# Do Air Filters Improve Student Learning? Experimental Evidence from Colombia*

Michael Kremer† Mauricio Romero‡ Santiago Saavedra§

July 10, 2025

**Abstract**

Laboratory tests suggest that air filters have the potential to inexpensively reduce indoor air pollution and reduce the transmission of infectious diseases. However, it is unclear whether they can be effective at scale in real-world conditions where filters may not always stay plugged in, windows may be open, and classroom construction may allow airflow from outside. The impact of air filters may vary depending on year-to-year fluctuations in pollution levels, infection risk, and the occurrence of pandemics. Consequently, we assess the expected benefit of a multi-year, large scale, light touch implementation of air filters under real world conditions through a randomized control trial which placed portable air filters in Bogotá schools. In 2023, students in treated schools scored $0.03\sigma$ higher (p-value 0.03) on a high-stakes exam taken 3-4 months after the filters' installation, but in 2024, they had no discernible effects on test scores. Since exams were administered off-site, these effects do not reflect test-day conditions. Filters led to a modest reduction in $PM_{2.5}$ pollution inside classrooms, at .47 $\mu g/m^3$ (p-value 0.082) from a base of 11.00 $\mu g/m^3$ in 2023 and by .66 $\mu g/m^3$ (p-value 0.066) from a base of 16.50 $\mu g/m^3$ in 2024. Effects did not vary by socio-economic status, baseline scores, or pollution levels. Filters' expected benefits would exceed costs, if the 2023 effect of filters on test scores occurs in only half of non-pandemic years, and at least 5% of the test score-wage correlation reflects human capital accumulation.

*We thank Secretaría de Ambiente and Secretaría de Educación for constant help during the project. Andrés Abella, Samuel Blanco, Sebastián Cuéllar Harker, Jack Ellington, Isabella Guerrero, Oscar Garcia, Iván de las Heras, Valentina Londoño, and Santiago Velasquez provided excellent research assistance. This project is funded by J-PAL's King Climate Action Initiative (K-CAI) and the Weiss Fund for Research in Development Economics. We also thank the seminar audiences at Stanford, UCSB, and Columbia for useful comments. Romero gratefully acknowledges financial support from the Asociación Mexicana de Cultura, A.C and the Jacobs Foundation. All errors are our own. AEA-RCT Registry AEARCTR-0009881 (https://www.socialscienceregistry.org/trials/9881). IRB approval from the University of Chicago (22-1163) and Universidad del Rosario (005 638 CS397).

†University of Chicago; kremermr@uchicago.edu

‡ITAM; J-PAL; mtromero@itam.mx

§Universidad del Rosario; santisaap@gmail.com

# 1 Introduction

Air pollution shortens life expectancy by 2.2 years globally (Greenstone, Hasenkopf, & Lee, 2022), and infectious respiratory diseases significantly impair health, education, and productivity, with particularly severe impacts during epidemics. Installing high-efficiency particulate air (HEPA) filters in schools may mitigate these adverse effects by reducing exposure to pollutants, curbing the transmission of routine diseases, and serving as a protective measure during pandemics, potentially reducing the need for school closures (Liu, Lee, & Gershenson, 2021; Lindsley, 2021; Ueki et al., 2022). While private adoption of air purifiers by households — particularly in high-income countries and upper-middle-income contexts — has become common, implementing air filtration systems in public facilities may offer greater cost-effectiveness due to economies of scale and broader reach.[1] However, there are reasons to be cautious about expecting large impacts from air filters in schools. Many classrooms in low- and middle-income countries have vents that allow outside air in, and students or teachers may turn off filters due to noise, limited electrical outlets, or open windows in hot weather, thereby reducing their effectiveness.

Evaluating the effectiveness of air filters involves addressing external validity over time (Rosenzweig & Udry, 2020). Pollution levels fluctuate annually due to factors such as weather variations and forest fires. The burden of infectious diseases also varies significantly, often exhibiting highly skewed impacts; COVID-19 alone is estimated to have reduced global lifetime earnings by US$21 trillion (Azevedo, Akmal, Cloutier, Rogers, & Wong, 2022). To address these concerns, this paper follows the approach of Rosenzweig and Udry (2020) by combining randomized controlled trial evidence from non-pandemic years with modeling to estimate expected impacts during pandemics. Additionally, this analysis explicitly shows how optimal policy decisions depend on policymakers' beliefs. As long as they assign even modest weight to either the economic value of human capital gains associated with test scores or the likelihood that air filters reduce school closures during pandemics, the expected benefits of air filters exceed their costs.

Specifically, we report the impact of placing portable HEPA filters in classrooms on student learning based on a randomized control trial conducted in Bogotá, Colombia, in 2023 and 2024. The study involved more than 42,000 students each year in 357 public schools. Learning outcomes were measured using scores on a nationwide standardized test (Saber 11). Historical test score data helped account for much of the variation in test scores,

[1] Approximately 1 in 4 of all American households owns an air purifier (Consumer Reports, 2019); In Delhi 24.1% of high-SES households own an air purifier device (Greenstone, Lee, & Sahai, 2021).

allowing relatively small treatment effects to be detected.[2] In this setting, students take the Saber 11 exam in central locations rather than their schools, allowing us to isolate the impact of prolonged exposure to air filters from any immediate performance boosts caused by cleaner air on the test day itself.

While reductions in pollution were modest, they were sufficient to yield measurable benefits in education, as evidenced by higher test scores in 2023. In 2023, students exposed to filters scored $.03\sigma$ (p-value 0.03) higher on the Saber 11 test, which was taken 3–4 months after filter installation. Effects did not vary by student characteristics, socio-economic status, or prevalent (outdoors) pollution levels. In 2024, we find no effect of air filters on test scores.

Air filters lowered the level of $PM_{2.5}$ in classrooms by .47 $\mu g/m^3$ (p-value 0.082) from a base of 11.00 $\mu g/m^3$ in 2023, and by .66 $\mu g/m^3$ (p-value 0.066) from a base of 16.50 $\mu g/m^3$ in 2024. Air filters reduce pollution more when the level of outdoor pollution is high, resulting in a -1.2 $\mu g/m^3$ (p-value 0.00) reduction from a base of 15.10 $\mu g/m^3$ in 2023 when outdoor pollution is above the World Health Organization (2021) recommended threshold of $5\mu g/m^3$.[3] In 2024, filters also significantly reduced $PM_{2.5}$ pollution by -1.2 $\mu g/m^3$ (p-value 0.01) from a higher base of 20.60 $\mu g/m^3$ when levels exceeded the WHO threshold.

Assuming filters affect test scores only by reducing exposure to particulate matter, placing air filters can be used as an instrument for indoor air quality. Estimates from 2023 indicate 1 $\mu g/m^3$ increase in indoor $PM_{2.5}$ reduces test scores by $.089\sigma$. If filters also affect test scores by reducing exposure to airborne infectious diseases, this estimate could be interpreted as an upper bound on the local average treatment effect of pollution reduction on learning.

The cost-effectiveness and benefit-cost ratio (BCR) of air filters vary based on assumptions about aggregate shocks and filter durability. In non-pandemic years, filters yield gains of 1.76 LAYS per $100 invested if the 2023 effect repeats annually, or 0.88 LAYS if it occurs half as often. Considering both non-pandemic learning gains and avoided school closures during pandemics, filters compare favorably with other educational interventions classified as cost-effective ("good buys") by the Global Education Evidence Advisory Panel (Angrist et al., 2025; GEEAP, 2023).

Further, even modest test score improvements from filters could significantly boost lifetime earnings. Using the established correlation between Saber 11 scores and Colombian wages, the annual per-student net present value of wage gains ranges between $43.10

[2] Controlling for a school's mean lagged test scores reduces the intra-class correlation coefficient by roughly 90 percent.

[3] These estimates are similar to experimental estimates of portable HEPA filters in Los Angeles, which reduced $PM_{2.5}$ by $0.581\mu g/m^3$ (Simona, Bartell, & Vieira, 2025).

and \$86.19, resulting in a BCR from learning-related wage gains of 20.08 to 40.15. Even if the effects of filters occurred only half of the non-pandemic years and only 5% of the test score-wage correlation reflects actual human capital gains, benefits would still exceed costs. Filters could also provide net benefits by mitigating learning losses due to pandemic school closures. Given estimated annual per-student earnings losses of \$31.08 due to pandemics (Azevedo et al., 2022), reducing school closure probability by just 7% would yield benefits exceeding the filter costs.

This study contributes to three strands of the literature. First, it advances research on policies to mitigate the effects of air pollution. Two studies have found that installing HEPA filters in schools in the US improves learning outcomes (Gilraine, 2023; Biasi, Lafortune, & Schonholzer, 2025).[4] This study differs in the nature and the context of the intervention. In those cases, dramatic changes were made to the ventilation systems in schools, the cost was 10 times higher, and the HEPA filters were not portable. These studies were also done in a high-income country with lower baseline pollution levels. In contrast, this paper evaluates a much lighter touch and less expensive program using portable filters in the context of a middle-income country with higher baseline levels of pollution. Our cost-benefit analysis suggests policymakers should consider expanding air filter installations in classrooms if they assign at least a 5% weight to the test score-wage relationship as reflecting genuine human capital gains, at least a 7% probability that filters avert pandemic-related closures, or any linear combination of these conditions.

Second, it contributes to the evidence on the effects of school health interventions. Interventions such as deworming (e.g., Miguel and Kremer (2004)), providing eyeglasses (e.g., Glewwe, Park, and Zhao (2016)), nutritional interventions (e.g., Glewwe and Miguel (2007); Field, Robles, and Torero (2009); Luo et al. (2012); Sylvia et al. (2013); Araújo, Carrillo, and Sampaio (2021)), and anti-malarial treatment (e.g., Fernando, De Silva, Carter, Mendis, and Wickremasinghe (2006)) are highly cost-effective ways to improve student outcomes. We complement these findings by establishing that air filters may be another such cost-effective policy alternative.[5]

[4]Multiple quasi-experimental and correlational studies find links between air pollution, including particulate matter, and test scores (e.g., Stafford (2015); Ebenstein, Lavy, and Roth (2016); Zhang, Chen, and Zhang (2018); Persico and Venator (2019); Heissel, Persico, and Simon (2020); Balakrishnan and Tsaneva (2021); Gilraine and Zheng (2024); Aguilar-Gomez, Dwyer, Graff Zivin, and Neidell (2022); Duque and Gilraine (2022)). Related papers estimate how contemporaneous pollution levels affect cognitive performance, see Ebenstein et al. (2016); S. Roth (2020); Künn, Palacios, and Pestel (2023).

[5]Routine respiratory diseases cause school absences in normal times (McLean, Peterson, King, Meece, & Belongia, 2017) and lead to learning losses. Further, the annual probability of a pandemic that is at least as severe as COVID-19 is approximately 1% — a one in 105-year event (Marani, Katul, Pan, & Parolari, 2021, 2023; Glennerster, Snyder, & Tan, 2023), potentially leading to large-scale school shutdowns and large learning losses (Azevedo, Hasan, Goldemberg, Geven, & Iqbal, 2021; Azevedo et al., 2022; Azevedo, Cojocaru, Talledo,

Third, if the results are unlikely to be driven by reductions in infectious disease, our results bolster the evidence of the negative effects of air pollution, particularly on cognitive performance and learning. This contribution is fourfold. First, while most papers focus on high pollution on the day of the exam, this paper focuses on reducing indoor pollution during learning sessions. Second, while most prior work analyzes outdoor or ambient air pollution, this paper focuses on indoor air pollution, as (Metcalfe & Roth, 2025). Third, while previous research in this area is based on quasi-experimental evidence, these results provide experimental evidence of how pollution affects learning. Finally, most of the evidence comes from developed countries; two notable exceptions are Zhang et al. (2018), which studies China, and Merkus (2024); Villalobos and Blackman (2025), which studies Colombia. However, the impact of pollution has been shown to vary across settings Arceo, Hanna, and Oliva (2016). This paper complements these results with further evidence from a middle-income country.

# 2 Context

## 2.1 Education in Bogotá

Schooling in Colombia is divided into basic primary education (grades 1-5), basic secondary education (grades 6-9), and upper secondary education (grades 10-11). Education is compulsory through grade 9; the final 2 years are not mandatory but are necessary to access university or technical education. Public schools in Colombia run from January through November.[6] In Bogotá 22.31% of all students and 61.76% of low-income students attend a public school.

