 2 (Columns 1–3), adding indoor air temperature and relative humidity as controls. All models include calendar date, strata-by-weekday, and strata-by-month fixed effects (where strata are school × grade, the randomization stratification, consistent with Panel A of Table 2). The sample is restricted to school days, matching Panel A of Table 2. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.6: Robustness of Panel A treatment effects to alternative inference methods

	(1) $PM_{2.5}$	(2) PM_{10}	(3) CO	(4) CO_2	(5) Temp.
Estimate	−4.615	−4.715	−0.195	31.017	−0.120
Clustered SE	(0.379)***	(0.403)***	(0.217)	(54.376)	(0.132)
Wild cluster bootstrap SE	(0.959)***	(0.976)***	(0.207)	(51.635)	(0.126)
Randomization SE	(1.417)***	(1.464)***	(0.321)	(76.579)	(0.198)
N. Obs	3,456	3,456	3,456	3,456	3,456
N. Clusters	30	30	30	30	30

Notes: This table reports point estimates and standard errors from three inference approaches for the average treatment effects on indoor environmental variables in Panel A of Table 2: conventional standard errors clustered at the classroom level; the wild cluster restricted bootstrap of Cameron et al. (2008) with Webb 6-point weights and $B = 9{,}999$ replications; and randomization inference via cluster-level permutation of treatment status across the 30 monitored classrooms ($B = 9{,}999$ draws). Significance stars correspond to the inference method in each row. All specifications include calendar date, strata-by-weekday, and strata-by-month fixed effects (strata = school × grade), and are estimated on school days. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.7: Treatment effects on indoor environmental variables: robustness to sample restrictions

	(1) $PM_{2.5}$	(2) PM_{10}	(3) CO	(4) CO_2	(5) Temp.
Panel A: School hours vs. non-school hours (school days only)					
School hours (8am–5pm)					
Estimate	−2.530***	−2.630***	−0.194	71.138	−0.188
	(0.327)	(0.375)	(0.221)	(73.945)	(0.216)
N.Obs	3,451	3,451	3,451	3,451	3,451
Control Mean	14.15	15.35	1.35	1159.2	21.80
Rel. Change %	−17.9	−17.1	−14.4	6.1	−0.9
Non-school hours (5pm–8am)					
Estimate	−5.941***	−6.043***	−0.222	−3.745	−0.077
	(0.447)	(0.467)	(0.212)	(43.035)	(0.100)
N.Obs	3,446	3,446	3,446	3,446	3,446
Control Mean	14.17	14.56	1.28	592.0	20.47
Rel. Change %	−41.9	−41.5	−17.3	−0.6	−0.4
Panel B: School days vs. non-school days (24-hour averages)					
School days					
Estimate	−4.615***	−4.715***	−0.195	31.017	−0.120
	(0.379)	(0.403)	(0.217)	(54.376)	(0.132)
N.Obs	3,456	3,456	3,456	3,456	3,456
Control Mean	14.15	14.85	1.31	805.9	20.97
Rel. Change %	−32.6	−31.8	−15.0	3.8	−0.6
Non-school days (weekends and holidays)					
Estimate	−3.536***	−3.533***	−0.236	−40.971	−0.259
	(0.421)	(0.429)	(0.216)	(40.621)	(0.160)
N.Obs	2,062	2,062	2,062	2,062	2,062
Control Mean	10.18	10.26	1.27	494.4	18.93
Rel. Change %	−34.7	−34.4	−18.6	−8.3	−1.4

Notes: Panel A compares treatment effects during school hours (8am–5pm) and non-school hours (5pm–8am) on school days, using hourly sensor data aggregated to classroom-day-period averages. Panel B compares treatment effects on school days (24-hour averages, matching the main specification in Table 2) and non-school days (weekends and holidays). All models include calendar date, strata-by-weekday, and strata-by-month fixed effects (strata = school × grade, consistent with Panel A of Table 2). Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.8: Treatment effect on absences: robustness to alternative estimators

	(1) LPM	(2) Probit	(3) Logit	(4) Poisson
Estimate	−0.0103**	−0.1030**	−0.2167**	−0.1920**
	(0.0049)	(0.0485)	(0.1027)	(0.0931)
p-value	0.037	0.034	0.035	0.039
Marginal effect	−0.0103	−0.0112	−0.0116	
Effect (%)	−17.0	−18.6	−19.3	−17.5
N.Obs	620,145	615,127	615,127	29,080
N.Clusters	95	95	95	95
Control Mean	0.060	0.060	0.060	1.29

Notes: All models pool school years 2022–23 and 2023–24 with the main specification: student (Columns 1–3) or classroom (Column 4) fixed effects, calendar date, strata-by-month, and strata-by-year fixed effects (where strata are school × grade), and strata-by-treatment-group linear time trends. Columns 1–3 are estimated at the student-day level with a binary absence indicator. Column 4 is a Poisson count model at the classroom-day level with the log number of students as an offset. Marginal effects for Probit and Logit are computed as the average of individual marginal effects. The effect (%) row expresses the estimate relative to the control mean (Columns 1–3) or as $(\exp(\hat{\beta}) - 1) \times 100$ (Column 4). Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.9: Treatment effects on primary outcomes: robustness to alternative inference methods

	(1) $PM_{2.5}$	(2) Daily Absence (LPM)	(3) Italian Score	(4) Math Score	(5) Stz Raven Score	(6) Mood Scale	(7) Aggressive Episodes
Estimate	−4.615	−0.0103	−0.826	0.560	−0.0277	0.0478	0.0293
Clustered SE	(0.379)***	(0.0049)**	(6.264)	(4.479)	(0.0628)	(0.0451)	(0.0345)
Wild cluster bootstrap SE	(0.959)***	(0.0050)*	(6.225)	(4.402)	(0.0625)	(0.0447)	(0.0348)
Randomization SE	(1.417)***	(0.0046)**	(6.103)	(4.761)	(0.0571)	(0.0476)	(0.0328)
N. Obs	3,456	620,145	701	711	1,808	1,651	1,770
N. Clusters	30	95	37	37	92	92	92