In the last year of high school (Grade 11), students must take a nationwide standardized test (Saber 11) to graduate. There are five sections: mathematics, Spanish, natural sciences, social studies, and English, each scored from 0 to 100, with a maximum global score of 500.[7] While no minimum score is required to graduate, the results are one of the primary factors that higher education institutions consider in admissions and scholarship decisions. The Instituto Colombiano para la Evaluación de la Educación (ICFES) administers the exam on a Sunday at central locations (e.g., at a local university). Students in our sample sat the exam in August.[8]

& Narayan, 2023; Schady, Holla, Sabarwal, Silva, & Chang, 2023).

[6] Schools in Colombia follow one of two academic calendars: Calendar A runs from January through November, while Calendar B runs from September through June. All public schools in Bogotá and 98.17% of private schools follow Calendar A.

[7] 'Global' refers to the cumulative score from the five sections: mathematics, Spanish, natural sciences, social studies, and English.

[8] This timing applies to all schools that follow Calendar A, which includes public schools.

## 2.2 Pollution in Bogotá

Particulate matter pollution, comprising airborne particles such as dust, pollen, mold, soot, and smoke, can have a significant impact on human health. Particulate matter is categorized based on the size of the particles. Fine particulate matter smaller than 2.5 micrometers ($PM_{2.5}$) is particularly harmful, compared to larger particles, as it is more easily absorbed deep into the brain (Chen et al., 2017), lungs, or bloodstream when inhaled. However, coarser particulate matter, smaller than 10 $\mu m$ ($PM_{10}$), is also known to be detrimental (Environmental Protection Agency, 2023). High levels of exposure can cause difficulty breathing and irritation to the eyes, throat, or lungs in the short term and chronic respiratory illnesses, cardiovascular disease, and cancer in the long term (World Health Organization, 2022).

Particulate matter concentrations in Bogotá exceed global health standards: the average particulate matter concentrations — 20.05 $\mu g/m^3$ for $PM_{2.5}$ and 55.64 $\mu g/m^3$ for $PM_{10}$ — exceed the World Health Organization (2021)'s guidelines of 5 $\mu g/m^3$ for $PM_{2.5}$ and 25 $\mu g/m^3$ for $PM_{10}$. While consistently above recommended limits, Bogotá's pollution is lower than that of many other capital cities, ranking 67th out of 115 in the 2023 World Air Quality Report (IQAir, 2023) —- Delhi, Hanoi, Jakarta, Beijing, Accra, and Lahore (among many others) have pollution levels that are 3–7 times those of Bogotá.[9]

Children in Bogotá are exposed to harmful levels of air pollution during school hours. Data from classroom monitors suggests that students are exposed to $\sim$11 $\mu g/m^3$ of $PM_{2.5}$ on average during school hours. Particulate matter in Bogotá tends to be at its highest from January to March due to the atmospheric conditions (Mura et al., 2020), months when schools are in session.

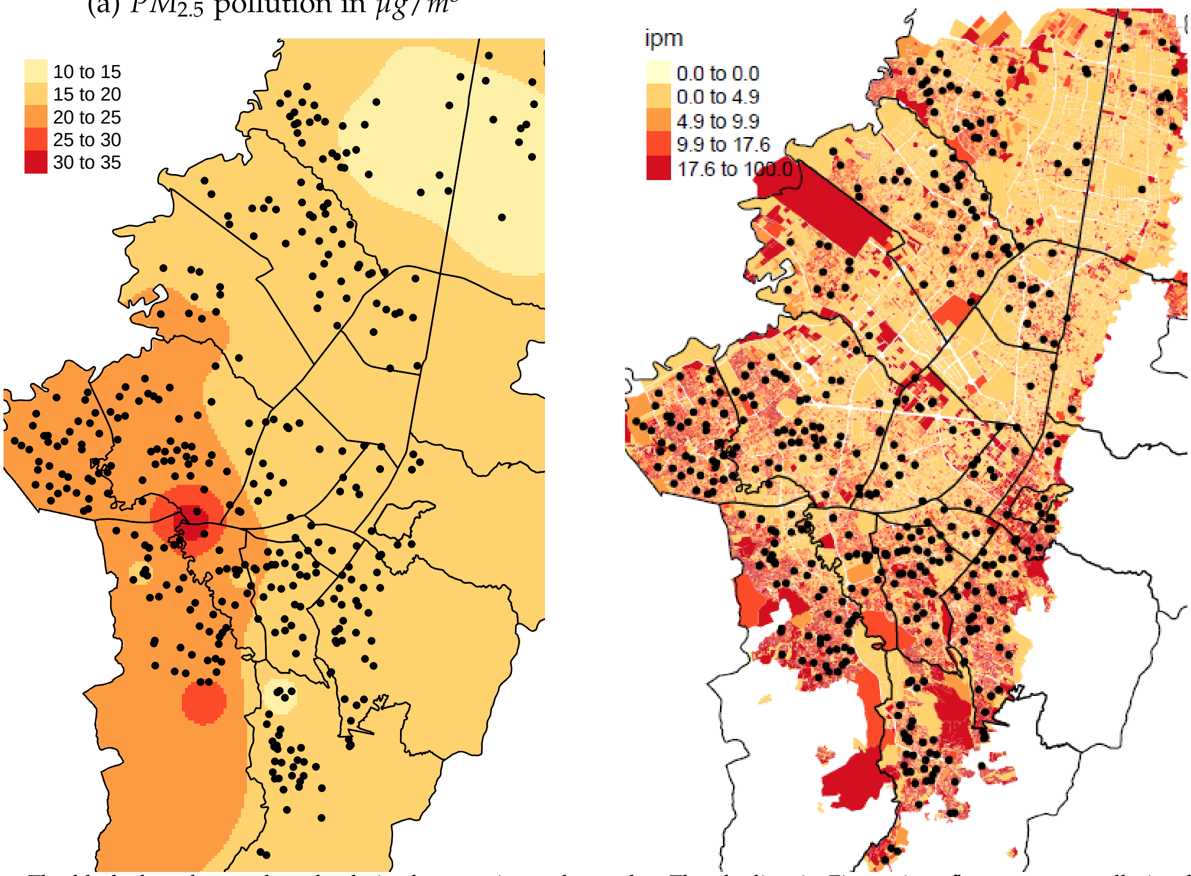
Air quality monitoring reveals stark socio-economic disparities in particulate matter exposure. Pollution levels, as measured by Colombia's Subsistema de Información sobre Calidad del Aire (SISAIRE) system, which operates 22 hourly monitoring stations across Bogotá, are highest in the city's south and west, where lower-income communities reside (Mura et al., 2020). Consequently, students from disadvantaged backgrounds face significantly worse air quality (see Figure 1 and Figure A.1 for graphical correlational evidence).

[9]See `https://www.who.int/data/gho/data/themes/air-pollution/who-air-quality-database/2022` and `https://www.iqair.com/world-most-polluted-cities` for information on pollution in different cities.

Figure 1: Pollution, poverty, and schools' locations

(a) $PM_{2.5}$ pollution in $\mu g/m^3$

(b) Multidimensional poverty index in 2018




*N*otes: The black dots denote the schools in the experimental sample. The shading in Figure 1a reflects average pollution levels in 2022. Raw data on pollution was obtained from SISAIRE, the government's official platform that maintains real-time records of multiple pollution measures `http://aircolombia.sisaire.gov.co/portal/index.col`. For Bogotá, this information is captured by 22 monitoring stations. Pollution levels across the city were estimated using a linear interpolation process. The multidimensional poverty in Figure 1b measures deprivation in five categories: housing and access to public utilities, education, health, work, and childhood conditions. The information is based on the 2018 Census and measures the percentage of poor households according to the multidimensional index (see `https://geoportal.dane.gov.co/visipm/` for more details). School locations were obtained from the Ministry of Education's publicly available databases `https://www.datos.gov.co/Educaci-n/ESTABLECIMIENTOS-EDUCATIVOS-COLOMBIA/upkm-vdjb`.

## 2.3 Respiratory diseases and learning

In a non-pandemic scenario, respiratory illnesses are a leading cause of student absenteeism (Vargas et al., 2020; McLean et al., 2017; Azor-Martínez et al., 2014; Neuzil, Hohlbein, & Zhu, 2002; Castillo-Rodríguez et al., 2022), which contributes to learning loss with long-term economic and social consequences (Liu et al., 2021). By reducing the spread of airborne pathogens, air filtration systems may improve educational outcomes by lowering rates of illness-related absences (Hammond, Khalid, Thornton, Woodall, & Hay, 2021; Curtius, Granzin, & Schrod, 2021; Burgmann & Janoske, 2021).

During a pandemic scenario, school closures can cause severe learning losses.[10] Colombian schools were closed for 25 weeks due to the COVID-19 pandemic (UNESCO, 2020). Similar length closures occurred globally, leading to an estimated global cost from learning losses of USD $21 trillion (Azevedo et al., 2021; Schady et al., 2023). The annual probability of a pandemic that is at least as severe as COVID-19 is approximately 1% — a one-in-105-year event. The likelihood of a pandemic at least half as severe as COVID-19 is approximately 1.54%, or once every 65 years (Marani et al., 2021, 2023; Glennerster et al., 2023). Conservative estimates of the expected present value of learning losses from future pandemics are about USD $1.7 trillion (Glennerster et al., 2023). Air filters may reduce pandemic-related learning losses by making it safer for schools to remain open during a pandemic.

However, rare but severe events, such as the COVID-19 pandemic, make it difficult to accurately generalize about the returns on pandemic preparedness investments without an unrealistically long study. For example, if pandemics occur once every hundred years and filters prevent school closures in half of those events, then over a hypothetical 500-year study, closures would be averted only two and a half times in expectation, and experimental estimates would still be very noisy.

# 3 Experimental design

## 3.1 Sample

The sample consists of 357 public high schools located in urban areas of Bogotá, with over 120,000 Grade 11 students (see Figure 2).[11] Randomization was stratified by locality, and the sample only included localities with at least four schools.

## 3.2 Randomization

Schools were randomly assigned to one of four groups: schools that received only filters, schools that received only monitors, schools that received both filters and monitors, and a control group. The allocation was stratified at the locality level. Seventy-five schools

[10]There is ample evidence of severe learning losses during the COVID-19 pandemic both in high-income (see Patrinos, Vegas, and Carter-Rau (2023); Moscoviz and Evans (2022); Betthäuser, Bach-Mortensen, and Engzell (2023) for reviews) as well as middle- and low-income countries (e.g., see Alasino, Ramírez, Romero, Schady, and Uribe (2024); Guariso and Björkman Nyqvist (2023); Ardington, Wills, and Kotze (2021); Lichand, Doria, Leal-Neto, and Fernandes (2022); Singh, Romero, and Muralidharan (2024)).

[11]Bogotá's administrative boundaries include a large rural area in the south. We exclude schools in this area since they typically have low pollution levels and due to logistical constraints.

were assigned to each treatment group, and the remaining 132 schools were assigned to the control group. Table 1, Panel A presents summary statistics for all eligible schools in 2022 (the year before the treatment). On average, students are 17 years old when they take the exam, and 64% of them come from the lowest socio-economic strata. Overall, the four experimental groups are balanced: the p-value from a joint significance test, where the null hypothesis is that all coefficients are equal to zero, using a system of seemingly unrelated equations, is 0.78. Balance holds at the student- and school-level when each treatment group is compared to the control group and in pairwise comparisons between treatment arms (see Table A.1 and Table A.2).

There are no concerns of differential attrition (see Table 1, Panel B and Panel C). The likelihood that we observe a school in the post-treatment years is balanced, as is the number of students that take the Saber 11 test.[12] In 2024, the test of equality of means across groups yields a p-value that is marginally significant at the 10% level. The observable characteristics of students (age, sex, maternal education, and SES strata) are also balanced in the post-treatment years (see Tables A.3, A.4, and A.5).

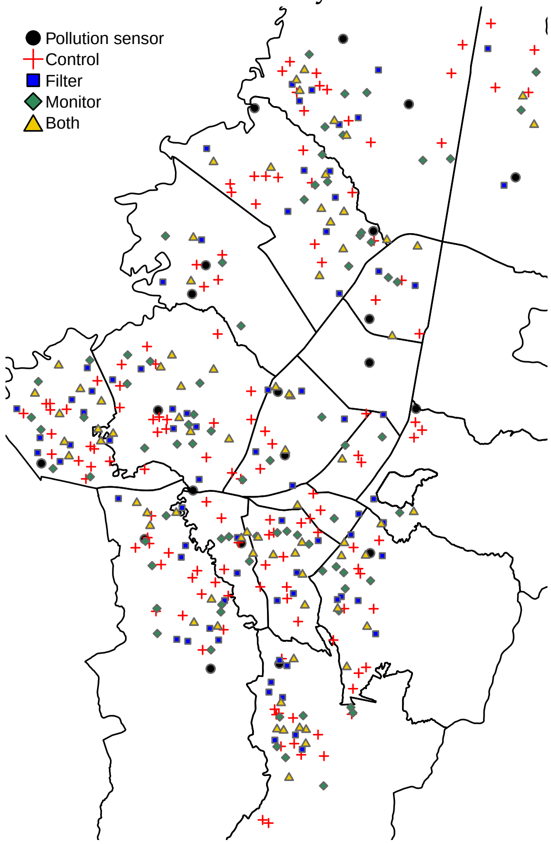
[12]Schools were selected based on 2021 data showing grade 11 students, but 11 schools had no students taking the Saber 11 exam in 2023 (4 control schools, 2 filter only schools, 3 monitor only schools, and 2 filter and monitor schools).

Table 1: Balance across experimental groups (2022)

| | Mean (SD) for each group | | | | p-value |
|---|---|---|---|---|---|
| | Control | Filter | Monitor | Both | (equality) |
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: Pre treatment (2022)** | | | | | |
| Global score | 0.01 | 0.01 | -0.00 | 0.01 | 0.934 |
| | (0.31) | (0.33) | (0.33) | (0.31) | |
| Reading score | 0.01 | 0.01 | 0.00 | 0.01 | 0.982 |
| | (0.26) | (0.28) | (0.28) | (0.25) | |
| Math score | 0.02 | 0.01 | -0.01 | 0.02 | 0.750 |
| | (0.28) | (0.31) | (0.30) | (0.29) | |
| SES strata 1 or 2 | 0.67 | 0.70 | 0.61 | 0.66 | 0.061* |
| | (0.25) | (0.24) | (0.26) | (0.24) | |
| Lower SES | 0.28 | 0.26 | 0.34 | 0.30 | 0.067* |
| | (0.25) | (0.24) | (0.27) | (0.25) | |
| Female | 0.55 | 0.54 | 0.56 | 0.54 | 0.251 |
| | (0.08) | (0.05) | (0.09) | (0.03) | |
| Age (exam date) | 17.39 | 17.47 | 17.49 | 17.46 | 0.900 |
| | (1.59) | (0.53) | (0.36) | (0.29) | |
| Mother some secondary education | 0.50 | 0.50 | 0.50 | 0.50 | 0.770 |
| | (0.06) | (0.05) | (0.06) | (0.04) | |
| PM2.5 | 17.54 | 17.56 | 17.45 | 17.53 | 0.835 |
| | (1.95) | (2.19) | (2.05) | (1.77) | |
| PM10 | 36.08 | 35.94 | 35.37 | 35.91 | 0.192 |
| | (5.49) | (5.72) | (5.81) | (5.52) | |
| Without Saber 11 data | 0.03 | 0.01 | 0.03 | 0.01 | 0.756 |
| | (0.17) | (0.12) | (0.16) | (0.12) | |
| Students (Grade 11) | 116.08 | 127.19 | 103.83 | 111.81 | 0.254 |
| | (73.32) | (90.82) | (60.08) | (66.06) | |
| F-test p-value | | | | | 0 .783 |
| **Panel B: Treatment (2023)** | | | | | |
| Without Saber 11 data | 0.03 | 0.03 | 0.04 | 0.03 | 0.958 |
| | (0.17) | (0.16) | (0.20) | (0.16) | |
| Students (Grade 11) | 121.94 | 132.07 | 104.01 | 114.49 | 0.111 |
| | (79.39) | (93.24) | (63.59) | (69.33) | |
| | [132] | [75] | [75] | [75] | |
| **Panel C: Treatment (2024)** | | | | | |
| Without Saber 11 data | 0.02 | 0.03 | 0.04 | 0.03 | 0.918 |
| | (0.15) | (0.16) | (0.20) | (0.16) | |
| Students (Grade 11) | 124.46 | 130.80 | 104.07 | 112.64 | 0.096* |
| | (78.45) | (97.42) | (65.59) | (65.90) | |
| | [132] | [75] | [75] | [75] | |

*Notes*: Columns 1–4 show the mean, standard deviation (in parentheses) and number of observations (in square brackets) for each of the following groups: Control, schools where filters were installed, schools where air quality monitors were installed, and schools where both filters and air quality monitors were installed. Column 5 reports the p-value of whether the mean is the same across all groups. The equality test considers the randomization design (i.e., it includes strata fixed effects and clusters standard errors at the school level). All variables except age, number of students, pollution measures, and test scores are % with respect to the number of students in Grade 11. $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

Figure 2: Distribution of study schools across the city




*N*otes: Red crosses indicate control schools; blue squares indicate schools treated only with filters; green rhomboids indicate schools treated with pollution monitors only; and golden triangles indicate schools treated with both filters and pollution monitors. Black dots indicate government pollution monitoring stations.

## 3.3 Interventions

### 3.3.1 Filters

Schools were provided Smart Air HEPA Sqair filters, which are recommended for use in rooms up to $40m^3$.[13] These are portable devices that consist of a fan with a HEPA filter pad attached, have low electricity consumption (6-38 watts), and are relatively quiet (23-52 decibels).[14] The equipment lasts up to 8 years and costs 287.63 USD (inclusive of the attached HEPA filter pad, import fees, transportation costs, and installation). However, HEPA filter pads have a shorter lifespan than the fan and should be replaced roughly every 9 months (if used 6 hours per school day on average) to maintain efficiency. The HEPA replacement pad costs 54 USD (including installation).

[13]See `https://smartairfilters.com/en/product/sqair-air-purifier/` for more details.

[14]For comparison, a typical air conditioner consumes between 500-1500 watts and generates noise levels of approximately 50-60 decibels, while an iPhone generates approximately 30 decibels during use.

Air filters were placed in Grade 11 classrooms in treated schools. We placed, on average, three filters in each school, with each filter located in a different classroom. Teachers and students in schools with filters received training on how to use them. They also received an infographic with key information, as well as reminders throughout the school year on the potential benefits of using the equipment. Schools were revisited in July to check filter usage and deliver a pamphlet with instructions on filter use (see Figure A.2). The mayor's office randomly selected 14 schools for additional in-person training sessions on best practices for filter usage.

Students may sit in the same classroom all day or rotate across classrooms for different subjects, depending on the school. If students rotated, we placed the filters in the classrooms where their exposure to the filter during core subjects (Math, Spanish, Science, and Social Sciences) was maximized.[15] We provided a total of 421 filters to 144 schools. On average, the schools that received filters had 123 students each, corresponding to approximately one filter for every 41 students.

### 3.3.2 Monitors

A local NGO, Trebola, installed Plantower PMS7003 monitors in treated schools.[16] The monitors were connected to the school's Wi-Fi when possible (85% of monitors). The rest collected the readings on a microSD card, which a student from Universidad del Rosario retrieved every 3 months. However, the data from these cards had to be excluded because they contained incorrect dates and times.[17] During the installation process, teachers and students were informed about how the monitors work and provided with a link to access live and historical data collected by the monitors. They were also provided a pamphlet with information about the monitors and a QR code to access the data (see Figure A.3).

In each school, one monitor was installed outdoors, typically in the school courtyard, and two indoors (in Grade 11 classrooms). In schools with air filters, the two monitors were randomly located in two of the three classrooms with filters. In schools without filters, the monitors were installed in Grade 11 classrooms that would have received a filter if the school had been assigned to the treatment group. A total of 426 air quality monitors were installed in 146 schools.

---

[15] When we arrived at the school, we asked if the students rotated classrooms. If possible, we requested that the schedule be adjusted to select classrooms that devote more hours to core subjects. If we were unable to obtain the schedule, we installed them in the classrooms where the school coordinator indicated Grade 11 students spent most of their time.

[16] See https://trebola.org/experiencias/ciencia-ciudadana/ for more information about Trebola and its work.

[17] The excluded (microSD) schools are not systematically different from the non-excluded (WiFi) schools.

## 3.4 Compliance

Most schools accepted the equipment, so this paper focuses on intent-to-treat effects throughout. Only five schools in the filters-only group, three in the monitor-only group, and one in the filters + monitors group declined the equipment.

## 3.5 Timeline

Filters were installed in all compliant treatment schools by early June 2023, while the Saber 11 exam took place on August 13, 2023. A quarter of schools assigned to the filter treatment had a filter by February 23, 50% had one by April 21, 75% by May 11, and 90% by May 23 (see Figure A.4). Therefore, by the 2023 exam date, filters had been in place for at least 2 months (one during the school year and one during the holidays) in all schools, and 3 - 6 months in a subset.

The monitors were installed between June and November 2023. Installation was initially slow due to contracting issues and problems connecting the monitors to the schools' Wi-Fi. Thus, as of July 13, a month before the test, only two schools allocated to receive air quality monitors had them installed. By August 2, 25% of treatment schools had monitors, 50% by August 14, 75% by September 1, and 100% by the end of November (see Figure A.5). Thus, students were barely exposed to the monitors in 2023 but were fully exposed in 2024. This timeline alters the empirical strategy for analyzing student outcomes in 2023 (see Section 3.7). Figure A.6 shows the analyzed sample and the distribution of filter and monitor assignments over time.

## 3.6 Data

### 3.6.1 Student learning

ICFES provides anonymized Saber 11 aggregate and subject-wise test scores for each student, along with school and socio-demographic information, such as socio-economic strata, employment status, access to a computer or internet at home, and parents' education levels.[18] The exam is administered in centralized locations (e.g., local universities), allowing us to capture the effects of filter exposure during instruction, rather than any benefits from cleaner air on the test day.

[18] In Colombia, households are classified into six socioeconomic strata (estratos) based on their location. This classification, established by Law 142 of 1994, is used to allocate utility subsidies and cross-subsidies. Households in lower strata (1, 2, and 3) receive subsidized utility rates, while those in higher strata (5 and 6) pay surcharges to fund these subsidies. Stratum 4 is considered neutral, paying cost-reflective rates. While the system has been criticized for inefficiencies and potential distortions in targeting, it is widely used as a proxy for socioeconomic status.

#### 3.6.2 Air quality and air filter usage

Monitors register air quality readings of $PM_{2.5}$ every 15 minutes. Data from July to December 2024 is included in the analysis. In addition to our pollution monitors, we use hourly data on various pollution measures ($PM_{2.5}$, $PM_{10}$, $NO_2$, $CO$, and $O_3$) obtained from the publicly available SISAIRE databases, which gather air quality data from 22 monitoring stations across Bogotá (see Figure 2). A simple interpolation (based on the quadratic distance from the monitoring station) is used to calculate hourly pollution measurements for each school.

City pollution monitoring station sensors provide more accurate readings than the Plantower PMS7003 monitors installed in schools, but require interpolation to estimate pollution near each school (and do not provide indoor air quality measures). Readings from SISAIRE and school monitors are tightly correlated (see Figure A.7). School outdoor monitors consistently report lower pollution levels than city sensors, likely due to their placement in courtyards shielded from direct traffic exposure.

#### 3.6.3 School and classrooms survey

Classroom-level data were collected during installation of air quality monitors or filters, including room size, number of windows, whether windows could open/close, subjects taught, and class frequency. In the treatment group, 66% of classrooms had vents around windows, allowing uncontrolled outdoor airflow that may reduce filter effectiveness (see Figure A.8). In 77% of treatment schools, Grade 11 students rotated classrooms, spending only part of the day in treated rooms.

### 3.7 Empirical strategy

#### 3.7.1 Main estimation

The intent-to-treat effects of the three treatments relative to the control group are estimated using ordinary least squares. The main estimates rely on an ANCOVA-style specification that interacts baseline measures of the outcome (demeaned) with treatment indicators, the most efficient estimator in this context (J. Roth & Sant'Anna, 2023).