Notes: This table reports point estimates and standard errors from three inference approaches for the average treatment effects in Table 3: conventional standard errors clustered at the classroom level; the wild cluster restricted bootstrap of Cameron et al. (2008) with Webb 6-point weights; and randomization inference via permutation of treatment status within randomization strata (school × grade). The bootstrap and randomization-inference use $B = 9{,}999$ replications for the cross-sectional outcomes in columns 1 and 3–7, and $B = 999$ replications for the linear probability model of daily absences in column 2, where the per-replication cost (large panel, strata-by-period fixed effects, and strata-by-treatment-group time trends) precludes a larger number of draws. Significance stars correspond to the inference method in each row. Specifications match those of Table 3. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.10: Heterogeneous treatment effects on absences by pre-treatment absences: robustness to different operationalizations of the moderator

	(1)	(2)	(3)	(4)	(5)
	Pre-treatment absence quartiles (split-sample)				*Continuous*
	Q1	Q2	Q3	Q4	
Quartile regressions (Cols. 1–4):					
Total TE (Treatment)	−0.0017	−0.0013	−0.0066	−0.0280**	
	(0.0049)	(0.0073)	(0.0070)	(0.0126)	
Continuous regression (Col. 5):					
Treatment (base, PreAbs= 0)					0.0015
					(0.0052)
Tr × PreAbs					−0.1872***
					(0.0584)
Post × PreAbs					−0.4019***
					(0.0336)
N.Obs	143,059	157,883	158,725	160,478	620,145
N.Clusters	93	94	94	91	95
Pre-abs rate (group mean)	0.008	0.033	0.060	0.132	—

Notes: The dependent variable is a binary indicator for daily absence, estimated with a Linear Probability Model pooling school years 2022–23 and 2023–24. Cols. 1–4 report separate split-sample treatment effect estimates for each pre-treatment absence quartile, using the fully interacted specification of Column 1 of Table 4: calendar date, strata-by-month, and strata-by-year fixed effects and strata-by-treatment-group linear time trends are all estimated separately within each quartile. Quartile membership is defined at the student level (time-invariant) using the same seed as Table 4. Col. 5 uses a pooled regression with the continuous pre-treatment absence rate interacted with treatment; a lower-order term Post × PreAbs is included because the continuous moderator has within-date student-level variation and is therefore identified (see footnote to Equation 3). Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.11: Heterogeneous treatment effects on absences by student characteristics

	(1)	(2)	(3)	(4)
	Gender		*Citizenship*	
	Male	Tr×Female	Italian	Tr×Foreign
Estimate	-0.0102^{**}	-0.0000	-0.0063	-0.0100^{*}
	(0.0050)	(0.0044)	(0.0050)	(0.0053)
N.Obs	620,145	620,145	620,145	620,145
N.Clusters	95	95	95	95
Control Mean ($M = 0$)	0.060		0.056	
Control Mean ($M = 1$)		0.058		0.063
Total TE ($M = 0$)	-0.0102^{**}		-0.0063	
	(0.0050)		(0.0050)	
Total TE ($M = 1$)	-0.0103^{*}		-0.0163^{***}	
	(0.0057)		(0.0062)	

Notes: The dependent variable is a binary indicator for daily absence. All models are estimated with a Linear Probability Model pooling school years 2022–23 and 2023–24, with student, calendar date, strata-by-month, and strata-by-year fixed effects (where strata are school × grade), and strata-by-treatment-group linear time trends. Columns 1–2 interact the treatment with an indicator for female students (male is the reference). Columns 3–4 interact the treatment with an indicator for foreign citizenship (Italian is the reference). Stars on interaction terms test whether the treatment effect differs across subgroups, not whether it differs from zero; the "Implied effect" rows report the total treatment effect for each subgroup. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.12: Heterogeneous treatment effects on absences by grade

	(1) Grade 1	(2) Tr×G2	(3) Tr×G3	(4) Tr×G4	(5) Tr×G5
Treatment	-0.0136^{**}				
	(0.0061)				
Tr × Grade		-0.0134	0.0089	0.0175	0.0035
		(0.0099)	(0.0107)	(0.0118)	(0.0109)
N.Obs	620,145	620,145	620,145	620,145	620,145
N.Clusters	95	95	95	95	95
Control Mean	0.058	0.070	0.053	0.058	0.056
Total TE	-0.0136^{**}	-0.0270^{***}	-0.0047	0.0039	-0.0101
	(0.0061)	(0.0083)	(0.0091)	(0.0104)	(0.0096)

Notes: The dependent variable is a binary indicator for daily absence. The model is estimated with a Linear Probability Model pooling school years 2022–23 and 2023–24, with student, calendar date, strata-by-month, and strata-by-year fixed effects (where strata are school × grade), and strata-by-treatment-group linear time trends. The specification interacts the treatment indicator with grade dummies, using grade 1 as the reference category; Tr×Gk reports the differential treatment effect for grade k relative to grade 1. The "Total TE" rows report the total treatment effect for each grade computed via the delta method, with corresponding standard errors. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.13: Heterogeneous treatment effects on absences by outdoor air pollution (quartile splits)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	7-day rolling average of $PM_{2.5}$				*WHO exceedances in last 7 days*			
	Q1	Tr×Q2	Tr×Q3	Tr×Q4	Q1	Tr×Q2	Tr×Q3	Tr×Q4
Estimate	−0.0182***	0.0050	0.0037	0.0126***	−0.0180***	0.0051	0.0050	0.0117**
	(0.0056)	(0.0038)	(0.0041)	(0.0044)	(0.0059)	(0.0039)	(0.0040)	(0.0048)
N.Obs	605,985	605,985	605,985	605,985	605,985	605,985	605,985	605,985
N.Clusters	95	95	95	95	95	95	95	95
Control Mean	0.060	0.062	0.061	0.053	0.061	0.060	0.062	0.054
Total TE	−0.0182***	−0.0132**	−0.0146***	−0.0056	−0.0180***	−0.0129**	−0.0130**	−0.0063
	(0.0056)	(0.0057)	(0.0056)	(0.0051)	(0.0059)	(0.0056)	(0.0053)	(0.0056)