Specifically, the following equation is estimated:

$$
\begin{aligned}
Y_{isl} \quad = \quad & \beta_F Filter_s + \beta_M Monitor_s + \beta_B Both_s + \\
& \alpha_1 \overline{Y_{s,-1}} + \alpha_2 \overline{Y_{s,-1}} \times Filter_s + \alpha_3 \overline{Y_{s,-1}} \times Monitor_s + \\
& \alpha_4 \overline{Y_{s,-1}} \times Both_s + \gamma_l + \epsilon_{isl}, \qquad (1)
\end{aligned}
$$

where $Y_{isl}$ is the standardized test score of student $i$, graduating from school $s$ located in locality $l$, $Filter_s (= 1)$ indicates that school $s$ received only filters, $Monitor_s (= 1)$ that school $s$ received only monitors, and $Both_s (= 1)$ that it received both. $\overline{Y_{s,-1}}$ are the demeaned average lagged scores of the school, and $\gamma_l$ are dummies for localities, the unit of stratification. Standard errors are clustered at the school level (the unit of randomization).[19]

Variations of these models are estimated by adding student controls (gender, age, parental education, and socio-economic strata of their household) and other school controls (pollution measured by the SISAIRE network). Coefficients of interest are $\beta_F$, $\beta_M$, and $\beta_B$, which estimate the effect of supplying filters, monitors, and both respectively.[20] Pooling filter vs non-filter schools.

### 3.7.2 Saber 2023

To analyze the 2023 scores, the analysis pools "monitor only" schools with controls and "monitor + filter" schools with "filter only", since most monitors were installed after the 2023 exam. The following equation is estimated:

$$Y_{isl} = \beta_F Filter_s + \alpha_1 \overline{Y_{s,-1}} + \alpha_2 \overline{Y_{l,-1}} \times Filter_s + \gamma_l + \epsilon_{isl}, \tag{2}$$

where all variables are defined as in equation (1).[21] In 2024, we find no effects of air quality monitors on student outcomes (i.e., if knowing the pollution concentration in the classrooms leads teachers or students to take additional measures to reduce pollution, such as keeping doors or windows closed), and so keep the main specification the same (i.e., pooling schools with filters and comparing them to all schools without filters, regardless of whether they have monitors). However, we also show results with the effects for the three original treatment groups in the Appendix. Overall, our results are quantitative and qualitatively similar.

### 3.7.3 Effect of air quality on test scores

As continuous monitoring of filter usage is not possible, the difference between indoor and outdoor air quality readings at schools with and without filters serves as a proxy for the actual use of filters. Restricting the sample to schools with air quality monitors, the analysis

[19] As J. Roth and Sant'Anna (2023) highlight, interacting the demeaned lagged scores with the treatment dummies is more efficient than simply controlling for the lagged scores (see Lin (2013)) as McKenzie (2012) proposed in the presence of heterogeneous treatment effects.

[20] There is no pre-treatment individual student performance data.

[21] The "controls + monitor only" and "filter + both" groups are generally balanced (see Tables A.6 and A.7).

estimates the effect of filter installation on indoor pollution by comparing schools with both filters and monitors to those with only monitors, using the following equation:

$$P_{slt} = \lambda_F Filter_s + \alpha_2 P^o_{slt} + \gamma_l + \gamma_t + \epsilon_{slt}, \tag{3}$$

where $P_{slt}$ is the air quality (measured by $PM_{2.5}$) inside school $s$ in locality $l$ at time $t$. $Filter_s(=1)$ indicates whether school $s$ was assigned to receive the filters (in addition to the monitors). $P^o_{slt}$ denotes outdoor pollution near the school, measured by the installed monitors and the city sensors. $\gamma_l$ and $\gamma_t$ are locality and time (e.g., hour and day of the week) fixed effects. Standard errors are clustered at the school level (the unit of randomization). The coefficient of interest is $\lambda_F$, which measures the effect of assigning a school to receive filters *and* monitors (vis-á-vis receiving only monitors).

The effect of air pollution on student outcomes is estimated using two-stage least squares, where air pollution is instrumented by the random assignment of filters. The sample is restricted to schools that received air quality monitors. The first stage is:

$$P_{slt} = \lambda_1 Filter_s + \theta_1 P^o_{slt} + \gamma_l + \gamma_t + \epsilon_{slt} \tag{4}$$

And the second stage is:

$$Y_{isl} = \beta_0 + \beta_{PM_{2.5}} \widehat{P_{slt}} + \alpha_1 \overline{Y_{s,-1}} + \alpha_2 \overline{Y_{s,-1}} \times Filter_s + \gamma_l + \epsilon_{isl} \tag{5}$$

The coefficient of interest is $\beta_{PM_{2.5}}$, which measures the effect of air quality ($PM_{2.5}$) on student outcomes. Under the assumption that filters affected test scores only through particulate matter, this approach estimates the impact of particulate matter pollution on learning. If filters also raise test scores by reducing infection rates of respiratory diseases, this could be considered an upper bound of the effect of particulate matter pollution on learning.

# 4 Results

## 4.1 Test scores

Students exposed to the filters for 3-4 months before the Saber 11 exam scored $.03\sigma$ (p-value 0.03) higher in 2023, as measured by the global score. The improvement primarily comes from scores in math, Spanish, and social sciences, with increments of

$.04\sigma$ (p-value 0.02), $.03\sigma$ (p-value 0.04), and $.03\sigma$ (p-value 0.02), respectively. These results are robust to the inclusion of pollution measures and student demographic controls (age, gender, parental education, and socioeconomic strata).

There is no heterogeneity in treatment effects by school characteristics (historical performance in the Saber test and pollution levels) or student characteristics (sex, age, and socio-economic strata) (see Table A.9). There are no heterogeneous effects by classroom features such as size, number of windows, and vents. (See Table A.10). Treatment effects are also estimated using unconditional quantile regressions (Firpo, Fortin, & Lemieux, 2009) with the RIF package in Stata (Rios-Avila, 2020). Positive effects appear across the test score distribution, with minimal heterogeneity A.9).

Table 2: Effect of filters on test scores

| Dep. var. | Math | Spanish | Science | Social Sciences | English | Global |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Lagged scores** | | | | | | |
| Treatment (filters) | 0.040** | 0.030** | 0.020 | 0.029** | 0.019 | 0.033** |
| | (0.0175) | (0.0143) | (0.0155) | (0.0130) | (0.0178) | (0.0154) |
| N. of obs. | 42,389 | 42,389 | 42,389 | 42,389 | 42,381 | 42,389 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.076 | 0.064 | 0.073 | 0.058 | 0.082 | 0.087 |
| **Panel B: Lagged scores + pollution** | | | | | | |
| Treatment (filters) | 0.038** | 0.027* | 0.013 | 0.024* | 0.020 | 0.029* |
| | (0.0180) | (0.0147) | (0.0153) | (0.0133) | (0.0189) | (0.0159) |
| N. of obs. | 42,389 | 42,389 | 42,389 | 42,389 | 42,381 | 42,389 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.076 | 0.064 | 0.074 | 0.058 | 0.082 | 0.087 |
| **Panel C: All controls** | | | | | | |
| Treatment (filters) | 0.030* | 0.023 | 0.0061 | 0.019 | 0.015 | 0.023 |
| | (0.0173) | (0.0140) | (0.0151) | (0.0133) | (0.0179) | (0.0153) |
| N. of obs. | 42,389 | 42,389 | 42,389 | 42,389 | 42,381 | 42,389 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.17 | 0.12 | 0.15 | 0.10 | 0.14 | 0.17 |

*Notes*: All regressions consider the randomization design (i.e., include strata fixed effects). Panel A controls for average school-level lagged scores (we do not have individual student-lagged scores). Panel B also controls for the average pollution levels of $PM_{2.5}$, $PM_{10}$, $NO_2$, $CO$, and $O_3$ during school hours (7 am–4 pm) in the 270 days (roughly the length of the school year) and 90 days (roughly the length of exposure to the filters) preceding the exam. Panel C controls for lagged scores, pollution levels, age, sex, mother's level of education, and socio-economic strata. Standard errors, clustered at the school level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

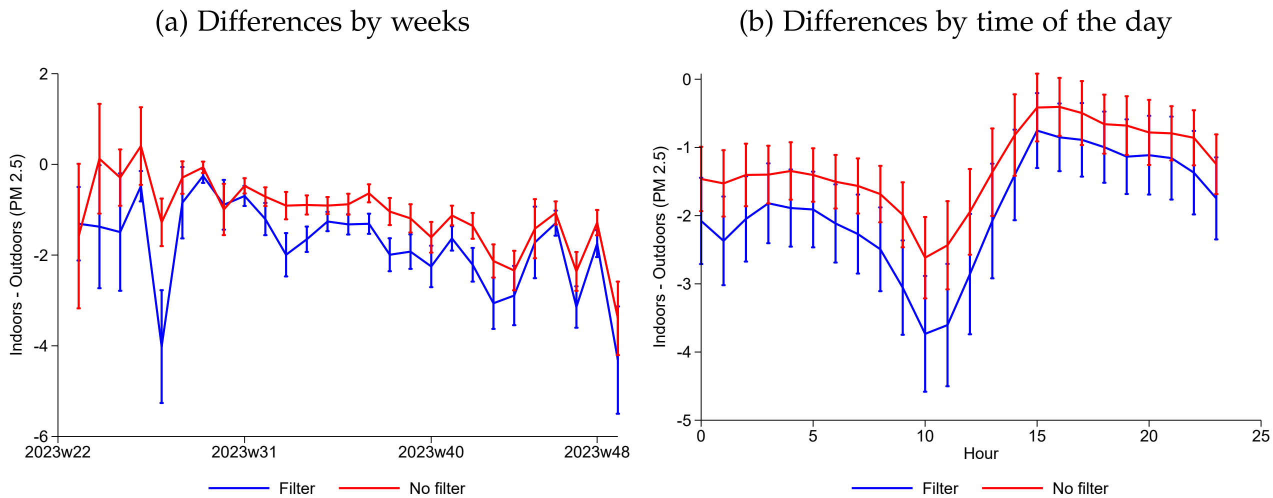
When estimating treatment effects using the original pre-specified arms, global test scores increase by $.04\sigma$ (p-value 0.09) for schools that received filters only and by $.03\sigma$ (p-value 0.08) for schools that received filters and monitors. No effect on test scores is found in schools that received monitors only (see Table A.8), as expected given the delayed installation.

In contrast, we find no effects of the filters (or the monitors) on test scores in 2024 (see Table B.1).

## 4.2 Air quality

To assess the impacts of the filters on indoor air pollution, the analysis is restricted to schools with monitors (those in the monitor-only and monitor + filter groups). The gap between indoor and outdoor pollution is typically negative, given that outdoor pollution tends to be higher. Schools with filters show a consistently wider indoor-outdoor gap than those with only monitors (Figure 3b), with this difference increasing over time 3a).[22] However, due to an internet outage in the schools, we do not have pollution measures in the weeks leading to the 2024 exam (see Figure B.2).

Figure 3: Difference between indoor and outdoor readings in schools with and without filters




*N*otes: These figures display the average difference between indoor and outdoor readings in each school, depending on whether they have filters. Figure 3a averages the data across weeks, while Figure 3b averages the data across hours of the day.

Filters reduce indoor $PM_{2.5}$ by .47 $\mu g/m^3$ (p-value 0.082) during school hours in 2023 (see Table 3) and by .66 $\mu g/m^3$ (p-value 0.066) in 2024 (see Table B.2).[23] The reduction including all hours and weekends is similar at .54 $\mu g/m^3$ (p-value 0.02) in 2023 and .8 $\mu g/m^3$ (p-value 0.02) in 2024, which may be driven by schools not shutting down the filters when students are not in the classroom.[24] Teachers and students were told it was particularly important to turn the filters on whenever pollution was above the WHO's recommended threshold of

[22]The confidence intervals vary by week since some schools turn off their electricity during school holidays, which shuts down the monitors. Figure A.10 displays the evolution of the readings over time. Table A.11 reports no differential attrition across treatment groups in the pollution monitor data. However, schools with filters and monitors are slightly more likely to have missing data from their outdoor monitors compared to those with only monitors.

[23]Since the standard deviation of indoors $PM_{2.5}$ is 9.2 $\mu g/m^3$, the treatment effect over school hours is equivalent to -.05$\sigma$.

[24]Some teachers mentioned they leave the filters on, regardless of whether there is a class, pollution is high, or it is the weekend.

5 $\mu g/m^3$ for $PM_{2.5}$ (World Health Organization, 2021). When outdoor pollution is above the WHO threshold, filters reduce indoor pollution by -1.2 $\mu g/m^3$ (p-value 0.00) in 2023 and -1.2 $\mu g/m^3$ (p-value 0.01) in 2024, suggesting they are effective when pollution is high (which could be driven by teachers turning the filters on when pollution was high, such as at the beginning of 2024 due to fires in nearby regions). As expected, filters have no effect on outdoor pollution (see Table A.12).

Table 3: Effect of filters on students' exposure to $PM_{2.5}$

| Dep. var. | $PM_{2.5}$ Indoor pollution ($\mu g/m^3$) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Filter and Monitor | -.47* | -.53** | -.54** | -.47* | -1.2*** |
| | (.24) | (.24) | (.23) | (.27) | (.42) |
| Control mean | 9.5 | 9.5 | 9.5 | 11 | 15 |
| N. of obs. | 1,405,455 | 1,405,455 | 1,405,447 | 410,059 | 205,102 |
| Locality FE | X | X | | | |
| Hour FE | X | | | | |
| Date FE | X | | | | |
| Hour $\times$ Date FE | | X | | | |
| Locality $\times$ Date $\times$ Hour FE | | | X | X | X |
| Hours | All | All | All | 7am–4pm | 7am–4pm |
| Days | All | All | All | Mon–Fri | Mon–Fri |
| Outdoor pollution | All | All | All | All | $> 5$ |

*Notes*: The outcome in all regressions is the indoor readings of $PM_{2.5}$. All regressions control for outdoor readings of $PM_{2.5}$ (demeaned), both as measured by the monitors we installed in schools and the city monitors and their interaction with the treatment dummy (following J. Roth and Sant'Anna (2023); Lin (2013)). All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

There is no heterogeneity in pollution reduction by outdoor pollution levels as measured by outdoor monitors and the city's monitoring systems. Results are similar whether outdoor pollution is measured using outdoor monitors installed for the study or the city's monitoring systems (see Table A.13 and Figures A.12, A.11, and A.13).

### 4.3 Effect of air quality on test scores

The effect of indoor $PM_{2.5}$ pollution on test scores is estimated using the experimental assignment of filters as an instrument for $PM_{2.5}$ levels. Three caveats apply: First, for 2023, $PM_{2.5}$, which is used to proxy filter usage, is only available for dates after the Saber 11 exam. Pre-exam usage is assumed to be similar to post-exam usage. Second, the instrumental variable approach assumes filters only affect scores via $PM_{2.5}$ (the exclusion restriction), but filters likely reduce exposure to all particulate matter (including $PM_{10}$ and $PM_{0.1}$), so this should be interpreted as an effect of reductions in all types of particulates, not $PM_{2.5}$ specifically. Finally, filters likely reduce the risk of respiratory infections ((Hammond et al., 2021; Curtius et al., 2021; Burgmann & Janoske, 2021)). Thus, the estimate may be an upper bound of the effect of particulate matter on learning.

The sample is restricted to 114 schools that have both Saber 11 scores and monitors (58 in the filter and monitor group and 56 in the monitor-only group). In 2023, in these schools, filters reduced indoor $PM_{2.5}$ by .78 $\mu g/m^3$ (p-value 0.02) from a base of 11.2 (Panel A, Table 4), indicative of an .089$\sigma$ (p-value 0.03) increase in test scores (Panel C, Table 4). In 2024, filters reduced indoor $PM_{2.5}$ by .96 $\mu g/m^3$ (p-value 0.04) from a base of 17.2 (Panel A, Table B.3), indicative of an .0051$\sigma$ (p-value 0.87) increase in test scores (Panel C, Table B.3).

Differences from sections 4.1–4.2 arise due to sample variation, the comparison between the filter + monitor and monitor-only groups, and the inclusion of outdoor pollution controls. The Wald estimator from the full sample in 2023 suggests a 1 $\mu g/m^3$ increase in indoor $PM_{2.5}$ reduces test scores by $\sim$0.045$\sigma$.

Table 4: Effect on 2023 Saber tests

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **Panel A: First Stage** | | | | | | |
| Dep. var. | $PM_{2.5}$ | | | | | |
| Treatment (filters) | -0.78** | | | | | |
| | (0.324) | | | | | |
| N. of obs. | 12,950 | | | | | |
| Only monitors mean | 11.2 | | | | | |
| N. Schools | 114 | | | | | |
| $R^2$ | 0.45 | | | | | |
| **Panel B: Reduced Form** | | | | | | |
| Dep. var. | Math | Spanish | Science | Social Sciences | English | Global |
| Treatment (filters) | 0.061** | 0.073*** | 0.046* | 0.070*** | 0.038 | 0.070*** |
| | (0.0263) | (0.0251) | (0.0234) | (0.0256) | (0.0264) | (0.0249) |
| N. of obs. | 12,950 | 12,950 | 12,950 | 12,950 | 12,948 | 12,950 |
| N. Schools | 114 | 114 | 114 | 114 | 114 | 114 |
| $R^2$ | 0.080 | 0.064 | 0.079 | 0.060 | 0.074 | 0.089 |
| **Panel C: IV Estimates** | | | | | | |
| Dep. var. | Math | Spanish | Science | Social Sciences | English | Global |
| Average PM 2.5 (indoors) | -0.079* | -0.094** | -0.059** | -0.090** | -0.048 | -0.089** |
| | (0.0425) | (0.0474) | (0.0299) | (0.0405) | (0.0365) | (0.0414) |
| N. of obs. | 12,950 | 12,950 | 12,950 | 12,950 | 12,948 | 12,950 |
| N. Schools | 114 | 114 | 114 | 114 | 114 | 114 |
| $R^2$ | 0.077 | 0.056 | 0.078 | 0.055 | 0.071 | 0.084 |

*Notes*: All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. The p-value of jothe int F-test for the first stage is 5.769. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## 4.4 Cost effectiveness

### 4.4.1 Methodology

Next, we assess whether scaling up air filters in schools is worthwhile under two scenarios of filter effectiveness: (1) assuming the 2023 effect applies in all non-pandemic years, and (2) assuming the 2023 effect applies in half of the non-pandemic years (since we do not find effects in 2024). This analysis evaluates the policy value of scaling up air filters in schools using two complementary approaches. First, from a cost-effectiveness perspective, by comparing the learning gains per 100 USD to other education interventions. Second, by directly comparing monetary benefits and costs. Table 5 shows the key parameters used in the analysis.

#### 4.4.2 Costs

Panel A presents the costs. One filter costs 288 USD, which includes the HEPA filter pad, import fees, transportation, and installation (Table 5, row 1). According to the manufacturer, they last at least 8 years — only 1 out of 421 broke in the first year of the experiment (Table 5, row 2). The yearly replacement HEPA filter pad costs 54 USD (Table 5, row 3), and is needed for every year, excluding the year of filter purchase.[25] The annual electricity cost per filter is \$11, assuming 8 hours of usage per day, 20 days per month, over 9 months of the academic year (Table 5, row 4). One filter serves 41 students (Table 5, row 5). A discount rate of 3% is used, based on the parametrization of Christiano, Eichenbaum, and Evans (2005) and López, Prada, and Rodríguez N. (2008) (Table 5, row 6). The annual per-student net present value of installing and maintaining one filter throughout its minimum expected lifespan is \$2.15 (5, row 6).

#### 4.4.3 Cost-Effectiveness in Non-Pandemic Years

Aiming to model filter effect variation under aggregate shocks, this section presents the analysis for two scenarios: one in which filters deliver the 2023 effect in all non-pandemic years (Table 5, column 1; hereafter, the *all-years* scenario), and another in which they do so in only half of those years (Table 5, column 2; hereafter, the *half-years* scenario). In the *all-years* scenario, air filters improve test scores by $.03\sigma$ (see Table 2, column 6 and Table 5, row 8), which is equivalent to a 0.038 LAYS gain per student year (Table 5, row 9), and an increase of 1.75 LAYS per 100 USD (Table 5, row 10); in the *half-years* scenario, the average test score gain is reduced to $.015\sigma$ (Table 5, row 8), corresponding to 0.019 LAYS per student year (Table 5, row 9), and 0.87 LAYS per 100 USD invested (Table 5, row 10).

#### 4.4.4 Cost-Effectiveness in Pandemic Years

Pandemic-related learning losses in Latin America and the Caribbean resulted in an average loss of 1.8 LAYS (Table 5, row 11) (Azevedo et al., 2022). The annual probability of a pandemic half as intense as COVID-19 is 1.54% (Based on Marani et al. (2021, 2023); Glennerster et al. (2023), Table 5, row 12). Thus, the annual expected pandemic LAYS loss per student year due to school closures is 0.028 (Table 5, row 13). Assuming a 25% probability that filters avert school closures in a pandemic, in a pandemic year, filters in the *all-years* scenario yield an increase of 2.07 LAYS per 100 USD (Table 5, row 14). In the *half-years* scenario, the estimated gain falls to 1.20 LAYS per 100 USD (Table 5, row 14).

[25] The pads need to be replaced roughly every 9 months, which matches the length of the academic year.

#### 4.4.5 Cost-Effectiveness in Non-Pandemic and Pandemic Years in Expectation

When accounting for benefits from non-pandemic learning gains and averted pandemic school closures, filters yield an expected return of 1.76 LAYS per 100 USD per student-year under the *all-years* scenario (Table 5, row 15). This places them above the average cost-effectiveness of educational interventions classified as "good buys" by the Global Education Evidence Advisory Panel (GEEAP) Report, 2023 (Angrist et al., 2025; GEEAP, 2023). However, even assuming a high probability that filters avert school closures during pandemics, the LAYS per 100 USD remains below the threshold for the most highly cost-effective interventions or "great buys". Under the *half-years* scenario, the expected LAYS per 100 USD falls to 0.88 (Table 5, row 15) — still above the median for "good buys," but considerably lower, and less clearly favorable by comparison.

#### 4.4.6 Benefit-Cost Ratio in Non-Pandemic Scenario

An alternative method for evaluating whether scaling up air filters in schools is worthwhile is to compare the dollar figures for benefits and costs. When calculating benefits from learning gains in non-pandemic years, the net present value of the estimated wage increment is computed by leveraging the correlation between an increase in Saber 11 scores and earnings, as estimated using a Mincerean equation by McLean et al. (2017) for college graduates. Bogotá's average annual wage for a population comparable to the one in this study is approximately 5,251 USD (Table 5, row 17). The share of the population with tertiary education attainment is 28% (OECD (2023); Table 5, row 18). According to 2023 Saber 11 data, a one-decile increase in test scores requires a 0.39 standard deviation gain, corresponding to a 3.3% rise in daily income (McLean et al. (2017), Table 7, Column 2). Assuming that the wage increase is linear for smaller test score gains, air filters increase lifetime earnings by 0.26% for each student served who attends tertiary education under the full-year scenario (Table 5, row 19). Under the *half-years* scenario, the corresponding increase in lifetime earnings is 0.13% (Table 5, row 19). We conservatively assume zero return for students served who do not pursue tertiary education. This yields a lifetime wage increment ranging from 43.10 USD to 86.19 USD per student-year of filter use (Table 5, Row 20), resulting in a Benefit-Cost Ratio ranging from 20.08 to 40.16 (Table 5, Row 21).

Table 5 also determines the fraction of the test score-wage relationship that must be causal to justify the investment, assuming no pandemic effect. If the 2023 effect holds in all non-pandemic years, even if only 2.49% of the correlation between test scores and wages in Colombia is causal, these gains would still be sufficient to cover the cost of filters; under the *half-years* scenario, the break even threshold rises to 5% (Table 5, row 22).

#### 4.4.7 Benefit-Cost Ratio in Pandemic Scenario

Filters may also provide net benefits by averting school closures during pandemics, thereby preventing learning losses and corresponding wage losses. The effect of lost learning from pandemics is estimated based on (Azevedo et al., 2022), who assumes that the Psacharopoulos-style Mincerian return to education reflects only human capital, excluding signaling or rationing effects related to access to high-return tertiary education. This differs from the non-pandemic analysis, which allows for the possibility that only a small share of the wage effect of test scores is due to human capital. Based on this approach, the net present value of earnings lost due to COVID-19 school closures is $2,014 per student-year (Azevedo et al. (2022) Table D7, Intermediate scenario for LAC, and Table 5, row 23). The corresponding annual expected pandemic earnings loss per student-year is $31.08 (Table 5, row 24). Therefore, filters are cost-effective if they reduce the probability of pandemic school closures by at least 7% (Table 5, row 25)

#### 4.4.8 Benefit-Cost Ratio and Policymakers' Optimal Decisions

The policymaker's optimal decision depends on beliefs about the extent to which test scores reflect human capital, signaling, or access to tertiary education, and how these mechanisms influence expected wage returns. Expanding air filter installation is optimal in this setting if at least 5% of the test score-wage correlation reflects human capital accumulation, if filters prevent school closures with at least 7% probability during a future pandemic, or for any linear combination of these effects.