Notes: The dependent variable is a binary indicator for daily absence. All models are estimated with a Linear Probability Model pooling school years 2022–23 and 2023–24, with student, calendar date, strata-by-month, and strata-by-year fixed effects (where strata are school × grade), and strata-by-treatment-group linear time trends. Columns 1–4 interact the treatment indicator with quartiles of the 7-day lagged rolling average of outdoor $PM_{2.5}$. Columns 5–8 interact the treatment with quartiles of the 7-day rolling sum of days with outdoor $PM_{2.5}$ exceeding the WHO daily threshold (15 $\mu g/m^3$). Q1 (lowest pollution) is the reference category. Both specifications include quartile-by-post interactions to allow for differential post-treatment shifts by pollution level. The bottom row reports the total treatment effect for each quartile (Q1 coefficient plus the interaction), with delta-method standard errors. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.14: Heterogeneous treatment effects on absences by outdoor pollen concentration

	(1) Q1	(2) Tr×Q2	(3) Tr×Q3	(4) Tr×Q4
Estimate	-0.0091^{*}	-0.0023	-0.0047	-0.0065
	(0.0051)	(0.0041)	(0.0044)	(0.0043)
N.Obs	620,145	620,145	620,145	620,145
N.Clusters	95	95	95	95
Control Mean	0.063	0.053	0.057	0.060
Total TE	-0.0091^{*}	-0.0113^{**}	-0.0138^{**}	-0.0156^{***}
	(0.0051)	(0.0052)	(0.0060)	(0.0058)

Notes: The dependent variable is a binary indicator for daily absence. The model interacts the treatment indicator with quartiles of the 10-day lagged rolling average of total outdoor pollen concentration. Q1 (lowest pollen) is the reference category. The specification includes quartile-by-post interactions to allow for differential post-treatment shifts by pollen level. Pollens includes: Alternaria, Alnus, Betula, Cladosporium, Ambrosia, Artemisia, Carpinus betulus, Corylus avellana, Cupressaceae and Taxaceae, Fagaceae, Gramineae, Oleaceae, Urticaceae. Data from the Association of Italian Territorial and Hospital Allergists and Immunologists (`https://www.pollinieallergia.net/`). All models are estimated with a Linear Probability Model pooling school years 2022–23 and 2023–24, with student, calendar date, strata-by-month, and strata-by-year fixed effects (where strata are school × grade), and strata-by-treatment-group linear time trends. The bottom row reports the total treatment effect for each quartile. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.15: Impact on cognitive and behavioral outcomes, robustness check with ANOVA

	(1)	(2)	(3)
	Stdz Raven score	Individual mood scale	Aggressive episodes
Estimate	0.034	0.055	0.057*
	(0.056)	(0.045)	(0.032)
N.Obs	1,552	1,274	1,464
N.Clusters	86	86	86
Control Mean	0.000	3.424	0.395
Control SD	0.998	0.772	0.489
Ex-post MDE	0.156	0.125	0.090

Notes: OLS estimates of the treatment effect on cognitive and behavioral outcomes using ANOVA (controlling for baseline outcome). Models control for gender and foreign nationality and include grade and school fixed effects. FDR-adjusted q-values (Benjamini et al., 2006) are reported in brackets, corrected within the family of primary outcomes. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.16: Impact on self-reported health symptoms, perceptions, and preferences, robustness check with ANOVA

Panel A: Impact on Self-Reported Health

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Symptoms							
	Runny nose	Blocked nose	Sneezing	Cough	Short breath	Tiredness	Headache	Stomach ache
Estimate	-0.068**	-0.023	-0.012	-0.005	0.023	0.002	0.029	0.060**
	(0.032)	(0.025)	(0.024)	(0.031)	(0.036)	(0.031)	(0.028)	(0.029)
N.Obs	1,184	1,178	1,168	1,193	1,108	1,136	1,145	1,135
N.Clusters	83	83	83	83	83	83	83	83
Control Mean	0.532	0.569	0.664	0.576	0.259	0.646	0.452	0.403
Control SD	0.499	0.496	0.473	0.494	0.438	0.479	0.498	0.491
Ex-post MDE	0.091	0.069	0.068	0.088	0.101	0.088	0.080	0.080

Panel B: Impact on Perceptions and Preferences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Air Quality Perception				Urban Policy Priorities				
	General	City	Class	Schoolyard	City cleaning	Green playgrounds	Sport infrastructure	Air quality	Less traffic
Estimate	0.115	-0.005	0.149**	-0.047	0.069	0.086	0.005	0.176***	-0.007
	(0.078)	(0.081)	(0.071)	(0.071)	(0.045)	(0.062)	(0.069)	(0.058)	(0.059)
N.Obs	1,367	1,383	1,379	1,379	1,379	1,351	1,331	1,345	1,320
N.Clusters	85	85	85	85	86	86	85	85	85
Control Mean	2.978	2.559	3.109	3.360	3.509	3.280	3.088	3.456	3.158
Control SD	0.880	0.983	0.793	0.808	0.871	0.980	1.074	0.977	1.014
Ex-post MDE	0.219	0.226	0.197	0.198	0.126	0.175	0.194	0.162	0.165