Table 5: Cost-Effectiveness and Benefit-Cost Ratio analysis under various assumptions

| Filter effectiveness scenarios | | 2023 effect in all non-pandemic years (1) | 2023 effect in half of non-pandemic years (2) | |
|---|---|---|---|---|
| Row # | Parameters | Estimates | | Source |
| | | A. Costs | | |
| 1 | Filter cost. | $288 | | Section 3.3.1. |
| 2 | Years before filter replacement. | 8 | | The manufacturer suggests filters last a minimum of 8 years. |
| 3 | Filter pad cost (annual, excluding the first year). | $54 | | Section 3.3.1. |
| 4 | Electricity cost (annual, per filter).[a] | $11 | | Assumes 8 hours of usage per day, 20 days per month, over 9 months of schooling per year. |
| 5 | Students served per filter. | 41 | | Section 3.3.1. |
| 6 | Discount rate. | 3% | | (Christiano et al., 2005; Lopez & Prada, 2009). |
| 7 | **Cost (per student year).**[b] | **$2.15** | | (NPV of Row(1, 3, 4) over Row(2) years)/(Row (2) * Row (5)). |
| | | B. Cost-effectiveness in LAYS per $100 | | |
| | **Non-Pandemic Scenario** | | | |

| | | | | |
|---|---|---|---|---|
| 8 | Average test scores increase (in SD) per student year. | 0.03 | 0.015 | (Table 2, Panel A) |
| 9 | LAYS gain per student year.[c] | 0.038 | 0.019 | Row (8) / 0.8. |
| 10 | **LAYS per 100 USD per student year.** | **1.75** | **0.87** | Row (9) * 100 / Row (7) |
| | **Pandemic Scenario** | | | |
| 11 | LAYS loss due to Covid-19 school closures (per student year). | 1.8 | | Based on Azevedo et al. (2022). Table D4 Intermediate scenario for LAC. |
| 12 | Annual probability of pandemic half as intense as COVID-19. | 1.54% | | Based on Marani et al. (2021, revised 2023) and Glennerster et al. (2023). |
| 13 | Annual expected pandemic LAYS loss due to school closures (per student year). | 0.028 | | Row (11) * Row (12). |
| 14 | LAYS per 100 USD per student year assuming a 25% probability that filters avert school closure conditional on pandemics. | 2.07 | 1.20 | (Row (9) + (0.25 * Row (13))) * 100 / Row (7) |
| 15 | **Expected LAYS per 100 USD per student year.** | **1.76** | **0.88** | (Row (10) * (1-Row(12)) + Row(14)*Row(12)) |
| 16 | Cost-effectiveness characterization according to the GEEAP smart buys report.[d] | Good Buy | Good Buy | |
| | | C. Benefit-Cost Ratio | | |
| | **Non-Pandemic Scenario** | | | |
| 17 | Average annual wage (Bogotá). | $5,251.33 | | Colombian Household Survey (GEIH). |
| 18 | Share of population with tertiary education attainment. | 28% | | OECD (2023) |
| 19 | Percentage increase in yearly earnings associated with 0.03 SD increase in Saber 11 score.[e] | 0.26% | 0.13% | Based on Bentley MacLeod et al. (2017) (Table 7, Column 2). |
| 20 | Estimated wage increment per student year. [f] | $86.19 | $43.10 | NPV of Row (17) * Row (18) * Row (19) over 37 years. |
| 21 | Benefit-Cost Ratio (based on wage increment only). | 40.16 | 20.08 | Row (20) / Row (7). |
| 22 | **Fraction of test score-wage relationship representing true increase in human capital for cost-effectiveness (no pandemic effect)** | **2.49%** | **4.98%** | 1 / (Row (22)). |
| | **Pandemic Scenario** | | | |
| 23 | NPV of earnings loss due to COVID-19 school closures (per student year).[g] | $2,014 | | Based on Azevedo et al. (2022). Table D7 Intermediate scenario for LAC. |
| 24 | Annual Expected pandemic earnings loss due to school closures (per student year). | $31.02 | | Row (12) * Row (23). |
| 25 | **Cutoff probability that filters avert school closure for E[benefits] >Costs (no additional learning gains)** | **6.91%** | | Row (7) / Row (24). |

*Notes:* This table reports the LAYS generated per 100 USD in expenditure and benefit-cost ratio of air filter investments, conditional on various assumptions. Each includes benefits from wage increases due to learning gains in non-pandemic years and from averted school closures in pandemic years. Column (1) reports calculations assuming the 2023 impact of filters on test scores in non-pandemic years, while column (2) reports calculations assuming the 2023 impact of filters on test scores in half of non-pandemic years and no impact in the other half of non-pandemic years. Panel A presents costs. Panel B reports LAYS gained per $100 in expenditure, including human capital benefits from learning gains in non-pandemic years and benefits from averting school closures in pandemic years. Panel C compares benefits and costs in non-pandemic years, while Panel D focuses on pandemic years. When accounting for benefits from non-pandemic learning gains and averted pandemic school closures, if filters perform as observed in the current study in all non-pandemic years, they would be comparable to other "effective but relatively expensive" interventions listed in the Global Education Evidence Advisory Panel Report, 2023. Even assuming the maximum probability that filters avert school closures during pandemics, the LAYS per $100 remains below the threshold for cost-effective interventions or "good buys" (see footnote d for more detail on GEEAP 2023). However, benefits from filters would exceed costs if 2.5% of the test score-wage correlation reflects human capital accumulation, if filters prevent school closures with 7% probability during a future pandemic, or for any linear combination of these effects.
[a] Electricity consumption is based on smart plug measurements and manufacturer specifications. The cost per kWh (COP 806.57) is taken from Bogotá Electric Power Company (ENEL) tariff schedule for non-residential, low-voltage Level 1 customers-applicable to public schools in Bogotá.
[b] The net present value of installing and maintaining a single filter includes a one-time purchase cost (Row 1), an annual electricity cost (Row 4), and an annual maintenance cost (excluding the first year) (Row 3) over the filter's lifespan (Row 2).
[c] Learning-Adjusted Years of Schooling is computed following Angrist et al. (2025). The estimated yearly learning gain for Singapore ($0.80\sigma$) is used as a high-performance benchmark to adjust for education quality in Colombia.
[d] GEEAP's Smart Buys report ranks education interventions in low- and middle-income countries into five categories: great buys, good buys, promising but limited evidence, effective but relatively expensive, and bad buys. For good buys, the average number of LAYS per $100 is approximately 1.3, with a median of 0.6. In contrast, for great buys, the average is about 27 LAYS per $100, with a median of 3.5.
[e] From 2023 Saber 11 data, a one-decile increase in test scores requires a 0.39 standard deviation gain and corresponds to a 3.3% rise in daily income (Bentley MacLeod et al., 2017: Table 7, Column 2). Bentley MacLeod et al. (2017) attempt to isolate the return to test scores from the return to tertiary initiation or reputation, estimating the parameters only for college graduates and excluding effects from tertiary education rationing. The calculated annual wage gain for students exposed to filters assumes similar wage increases for smaller learning gains.
[f] The net present value of the earnings increase resulting from a 0.03 SD increase in the Saber 11 score for college graduates is calculated directly from Row (19) parameter and assumes that individuals enter the workforce at an average age of 26 and retire at 62, in accordance with Colombian law. Since the correlation between test-score and wages is computed only for college graduates, we conservatively assume zero return for students who don't pursue tertiary education.
[g] The net present value of earnings loss, due to Covid-19 school closures (per student year) based on Azevedo 2022 Table D7 Intermediate scenario for LAC.

# 5 Conclusions and discussion

This paper finds evidence that installing high-efficiency particulate (HEPA) air filters in classrooms in Bogotá, Colombia, increased learning levels even after a short period of exposure. In schools randomly assigned to receive filters 3 - 4 months before a high-stakes exam, conducted outside school premises, students scored .03 standard deviations higher. Data from air pollution monitors suggests that students in schools with filters were exposed to lower levels of $PM_{2.5}$ (a reduction of .47 $\mu g/m^3$ from a base of 11.00 $\mu g/m^3$). This paper also suggests that, when accounting for benefits from both non-pandemic learning gains and averted pandemic school closures, air filters would be comparable to other cost-effective educational interventions.

These results are specific to a particular context, and one cannot rule out sampling variation; however, at a minimum, they suggest that further exploration of the learning impact of filters would be warranted. The small absolute impacts highlight the advantages of conducting large-scale studies in contexts with good administrative data. Future research could examine more advanced filter systems, could identify and measure impacts on subpopulations at particular risk, such as students with asthma, and could directly measure impacts on infectious disease transmission within classrooms, which could shed light on mechanisms, and thus on whether air filters might deliver learning benefits in settings with much lower air pollution. Ideally, this research would be conducted in a range of different environments with varying pollution levels to evaluate the differences in impact.

# References


Abadie, A. (2002). Bootstrap tests for distributional treatment effects in instrumental variable models. *Journal of the American Statistical Association*, *97*(457), 284–292. Retrieved from http://www.jstor.org/stable/3085782

Abdulkadiroğlu, A., Pathak, P. A., & Walters, C. R. (2018, January). Free to choose: Can school choice reduce student achievement? *American Economic Journal: Applied Economics*, *10*(1), 175-206.

Aguilar-Gomez, S., Dwyer, H., Graff Zivin, J., & Neidell, M. (2022). This is air: The "nonhealth"' effects of air pollution [Journal Article]. *Annual Review of Resource Economics*, *14*(Volume 14, 2022), 403-425. Retrieved from https://www.annualreviews.org/content/journals/10.1146/annurev-resource-111820-021816 doi: https://doi.org/10.1146/annurev-resource-111820-021816

Alasino, E., Ramírez, M. J., Romero, M., Schady, N., & Uribe, D. (2024). Learning losses during the COVID-19 pandemic: Evidence from Mexico. *Economics of Education Review*, *98*, 102492.

Angrist, N., Evans, D. K., Filmer, D., Glennerster, R., Rogers, H., & Sabarwal, S. (2025). How to improve education outcomes most efficiently? a review of the evidence using a unified metric. *Journal of Development Economics*, *172*, 103382. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304387824001317 doi: https://doi.org/10.1016/j.jdeveco.2024.103382

Araújo, D., Carrillo, B., & Sampaio, B. (2021). The long-run economic consequences of

iodine supplementation. *Journal of Health Economics*, *79*, 102490. Retrieved from https://www.sciencedirect.com/science/article/pii/S0167629621000758 doi: https://doi.org/10.1016/j.jhealeco.2021.102490
Arceo, E., Hanna, R., & Oliva, P. (2016, 01). Does the effect of pollution on infant mortality differ between developing and developed countries? Evidence from Mexico City. *The Economic Journal*, *126*(591), 257-280. Retrieved from https://doi.org/10.1111/ecoj.12273 doi: 10.1111/ecoj.12273
Ardington, C., Wills, G., & Kotze, J. (2021). COVID-19 learning losses: Early grade reading in South Africa. *International Journal of Educational Development*, *86*, 102480. Retrieved from https://www.sciencedirect.com/science/article/pii/S0738059321001334 doi: https://doi.org/10.1016/j.ijedudev.2021.102480
Azevedo, J. P., Akmal, M., Cloutier, M.-H., Rogers, H., & Wong, Y. N. (2022). Learning losses during COVID-19. *Policy Research Working Paper*(10218).
Azevedo, J. P., Cojocaru, A., Talledo, V. M., & Narayan, A. (2023, March). *COVID-19 school closures, learning losses and intergenerational mobility* (Policy Research Working Paper No. 10381). Washington, DC: World Bank. Retrieved from https://documents.worldbank.org/curated/en/991411638314548186/COVID-19-School-Closures-Learning-Losses-and-Intergenerational-Mobility
Azevedo, J. P., Hasan, A., Goldemberg, D., Geven, K., & Iqbal, S. A. (2021). Simulating the potential impacts of COVID-19 school closures on schooling and learning outcomes: A set of global estimates. *The World Bank Research Observer*, *36*(1), 1–40. Retrieved from https://academic.oup.com/wbro/article/36/1/1/6174606 doi: 10.1093/wbro/lkab003
Azor-Martínez, E., Gonzalez-Jimenez, Y., Seijas-Vazquez, M. L., Cobos-Carrascosa, E., Santisteban-Martínez, J., Martínez-López, J. M., ... others (2014). The impact of common infections on school absenteeism during an academic year. *American journal of infection control*, *42*(6), 632–637.
Balakrishnan, U., & Tsaneva, M. (2021). Air pollution and academic performance: Evidence from India. *World Development*, *146*, 105553. doi: https://doi.org/10.1016/j.worlddev.2021.105553
Betthäuser, B. A., Bach-Mortensen, A. M., & Engzell, P. (2023). A systematic review and meta-analysis of the evidence on learning during the COVID-19 pandemic. *Nature Human Behaviour*, *7*(3), 375–385.
Biasi, B., Lafortune, J., & Schonholzer, D. (2025, 02). What works and for whom? effectiveness and efficiency of school capital investments across the U.S. *The Quarterly Journal of Economics*, qjaf013. Retrieved from https://doi.org/10.1093/qje/qjaf013 doi: 10.1093/qje/qjaf013
Burgmann, S., & Janoske, U. (2021). Transmission and reduction of aerosols in classrooms using air purifier systems. *Physics of fluids*, *33*(3).
Castillo-Rodríguez, L., Malo-Sánchez, D., Díaz-Jiménez, D., García-Velásquez, I., Pulido, P., & Castañeda-Orjuela, C. (2022). Economic costs of severe seasonal influenza in Colombia, 2017–2019: A multi-center analysis. *PLoS One*, *17*(6), e0270086.
Cattaneo, M. D., Crump, R. K., Farrell, M. H., & Feng, Y. (2024, May). On binscatter. *American Economic Review*, *114*(5), 1488-1514. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.20221576 doi: 10.1257/aer.20221576
Chen, J.-C., Wang, X., Serre, M., Cen, S., Franklin, M., & Espeland, M. (2017). Particulate

air pollutants, brain structure, and neurocognitive disorders in older women. *Research Reports: Health Effects Institute*, *2017*.
Christiano, L. J., Eichenbaum, M., & Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of political Economy*, *113*(1), 1–45.
Consumer Reports. (2019). *Air purifiers and the cost of clean air.* Retrieved 2025-05-05, from https://www.consumerreports.org/appliances/air-purifiers/air-purifiers-and-the-cost-of-clean-air-a6152505326/ (Accessed 5 May 2025)
Curtius, J., Granzin, M., & Schrod, J. (2021). Testing mobile air purifiers in a school classroom: Reducing the airborne transmission risk for SARS-CoV-2. *Aerosol Science and Technology*, *55*(5), 586–599.
Duque, V., & Gilraine, M. (2022). Coal use, air pollution, and student performance. *Journal of Public Economics*, *213*, 104712. Retrieved from https://www.sciencedirect.com/science/article/pii/S0047272722001141 doi: https://doi.org/10.1016/j.jpubeco.2022.104712
Ebenstein, A., Lavy, V., & Roth, S. (2016, 10). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. *American Economic Journal: Applied Economics*, *8*(4), 36-65. doi: 10.1257/app.20150213
Environmental Protection Agency, U. S. (2023). *Particulate matter (pm) basics.* Web.
Fernando, D., De Silva, D., Carter, R., Mendis, K. N., & Wickremasinghe, R. (2006). A randomized, double-blind, placebo-controlled, clinical trial of the impact of malaria prevention on the educational attainment of school children. *American Journal of Tropical Medicine and Hygiene*, *74*(3), 386–393.
Field, E., Robles, O., & Torero, M. (2009, October). Iodine deficiency and schooling attainment in Tanzania. *American Economic Journal: Applied Economics*, *1*(4), 140-69. Retrieved from https://www.aeaweb.org/articles?id=10.1257/app.1.4.140 doi: 10.1257/app.1.4.140
Firpo, S., Fortin, N. M., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica*, *77*(3), 953–973. Retrieved 2024-06-10, from http://www.jstor.org/stable/40263848
GEEAP, G. E. E. A. P. (2023, May). *Cost-effective approaches to improve global learning: Smart buys for improving learning in low- and middle-income countries* (Tech. Rep.). London, Washington D.C., New York: FCDO, World Bank, UNICEF, USAID. Retrieved from https://documents.worldbank.org/curated/en/495041468224995463/pdf/WPS4998.pdf
Gilraine, M. (2023). Air filters, pollution, and student achievement. *Journal of Human Resources*, 0421-11642R2. doi: 10.3368/jhr.0421-11642R2
Gilraine, M., & Zheng, A. (2024). Air pollution and student performance in the U.S. *Journal of Urban Economics*, *143*, 103686. Retrieved from https://www.sciencedirect.com/science/article/pii/S0094119024000561 doi: https://doi.org/10.1016/j.jue.2024.103686
Glennerster, R., Snyder, C. M., & Tan, B. J. (2023, June). *Calculating the costs and benefits of advance preparations for future pandemics* (Working Paper No. 30565). Cambridge, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w30565 (Revised version of October 2022 paper)
Glewwe, P., & Miguel, E. A. (2007). Chapter 56 - the impact of child health and nutrition on education in less developed countries. In T. P. Schultz & J. A. Strauss (Eds.), (Vol. 4,

p. 3561-3606). Elsevier. Retrieved from https://www.sciencedirect.com/science/article/pii/S1573447107040569 doi: https://doi.org/10.1016/S1573-4471(07)04056-9

Glewwe, P., Park, A., & Zhao, M. (2016). A better vision for development: Eyeglasses and academic performance in rural primary schools in China. *Journal of Development Economics*, *122*, 170–182.

Greenstone, M., Hasenkopf, C., & Lee, K. (2022, June). *Air quality life index.*

Greenstone, M., Lee, K., & Sahai, H. (2021, May). Indoor air quality, information, and socioeconomic status: Evidence from Delhi. *AEA Papers and Proceedings*, *111*, 420-24. Retrieved from https://www.aeaweb.org/articles?id=10.1257/pandp.20211006 doi: 10.1257/pandp.20211006

Guariso, A., & Björkman Nyqvist, M. (2023). The impact of the COVID-19 pandemic on children's learning and wellbeing: Evidence from India. *Journal of Development Economics*, *164*, 103133. Retrieved from https://www.sciencedirect.com/science/article/pii/S0304387823000883 doi: https://doi.org/10.1016/j.jdeveco.2023.103133

Hammond, A., Khalid, T., Thornton, H. V., Woodall, C. A., & Hay, A. D. (2021). Should homes and workplaces purchase portable air filters to reduce the transmission of SARS-CoV-2 and other respiratory infections? a systematic review. *PLoS One*, *16*(4), e0251049.

Heissel, J. A., Persico, C., & Simon, D. (2020). Does pollution drive achievement? the effect of traffic pollution on academic performance. *Journal of Human Resources*. Retrieved from https://jhr.uwpress.org/content/early/2020/04/03/jhr.57.3.1218-9903R2 doi: 10.3368/jhr.57.3.1218-9903R2

IQAir. (2023). *World air quality report 2023.* Retrieved from https://www.iqair.com/world-air-quality-report (Accessed: 2024-08-14)

Künn, S., Palacios, J., & Pestel, N. (2023). Indoor air quality and strategic decision making. *Management Science*, *69*(9), 5354–5377.

Lichand, G., Doria, C. A., Leal-Neto, O., & Fernandes, J. P. C. (2022). The impacts of remote learning in secondary education during the pandemic in Brazil. *Nature Human Behaviour*, 1–8.

Lin, W. (2013). Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique. *The Annals of Applied Statistics*, *7*(1), 295 – 318. Retrieved from https://doi.org/10.1214/12-AOAS583 doi: 10.1214/12-AOAS583

Lindsley, W. G. (2021). Efficacy of portable air cleaners and masking for reducing indoor exposure to simulated exhaled SARS-CoV-2 aerosols - United States, 2021. *MMWR. Morbidity and mortality weekly report*, *70*.

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

López, M. R., Prada, J. D., & Rodríguez N., N. (2008). Financial accelerator mechanism in a small open economy. *Borradores de Economía*(525).

Luo, R., Shi, Y., Zhang, L., Liu, C., Rozelle, S., Sharbono, B., ... Martorell, R. (2012). Nutrition and educational performance in rural China's elementary schools: Results of a randomized control trial in Shaanxi province. *Economic development and cultural change*, *60*(4), 735–772.

Marani, M., Katul, G. G., Pan, W. K., & Parolari, A. J. (2021). Intensity and frequency

of extreme novel epidemics. *Proceedings of the National Academy of Sciences*, *118*(35), e2105482118.
Marani, M., Katul, G. G., Pan, W. K., & Parolari, A. J. (2023). Correction for marani et al., intensity and frequency of extreme novel epidemics. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(12), e2302169120. (Epub 2023 Mar 14. Erratum for: Proc Natl Acad Sci U S A. 2021 Aug 31;118(35):e2105482118.) doi: 10.1073/pnas.2302169120
McKenzie, D. (2012). Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*, *99*(2), 210-221. Retrieved from https://www.sciencedirect.com/science/article/pii/S030438781200003X doi: https://doi.org/10.1016/j.jdeveco.2012.01.002
McLean, H. Q., Peterson, S. H., King, J. P., Meece, J. K., & Belongia, E. A. (2017). School absenteeism among school-aged children with medically attended acute viral respiratory illness during three influenza seasons, 2012-2013 through 2014-2015. *Influenza and other respiratory viruses*, *11*(3), 220–229.
Merkus, E. (2024). Hold your breath! air pollution and cognitive performance in Colombia. *World Development Sustainability*, *4*, 100123. Retrieved from https://www.sciencedirect.com/science/article/pii/S2772655X24000016 doi: https://doi.org/10.1016/j.wds.2024.100123
Metcalfe, R. D., & Roth, S. (2025, February). *Making the invisible visible: The impact of revealing indoor air pollution on behavior and welfare* (Working Paper No. 33510). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w33510 doi: 10.3386/w33510
Miguel, E., & Kremer, M. (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica*, *72*(1), 159–217.
Moscoviz, L., & Evans, D. K. (2022). Learning loss and student dropouts during the COVID-19 pandemic: A review of the evidence two years after schools shut down. *Center for Global Development, Working Paper*, *609*.
Mura, I., Franco, J. F., Bernal, L., Melo, N., Díaz, J. J., & Akhavan-Tabatabaei, R. (2020). A decade of air quality in Bogotá: A descriptive analysis. *Frontiers in Environmental Science*, *8*. doi: 10.3389/fenvs.2020.00065
Neuzil, K. M., Hohlbein, C., & Zhu, Y. (2002). Illness among schoolchildren during influenza season: effect on school absenteeism, parental absenteeism from work, and secondary illness in families. *Archives of pediatrics & adolescent medicine*, *156*(10), 986–991.
OECD. (2023). Education at a glance 2023: Colombia - country note. *Education at a Glance 2023*. (Available at https://doi.org/10.1787/d7f76adc-en)
Patrinos, H., Vegas, E., & Carter-Rau, R. (2023, 08). *An analysis of COVID-19 student learning loss.* Oxford University Press. Retrieved from https://oxfordre.com/economics/view/10.1093/acrefore/9780190625979.001.0001/acrefore-9780190625979-e-893 doi: 10.1093/acrefore/9780190625979.013.893
Persico, C. L., & Venator, J. (2019). The effects of local industrial pollution on students and schools. *Journal of Human Resources*. Retrieved from https://jhr.uwpress.org/content/early/2019/08/02/jhr.56.2.0518-9511R2 doi: 10.3368/jhr.56.2.0518-9511R2
Rios-Avila, F. (2020). Recentered influence functions (RIFs) in stata: RIF regression and

RIF decomposition. *The Stata Journal*, *20*(1), 51-94. Retrieved from https://doi.org/10.1177/1536867X20909690 doi: 10.1177/1536867X20909690

Rosenzweig, M. R., & Udry, C. (2020). External validity in a stochastic world: Evidence from low-income countries. *The Review of Economic Studies*, *87*(1), 343–381.

Roth, J., & Sant'Anna, P. H. C. (2023). Efficient estimation for staggered rollout designs. *Journal of Political Economy Microeconomics*, *1*(4), 669-709. Retrieved from https://doi.org/10.1086/726581 doi: 10.1086/726581

Roth, S. (2020). The effect of indoor air pollution on cognitive performance: Evidence from the UK. *Manuscript, London School of Economics https://personal. lse. ac. uk/roths/JMP. pdf*.

Schady, N., Holla, A., Sabarwal, S., Silva, J., & Chang, A. Y. (2023). *Collapse and recovery: How the COVID-19 pandemic eroded human capital and what to do about it*. Washington, DC: World Bank. Retrieved from https://openknowledge.worldbank.org/entities/publication/94cd910e-033e-42f4-9e8a-92a0bc316bbb (Accessed: 2025-03-11)

Silverman, B. W. (2018). *Density estimation for statistics and data analysis*. Routledge.