Notes: Panel A reports OLS estimates of the treatment effect on self-reported health symptoms, air quality perceptions, and urban policy priorities using ANOVA (controlling for baseline outcomes). The dependent variables equal one if the student reported the symptom at least some time over the previous week and zero otherwise. Panel B reports estimates for air quality perceptions, using an air quality perception index scoring from 1 (very bad) to 4 (very good), and urban policy priorities (last five columns), using an index scoring from 1 (not important at all) to 4 (very important). All models control for gender and foreign nationality and include grade and school fixed effects. Standard errors are clustered at the classroom level. FDR-adjusted q-values are omitted here but are available upon request. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table A.17: Impact on self-reported health symptoms, robustness check with ANOVA

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Symptoms							
	Runny nose	Blocked nose	Sneezing	Cough	Short breath	Tiredness	Headache	Stomach ache
Estimate	-0.068**	-0.023	-0.012	-0.005	0.023	0.002	0.029	0.060**
	(0.032)	(0.025)	(0.024)	(0.031)	(0.036)	(0.031)	(0.028)	(0.029)
N.Obs	1,184	1,178	1,168	1,193	1,108	1,136	1,145	1,135
N.Clusters	83	83	83	83	83	83	83	83
Control Mean	0.532	0.569	0.664	0.576	0.259	0.646	0.452	0.403
Control SD	0.499	0.496	0.473	0.494	0.438	0.479	0.498	0.491
Ex-post MDE	0.091	0.069	0.068	0.088	0.101	0.088	0.080	0.080

Notes: OLS estimates of the treatment effect on self-reported health symptoms using ANOVA (controlling for baseline outcome). The dependent variables equal one if the student reported the symptom at least some time over the previous week and zero otherwise. Models control for gender and foreign nationality and include grade and school fixed effects. FDR-adjusted q-values (Benjamini et al., 2006) are reported in brackets, corrected within the family of health symptoms. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.18: Impact on air quality perceptions, robustness check with ANOVA

	(1) General	(2) City	(3) Class	(4) Schoolyard
Estimate	0.115	-0.005	0.149**	-0.047
	(0.078)	(0.081)	(0.071)	(0.071)
N.Obs	1,367	1,383	1,379	1,379
N.Clusters	85	85	85	85
Control Mean	2.978	2.559	3.109	3.360
Control SD	0.880	0.983	0.793	0.808
Ex-post MDE	0.219	0.226	0.197	0.198

Notes: OLS estimates of the treatment effect on air quality perceptions using ANOVA (controlling for baseline outcome). The air quality perception index is scored from 1 (very bad) to 4 (very good). Models control for gender and foreign nationality and include grade and school fixed effects. FDR-adjusted q-values (Benjamini et al., 2006) are reported in brackets, corrected within the family of air quality perceptions. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.19: Impact on urban policy priorities, robustness check with ANOVA

	(1) City cleaning	(2) Green playgrounds	(3) Sport infrastructure	(4) Air quality	(5) Less traffic
Estimate	0.069	0.086	0.005	0.176***	-0.007
	(0.045)	(0.062)	(0.069)	(0.058)	(0.059)
N.Obs	1,379	1,351	1,331	1,345	1,320
N.Clusters	86	86	85	85	85
Control Mean	3.509	3.280	3.088	3.456	3.158
Control SD	0.871	0.980	1.074	0.977	1.014
Ex-post MDE	0.126	0.175	0.194	0.162	0.165

Notes: OLS estimates of the treatment effect on urban policy priorities using ANOVA (controlling for baseline outcome). The policy priority score is scored from 1 (not important at all) to 4 (very important). Models control for gender and foreign nationality and include grade and school fixed effects. FDR-adjusted q-values (Benjamini et al., 2006) are reported in brackets, corrected within the family of urban policy priorities. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table A.20: Association between baseline symptoms and pre-treatment absences

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline symptom reported							
	Runny nose	Blocked nose	Sneezing	Cough	Short breath	Tiredness	Headache	Stomach ache
Pre-treat absences	0.429*	0.308	0.220	0.162	0.341	0.194	0.241	−0.066
	(0.239)	(0.241)	(0.232)	(0.239)	(0.214)	(0.239)	(0.245)	(0.241)
N.Obs	1,451	1,437	1,430	1,433	1,375	1,397	1,405	1,389

Notes: OLS estimates of the association between pre-treatment absence rates and baseline self-reported health symptoms. The dependent variables equal one if the student reported the symptom at least some time over the previous week and zero otherwise. Pre-treatment absences are calculated over the school year 2022–23 and September–October 2023. Models control for gender and foreign citizenship. Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

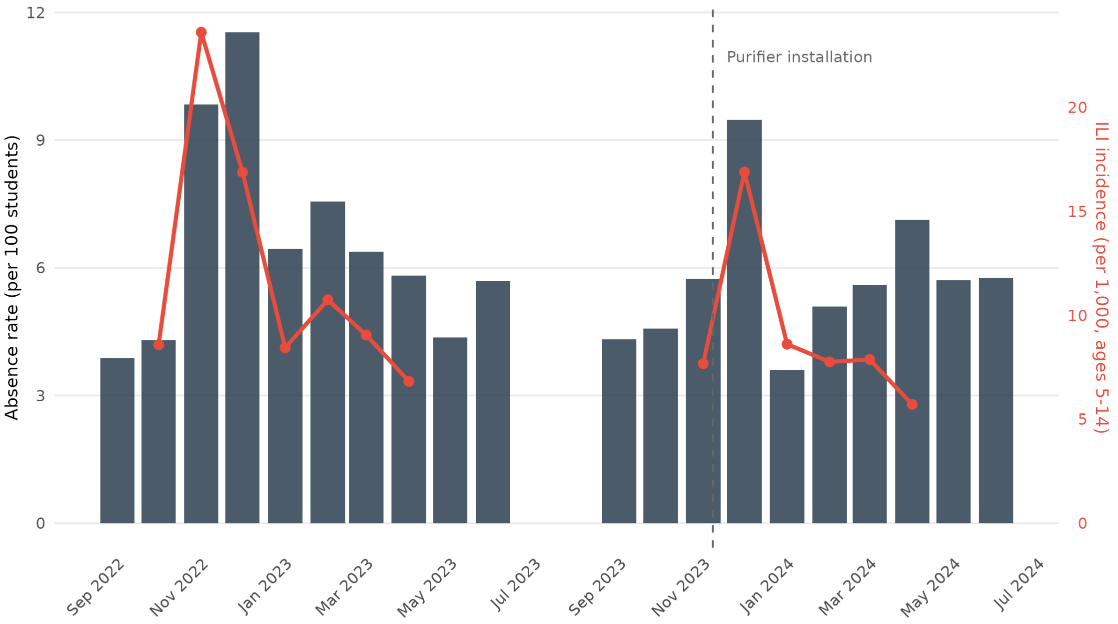


Figure A.1: School absences and influenza-like illness incidence in Lombardy

Notes: Navy bars show the monthly absence rate in the control group (left axis, absences per 100 students). The red line shows the monthly average influenza-like illness (ILI) incidence among children aged 5–14 in Lombardy (right axis, cases per 1,000). ILI data from the Italian National Institute of Health (ISS) RespiVirNet surveillance system (Branda, 2024). The dashed green line marks the purifier installation date (November 8, 2023). Absence peaks in winter months coincide with the flu season, consistent with the weaker treatment effects during winter documented in Table 4 Column 3.

B Purifiers and monitors' technical features

The study utilizes NETCO NIVEUS NV100 purifiers, shown in the left panel of Appendix Figure B.1. These purifiers feature U15 Ultra Low Particulate Air (ULPA) filters, which capture up to 99.99% of particles larger than 0.026 microns. ULPA filtration is the highest standard of mechanical air purification, certified and recognized internationally. It surpasses the more common HEPA filters, providing 10 to 100 times greater efficiency (see Appendix Table B.1).

The purifiers are energy-efficient, consuming only 4W per hour at the operating speed used in the study, comparable to a 60-watt incandescent bulb. They also operate quietly, with average acoustic pressure levels ranging from 29 to 45 dB(A). Air enters the device and passes through the ULPA filter, made of layers of ultrafine material, followed by an activated carbon filter, before being recirculated into the environment. The efficiency of the purifiers is measured primarily by the Clean Air Delivery Rate (CADR), expressed in cubic meters per hour (m^3/h), and the Air Exchange Rate, which indicates how many times per hour the purifier can filter all the air in a given room. Following the manufacturer's recommendations, we selected a model suitable for the average classroom volume. The installed units have a CADR of 200 m^3/h, yielding an average Air Exchange Rate of 1.04 across classrooms (ranging from 0.7 to 1.5). To balance effectiveness and noise reduction, purifiers operate at 60% capacity (speed 3 of 5), producing an acoustic pressure of 33.5 dB(A)—below the WHO recommended limit of 35 dB(A) for classrooms. The product has a 3-year warranty and has an expected lifespan of ten years.[32]

In addition to the purifiers, we installed 31 ENVIRA Nanoenvi indoor air quality sensors, as shown in the right panel of Appendix Figure B.1. These sensors measure concentrations of CO_2 (ppm), $PM_{2.5}$, PM_{10}, and CO (ppm), along with temperature ($°C$), humidity (%), and atmospheric pressure (hPa). Their technical specifications are detailed in Appendix Table B.2. Once powered and connected to the internet, the sensors transmit measurements every 30 seconds to an online data platform. Each device features a small LED display that visually represents indoor air quality using a four-level color-coded scale based on the Indoor Ambient Air Quality Index defined by the manufacturer. To ensure comparability between classrooms with and without sensors—and to minimize the risk of influencing behavior—we covered the LED displays with anti-tampering tape so that students and teachers could not see real-time air quality readings. When the sensors were collected at the end of the study period, the tape was still intact on all devices.

---

[32] The main mechanical component, the ventilator, has a certified lifespan of 40,000 hours at full speed.

Figure B.1: Purifiers and sensors installed

Table B.1: Efficiency of different mechanical filter technologies

Filter Group	Class	MPSS INTEGRAL VALUES		MPSS INTEGRAL VALUES	
		Efficiency (%)	Penetration (%)	Efficiency (%)	Penetration (%)
EPA	E10	85	15	-	-
	E11	95	5	-	-
	E12	99.5	0.5	-	-
HEPA	H13	99.95	0.05	99.75	0.25
	H14	99.995	0.005	99.975	0.025
ULPA	U15	99.9995	0.0005	99.9975	0.0025
	U16	99.99995	5E-05	99.99975	0.00025
	U17	99.99995	5E-05	99.9999	0.0001

Table B.2: Technical specifications of the low-cost sensors employed

Pollutant/Parameter	**Precision**	**Measuring range**
Carbon monoxide (CO)	±5%	0 - 5000 ppm
Particulate matter ($PM_{2.5}$)	±10 μg/m^3	0 - 1000 μg/m^3
Carbon dioxide (CO_2)	±30 ppm	0 - 40000 ppm
Temperature	±0.02 ºC	0 - 65 ºC
Relative humidity	±2%	10 - 95%
Atmospheric pressure	±10 hPa	500 - 1150 hPa

Notes: This table presents the precision and measuring range of low-cost sensors by measured pollutant (CO, CO_2, $PM_{2.5}$) and environmental parameter (temperature, relative humidity, atmospheric pressure). *Source:* (ENVIRA, 2024a,b).

C Indoor sensors' intercomparison

We conducted a sensor-to-sensor intercomparison study to evaluate the performance and consistency of the ENVIRA Nanoenvi low-cost sensors. To ensure comparability, we co-located all sensors in an indoor, non-laboratory environment and operated them continuously for four days under nearly identical conditions. We positioned the sensors at a uniform height, spaced approximately 30 cm apart, to minimize differences in exposure to air volume and environmental factors. We computed hourly averages for all monitored variables. Appendix Table C.1 presents summary statistics, including sample size, mean, standard deviation, minimum, maximum, and quartiles. We identified one monitor as an outlier because of its low variability and extremely low pollutant readings. We excluded this monitor from all analyses.

To assess whether imprecise measurements bias our main impact estimates, we calculated the absolute hourly deviation from the overall mean monitor reading for each sensor, using the average across all sensors. We then regressed these absolute deviations on a treatment indicator and hour fixed effects. Results presented in Appendix Table C.2 indicate no statistically significant differences in measurement accuracy between sensors installed in treatment and control classrooms.

Table C.1: Hourly-level sensor-to-sensor intercomparison summary statistics

Parameter	N	Mean	Sd	Min	25th	50th	75th	Max
$PM_{2.5}$	2418	13.78	6.81	5.49	10.09	12.27	14.79	84.34
PM_{10}	2418	13.85	7.04	5.49	10.10	12.29	14.81	96.69
CO	2418	1.23	0.83	0.00	0.41	1.34	1.96	3.57
CO_2	2331	853.40	179.36	380.68	751.45	847.91	936.66	1864.11
Temperature	2418	21.16	0.41	20.05	20.85	21.16	21.45	22.42
Humidity	2418	65.96	2.26	58.00	64.49	66.03	67.60	71.38
Atmospheric pressure	2418	1010.09	1.29	1006.57	1009.17	1010.02	1010.92	1014.07

Notes: The sample size for CO_2 is lower due to missing data in two sensors.

Table C.2: Average treatment effect on absolute deviations in sensor-to-sensor intercomparison

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Absolute deviation from hourly mean						
	Indoor air quality			Other environmental variables			
	$PM_{2.5}$	PM_{10}	CO	CO_2	Temp.	Humidity	Pressure
Estimate	0.084	0.087	−0.143	4.50	0.011	−0.093	0.012
	(0.089)	(0.095)	(0.133)	(5.05)	(0.056)	(0.174)	(0.162)
N.Obs	2,418	2,418	2,418	2,331	2,418	2,418	2,418
Control Mean	4.103	4.176	0.805	123.440	0.334	1.857	1.026

Notes: The table reports the average treatment effects (ATE) on indoor air quality ($PM_{2.5}$, PM_{10}, CO) and other environmental variables (CO_2 and temperature, humidity, and atmospheric pressure) absolute deviations from the hourly means. The sample is restricted to sensor-to-sensor intercomparison days. All models include hour fixed effects. Standard errors are clustered at the sensor level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

D Detecting ventilation episodes

Natural ventilation through open windows and doors alters the air exchange rate between indoor and outdoor environments, which can impact the effectiveness of air purifiers. We investigate how air purifiers affect ventilation behavior in treatment and control classrooms. The installation of purifiers can lead to conflicting outcomes: teachers in treated classrooms may reduce ventilation to optimize purifier performance and limit the infiltration of outdoor pollutants. Conversely, the protection offered by purifiers might encourage more frequent ventilation, even during high pollution periods, based on the assumption that the device lessens associated risks. To identify ventilation events, we leverage the characteristic sharp declines in indoor CO_2 concentrations that typically occur during air exchange with the external environment. These events may also accompany temperature drops when outdoor temperatures fall below indoor levels.

We focus our analysis on school hours and exclude the last 90 minutes before student dismissal to prevent capturing air quality changes linked to end-of-day routines. To reduce measurement noise, we apply a five-minute rolling average to the temperature time series. We define "ventilation episodes" based on specific thresholds for the magnitude and duration of decreases in indoor CO_2 concentrations and, when applicable, temperature. It is important to note that a single instance of window or door opening may result in the detection of multiple ventilation episodes. Figure D.1 illustrates the distribution of detected ventilation episodes under various threshold definitions.

To examine whether air purifiers influenced ventilation behavior, we estimate model 1 using the daily number of ventilation episodes as the dependent variable. Table D.1 reports average treatment effects based on different threshold definitions. For instance, Column 1 shows the treatment effect on the number of episodes per day where CO_2 concentrations decrease by at least 25 ppm per minute and temperature drops by at least 0.005°C per minute for at least a minute. Columns 7 to 12 define ventilation episodes based solely on CO_2 drops, noting that during spring and fall, outdoor temperatures may exceed indoor temperatures, potentially leading to an increase in indoor temperature during ventilation.

Across all specifications, we find no statistically significant differences between the treatment and control groups. This indicates that installing air purifiers did not change ventilation behavior. Since ventilation patterns may vary seasonally with outdoor temperatures, we further split the sample into winter (November–March) and spring (April–June) months. Even when disaggregated by season, we find no significant differences in ventilation behavior between treated and control classrooms.

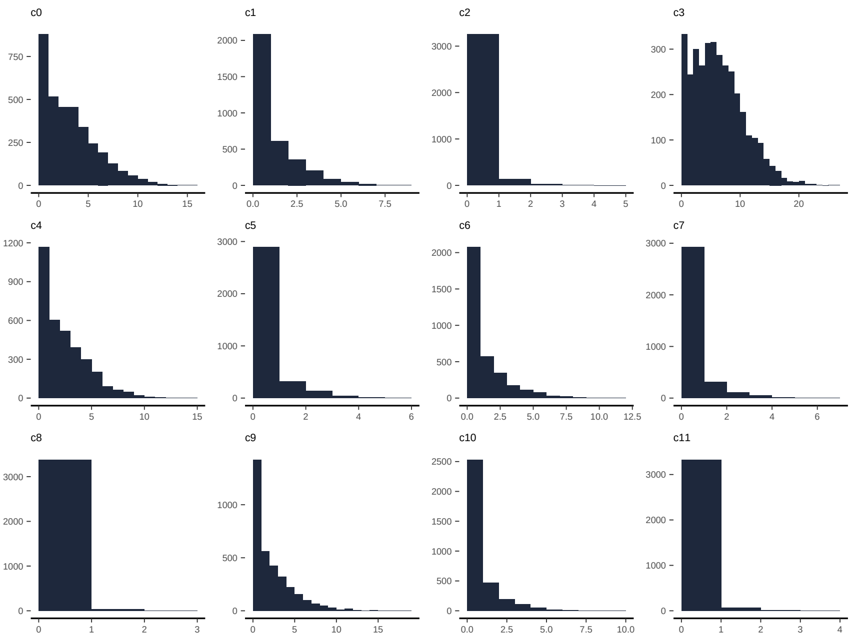


Figure D.1: Distribution of detected ventilation episodes for different thresholds.

Table D.1: Average treatment effect on the number of ventilation episodes

Panel A: CO_2 and temperature thresholds	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta CO_2 < -25$	$\Delta CO_2 < -25$	$\Delta CO_2 < -25$	$\Delta CO_2 < -50$	$\Delta CO_2 < -50$	$\Delta CO_2 < -50$
	$\Delta T < -0.005$	$\Delta T < -0.005$	$\Delta T < -0.005$	$\Delta T < -0.005$	$\Delta T < -0.005$	$\Delta T < -0.005$
	> 1 min	> 2 min	> 5 min	> 1 min	> 2 min	> 5 min
Estimate	0.290	0.172	0.046	-0.014	0.018	0.001
	(0.248)	(0.141)	(0.053)	(0.183)	(0.094)	(0.032)
N.Obs	3,461	3,461	3,461	3,461	3,461	3,461
Control Mean	3.5	1.4	0.2	1.5	0.6	0.1
Panel B: CO_2 thresholds only	(7)	(8)	(9)	(10)	(11)	(12)
	$\Delta CO_2 < -25$	$\Delta CO_2 < -25$	$\Delta CO_2 < -25$	$\Delta CO_2 < -50$	$\Delta CO_2 < -50$	$\Delta CO_2 < -50$
	> 1 min	> 2 min	> 5 min	> 1 min	> 2 min	> 5 min
Estimate	0.327	0.289	0.146	-0.019	0.022	0.005
	(0.412)	(0.232)	(0.106)	(0.268)	(0.144)	(0.053)
N.Obs	3,461	3,461	3,461	3,461	3,461	3,461
Control Mean	6.6	2.7	0.6	2.6	1.0	0.2

Notes: The dependent variable is the number of ventilation episodes per classroom-day, identified by sustained decreases in indoor CO_2 (ppm/min) and temperature (°C/min) lasting at least the specified duration. Panel A defines episodes using joint CO_2 and temperature drops. Panel B uses CO_2 drops only, as outdoor temperatures may exceed indoor temperatures in spring, causing indoor warming during ventilation. All models include calendar date, strata-by-weekday, and strata-by-month fixed effects (strata = school × grade, consistent with Panel A of Table 2). The sample is restricted to school days. Standard errors are clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

E Pre-specified analysis and deviations from the pre-analysis plan

The current study presents deviations from the pre-analysis plan (PAP) uploaded to the Registry of the American Economic Association.[33] We document and explain the choices made in this paper and provide evidence for the commitments outlined in the PAP that were not implemented in the final version.

Main outcomes In the PAP, the main outcomes were: $PM_{2.5}$, absences, cognitive skills, achievement, mood, and aggressive episodes. The manuscript presents evidence on all main outcomes.

Secondary outcomes Most secondary outcomes outlined in the PAP are included in the manuscript, except for data on non-standardized student grades calculated at the end of each semester. After extensive discussions with the directors, we concluded that these evaluations are highly subjective and specific to individual teachers and their classes. Hence, they were not collected.

Data As committed in the PAP, we use absence data from both school years (2022–23 and 2023–24). The two-year specification includes strata-by-treatment-group linear time trends to correct for differential pre-treatment dynamics documented in Appendix F. This correction was not anticipated in the PAP but is motivated by observable patterns in the pre-treatment data.

Models The specification in Section 4.4 for survey outcomes differs from that in the PAP. In this specification, we regress endline outcomes on a treatment dummy, individual characteristics, grade, and school fixed effects. In the PAP, we committed to a two-way fixed effects model that included student and survey wave fixed effects. The reasons for these changes are: i. We conducted only two survey waves instead of the planned three due to budget constraints and the inability to conduct a midline survey in some schools. ii. We did not anticipate attrition rates of 11% and 12% in the first and second waves, respectively, or the missing value patterns affecting certain variables (see Appendix Tables A.1 and A.2). Combining both waves in the same model (panel or ANOVA) would result in a sample reduction of 33% to 45% (depending on the variable considered), leading to a loss of statistical power.

---

[33] It is available at `https://www.socialscienceregistry.org/trials/11960`.

Heterogeneity In the PAP, we committed to examining heterogeneous treatment effects on absences based on pre-treatment absence levels (median split), seasonality (winter vs. spring and fall), and outdoor air pollution (seven-day rolling average, median split). The main text presents all three analyses using the pre-registered median splits (Table 4). Pre-treatment absences refer to the pre-treatment period (school-year 2022–23 and September–October 2023) rather than the previous school year only. Appendix Table A.13 presents finer quartile decompositions of the pollution heterogeneity; results are consistent across rolling window lengths (3, 5, 7, and 10 days).

The manuscript introduces additional dimensions of heterogeneity that were not pre-registered: students' gender and nationality (Appendix Table A.11). These address readers' interest in exploring aspects of heterogeneity that we did not anticipate when designing the study.

Robustness checks As promised in the PAP, we include indoor temperature and humidity in our analysis of air purifiers' impact on indoor pollution. Appendix Table A.5 replicates the results from Panel A of Table 2 (Columns 1-3). Results are consistent. We exclude class characteristics (floor and orientation) because these data are unavailable. We do not repeat this exercise for absences because of sample size limitations and statistical power issues, as the sample is restricted to classes equipped with sensors.

F Trend correction

The PAP specified using absence data from both school years (2022–23 and 2023–24). When pooling both years without adjustment, the LPM treatment effect attenuates from the main estimate of −0.0103 to −0.0017 ($p = 0.522$). This attenuation is driven by a smaller point estimate, not by inflated standard errors: the standard error actually falls from 0.0049 to 0.0027 with the larger sample. This section documents the source of the attenuation and validates the strata-by-treatment trend correction used in the main specification.

Appendix Table F.1 presents three specifications. Column 1 pools both school years without trend correction, yielding the attenuated estimate. Column 2 adds a single aggregate treatment-group differential trend (1 parameter), which recovers the full effect (−0.0102, $p = 0.035$). Column 3 replaces this with strata-by-treatment-group trends (50 parameters), yielding a nearly identical estimate (−0.0103, $p = 0.037$). The consistency between the one-parameter and fifty-parameter corrections confirms that the adjustment is driven by a group-level differential, not by overfitting.

Table F.1: Treatment effect on absences: progressive trend corrections (LPM)

	(1) 2 years	(2) 2 years + agg. trend	(3) 2 years + strata trends
Estimate	−0.0017	−0.0102**	−0.0103**
	(0.0027)	(0.0048)	(0.0049)
p-value	0.522	0.036	0.037
Effect (%)	−2.9	−17.2	−17.4
N.Obs	620,145	620,145	620,145
N.Clusters	95	95	95
Extra trend parameters	0	1	50
Student FE	Yes	Yes	Yes
Date FE	Yes	Yes	Yes
Strata × month FE	Yes	Yes	Yes
Strata × year FE	Yes	Yes	Yes
Treatment × trend	No	Yes	No
Strata × treat. trend	No	No	Yes

Notes: The dependent variable is a binary indicator for daily absence, estimated with a Linear Probability Model pooling school years 2022–23 and 2023–24. Column 1: no trend correction. Column 2: single treatment-group differential linear trend (1 parameter). Column 3: strata-by-treatment-group linear trends (50 parameters), where strata are school × grade (the randomization stratification). All specifications include student, calendar date, strata-by-month, and strata-by-year fixed effects. Standard errors clustered at the classroom level. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

To understand the source of the attenuation, we examine pre-treatment dynamics within randomization strata. We estimate the differential trend between treated and control classrooms within each school × grade stratum using only 2022–23 data (a full year before treatment). Appendix Table F.2 reports the results. Of 20 testable strata (Grade 1 is excluded because those students were not enrolled in 2022–23), 7 show nominally significant differential pre-treatment trends at the 5% level under cluster-robust standard errors, concentrated in School 2 (all four grades), School 3 (Grades 2 and 4), and one stratum in School 5 (Grade 5). Because strata contain only 2–6 classrooms, individual stratum tests are unreliable: valid permutation inference yields minimum achievable p-values of 0.10–1.00 depending on stratum size. An F-test on the full two-year pooled model (both school years, all grades including Grade 1) rejects the null that the 50 strata × treatment-group trends are jointly zero ($p < 0.001$, 95 clusters).

Since randomization was stratified by school × grade with only 2–6 classrooms per cell, such chance imbalances in classroom-level dynamics are not unexpected. If treated classrooms were trending upward relative to control within their strata, the standard two-year specification is biased toward zero.

Table F.2: Pre-treatment differential trends by randomization stratum (2022–23 only)

Stratum	Diff. slope ($\times 10^4$)	p-value	Treated classes	Control classes	Significant ($p < 0.05$)
S1 Grade 2	+0.65	0.158	2	3	
S1 Grade 3	+1.16	0.154	2	3	
S1 Grade 4	−0.42	0.210	2	3	
S1 Grade 5	−0.68	0.294	2	3	
S2 Grade 2	+1.96	<0.001	1	1	✓
S2 Grade 3	+0.40	<0.001	1	1	✓
S2 Grade 4	−2.42	<0.001	1	1	✓
S2 Grade 5	−0.79	<0.001	1	1	✓
S3 Grade 2	+1.31	0.017	2	2	✓
S3 Grade 3	+0.92	0.136	2	2	
S3 Grade 4	+2.26	<0.001	2	2	✓
S3 Grade 5	+0.10	0.791	2	2	
S4 Grade 2	+0.32	0.703	2	3	
S4 Grade 3	+1.50	0.292	2	3	
S4 Grade 4	+0.02	0.983	2	3	
S4 Grade 5	+1.36	0.298	2	3	
S5 Grade 2	+1.90	0.095	1	2	
S5 Grade 3	+0.04	0.930	1	2	
S5 Grade 4	−0.85	0.079	1	2	
S5 Grade 5	+0.39	0.017	2	1	✓

Notes: Each row reports the coefficient on $AirPurifier_c \times t$ from a linear regression of daily absences within the indicated stratum, using only 2022–23 data. Standard errors are clustered at the classroom level; because strata contain only 2–6 classrooms, these individual tests are unreliable. Under exact permutation inference, the minimum achievable p-value ranges from 0.10 to 1.00 depending on stratum size, so no individual stratum can achieve significance at the 5% level. The p-values are reported for completeness; the valid evidence of differential pre-trends is the joint F-test across all 95 clusters ($p < 0.001$). A positive coefficient means treated classrooms were trending upward relative to control classrooms. Grade 1 strata are excluded because those students were not enrolled in 2022–23.