Simona, S. C., Bartell, S. M., & Vieira, V. M. (2025). Classroom air quality in a randomized crossover trial with portable HEPA air cleaners. *Journal of Exposure Science & Environmental Epidemiology*. Retrieved from https://doi.org/10.1038/s41370-025-00743-9 doi: 10.1038/s41370-025-00743-9

Singh, A., Romero, M., & Muralidharan, K. (2024). COVID-19 learning loss and recovery. *Journal of Human Resources*. Retrieved from https://jhr.uwpress.org/content/early/2024/04/01/jhr.0723-13025R2 doi: 10.3368/jhr.0723-13025R2

Stafford, T. M. (2015). Indoor air quality and academic performance. *Journal of Environmental Economics and Management*, *70*, 34-50. Retrieved from https://www.sciencedirect.com/science/article/pii/S0095069614001016 doi: https://doi.org/10.1016/j.jeem.2014.11.002

Sylvia, S., Luo, R., Zhang, L., Shi, Y., Medina, A., & Rozelle, S. (2013). Do you get what you pay for with school-based health programs? evidence from a child nutrition experiment in rural China. *Economics of Education Review*, *37*, 1–12.

Ueki, H., Ujie, M., Komori, Y., Kato, T., Imai, M., & Kawaoka, Y. (2022). Effectiveness of HEPA filters at removing infectious SARS-CoV-2 from the air. *Msphere*, *7*(4), e00086–22.

UNESCO. (2020). *Global monitoring of school closures caused by COVID-19.* Retrieved from https://covid19.uis.unesco.org/global-monitoring-school-closures-covid19/ (Accessed: 2024-12-06)

Vargas, S. L., Matiz, M. I., Olano, V. A., Sarmiento-Senior, D., Jaramillo, J. F., Neal, A., ... Overgaard, H. J. (2020). Razones de ausentismo en escuelas primarias rurales de dos municipios colombianos. *Revista Salud Bosque*(1).

Villalobos, L., & Blackman, A. (2025). Air pollution and education in the global south: Fine particulates reduce high-stakes test scores in Bogotá. *Applied Economics Letters*, *0*(0), 1–5. Retrieved from https://doi.org/10.1080/13504851.2025.2459883 doi: 10.1080/13504851.2025.2459883

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

World Health Organization. (2022). *Ambient (outdoor) air pollution.*

Zhang, X., Chen, X., & Zhang, X. (2018). The impact of exposure to air pollution on cognitive performance. *Proceedings of the National Academy of Sciences*, *115*(37), 9193–9197.

# A Additional tables and figures

## A.1 Tables

Table A.1: Student-level balance (2022)

| | Mean/SD for each group | | | | Treatment effect (vs control) | | | Comparing treatments | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Control (C) | Filter (F) | Monitor (M) | Both (B) | F vs C (2)vs(1) | M vs C (3)vs(1) | B vs C (4)vs(1) | F vs M (3)vs(2) | F vs B (4)vs(2) | M vs B (4)vs(3) |
| Global score | -0.01 | 0.04 | -0.02 | 0.00 | 0.06 | -0.04 | -0.00 | -0.10** | -0.06 | 0.04 |
| | (1.01) | (0.99) | (1.00) | (1.00) | (0.04) | (0.04) | (0.05) | (0.05) | (0.05) | (0.05) |
| | [15,323] | [9,539] | [7,787] | [8,386] | [24,862] | [23,110] | [23,709] | [17,326] | [17,925] | [16,173] |
| Reading score | -0.01 | 0.03 | -0.02 | 0.00 | 0.04 | -0.03 | -0.00 | -0.08* | -0.05 | 0.03 |
| | (1.01) | (0.99) | (1.00) | (1.00) | (0.04) | (0.04) | (0.04) | (0.04) | (0.04) | (0.05) |
| | [15,323] | [9,539] | [7,787] | [8,386] | [24,862] | [23,110] | [23,709] | [17,326] | [17,925] | [16,173] |
| Math score | -0.01 | 0.05 | -0.05 | 0.00 | 0.07 | -0.07 | 0.00 | -0.13*** | -0.06 | 0.07 |
| | (1.01) | (1.00) | (0.98) | (0.99) | (0.04) | (0.04) | (0.04) | (0.05) | (0.05) | (0.05) |
| | [15,323] | [9,539] | [7,787] | [8,386] | [24,862] | [23,110] | [23,709] | [17,326] | [17,925] | [16,173] |
| SES strata 1 or 2 | 0.63 | 0.68 | 0.60 | 0.64 | 0.03 | -0.01 | 0.02 | -0.05* | -0.02 | 0.03 |
| | (0.48) | (0.47) | (0.49) | (0.48) | (0.02) | (0.02) | (0.02) | (0.03) | (0.03) | (0.03) |
| | [15,323] | [9,539] | [7,787] | [8,386] | [24,862] | [23,110] | [23,709] | [17,326] | [17,925] | [16,173] |
| SES strata 3 or 4 | 0.30 | 0.27 | 0.34 | 0.30 | -0.02 | 0.02 | -0.01 | 0.04 | 0.02 | -0.03 |
| | (0.46) | (0.44) | (0.47) | (0.46) | (0.02) | (0.02) | (0.02) | (0.03) | (0.03) | (0.03) |
| | [15,323] | [9,539] | [7,787] | [8,386] | [24,862] | [23,110] | [23,709] | [17,326] | [17,925] | [16,173] |
| Female | 0.55 | 0.53 | 0.56 | 0.53 | -0.03* | 0.00 | -0.02* | 0.03* | 0.00 | -0.03* |
| | (0.50) | (0.50) | (0.50) | (0.50) | (0.01) | (0.02) | (0.01) | (0.02) | (0.01) | (0.02) |
| | [15,323] | [9,539] | [7,787] | [8,386] | [24,862] | [23,110] | [23,709] | [17,326] | [17,925] | [16,173] |
| Age (exam date) | 17.62 | 17.48 | 17.62 | 17.58 | -0.14** | 0.01 | -0.04 | 0.15* | 0.10 | -0.04 |
| | (1.59) | (1.30) | (1.52) | (1.42) | (0.07) | (0.08) | (0.08) | (0.08) | (0.08) | (0.09) |
| | [15,107] | [9,438] | [7,714] | [8,295] | [24,545] | [22,821] | [23,402] | [17,152] | [17,733] | [16,009] |
| Mother some secondary education | 0.48 | 0.50 | 0.49 | 0.49 | 0.01 | 0.01 | 0.01 | 0.00 | 0.00 | 0.00 |
| | (0.50) | (0.50) | (0.50) | (0.50) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| | [15,323] | [9,539] | [7,787] | [8,386] | [24,862] | [23,110] | [23,709] | [17,326] | [17,925] | [16,173] |

*Notes*: Columns 1–4 show the mean, the standard deviation (in parentheses), and the number of observations (in square brackets) for each of the experimental groups: Control, Filters (F), Monitors (M), and Both (B). Columns 5–7 report the difference between the three treatment groups and the control, the standard error of the difference (in parentheses), and the number of observations used for the estimation (in square brackets). Columns 8–10 show the difference between treatment groups, the standard error of the difference (in parentheses), and the number of observations used for the estimation (in square brackets). All the differences between groups (Columns 4–10) consider the randomization design (i.e., include strata fixed effects). Standard errors are clustered at the school level. The p-value from a joint test that all coefficients in Columns 1–4 are zero (from a multinomial logit with the treatment groups on the left-hand side and all the covariates on the right-hand side) is .126. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table A.2: School-level balance (2022)

| | Mean/SD for each group | | | | Treatment effect (vs control) | | | Comparing treatments | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Control (C) (1) | Filter (F) (2) | Monitor (M) (3) | Both (B) (4) | F vs C (2)vs(1) (5) | M vs C (3)vs(1) (6) | B vs C (4)vs(1) (7) | F vs M (3)vs(2) (8) | F vs B (4)vs(2) (9) | M vs B (4)vs(3) (10) |
| PM2.5 | 17.54 | 17.56 | 17.45 | 17.53 | 0.02 | -0.06 | 0.05 | -0.08 | 0.03 | 0.11 |
| | (1.95) | (2.19) | (2.05) | (1.77) | (0.15) | (0.12) | (0.11) | (0.15) | (0.15) | (0.12) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |
| PM10 | 36.08 | 35.94 | 35.37 | 35.91 | -0.10 | -0.61* | -0.01 | -0.51 | 0.09 | 0.60* |
| | (5.49) | (5.72) | (5.81) | (5.52) | (0.34) | (0.32) | (0.33) | (0.35) | (0.36) | (0.34) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |
| O3 | 26.90 | 26.41 | 27.10 | 26.81 | -0.53 | 0.12 | -0.02 | 0.65 | 0.51 | -0.13 |
| | (4.51) | (3.93) | (3.93) | (3.88) | (0.36) | (0.38) | (0.39) | (0.40) | (0.41) | (0.42) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |
| Without Saber 11 data | 0.03 | 0.01 | 0.03 | 0.01 | -0.02 | -0.00 | -0.02 | 0.01 | -0.00 | -0.01 |
| | (0.17) | (0.12) | (0.16) | (0.12) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) | (0.02) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |
| Students (Grade 11) | 116.08 | 127.19 | 103.83 | 111.81 | 12.40 | -11.69 | -3.09 | -24.08* | -15.49 | 8.60 |
| | (73.32) | (90.82) | (60.08) | (66.06) | (11.84) | (9.27) | (9.53) | (12.28) | (12.43) | (10.14) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |

*Notes*: Columns 1–4 show the mean, the standard deviation (in parentheses), and the number of observations (in square brackets) for each of the experimental groups: Control, Filters (F), Monitors (M), and Both (B). Columns 5–7 report the difference between the three treatment groups and the control, the standard error of the difference (in parentheses), and the number of observations used for the estimation (in square brackets). Columns 8–10 show the difference between treatment groups, the standard error of the difference (in parentheses), and the number of observations used for the estimation (in square brackets). All the differences between groups (Columns 4–10) consider the randomization design (i.e., include strata fixed effects). Standard errors are clustered at the school level. The p-value from a joint test that all coefficients in Columns 1–4 are zero (from a multinomial logit with the treatment groups on the left-hand side and all the covariates on the right-hand side) is .376. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table A.3: Balance across experimental groups (2023)

| | Mean/SD for each group | | | | p-value |
|---|---|---|---|---|---|
| | Control | Filter | Monitor | Both | (equality) |
| | (1) | (2) | (3) | (4) | (5) |
| **Panel A: Student level (2023)** | | | | | |
| SES strata 1 or 2 | 0.60 | 0.63 | 0.57 | 0.60 | 0.432 |
| | (0.49) | (0.48) | (0.50) | (0.49) | |
| | [16,096] | [9,905] | [7,801] | [8,587] | |
| SES strata 3 or 4 | 0.30 | 0.29 | 0.35 | 0.31 | 0.643 |
| | (0.46) | (0.45) | (0.48) | (0.46) | |
| | [16,096] | [9,905] | [7,801] | [8,587] | |
| Female | 0.54 | 0.52 | 0.57 | 0.54 | 0.033** |
| | (0.50) | (0.50) | (0.50) | (0.50) | |
| | [16,096] | [9,905] | [7,801] | [8,587] | |
| Age (exam date) | 17.55 | 17.39 | 17.48 | 17.47 | 0.263 |
| | (1.69) | (1.28) | (1.38) | (1.39) | |
| | [15,737] | [9,782] | [7,703] | [8,466] | |
| Mother some secondary education | 0.48 | 0.48 | 0.47 | 0.49 | 0.202 |
| | (0.50) | (0.50) | (0.50) | (0.50) | |
| | [16,096] | [9,905] | [7,801] | [8,587] | |
| **Panel B: School level (2023)** | | | | | |
| PM2.5 | 15.78 | 15.68 | 15.67 | 15.77 | 0.569 |
| | (2.04) | (2.06) | (2.09) | (2.00) | |
| | [132] | [75] | [75] | [75] | |
| PM10 | 31.80 | 31.71 | 31.25 | 31.76 | 0.199 |
| | (4.87) | (5.00) | (4.99) | (4.96) | |
| | [132] | [75] | [75] | [75] | |
| O3 | 27.16 | 26.69 | 27.49 | 26.93 | 0.269 |
| | (4.73) | (4.69) | (4.57) | (4.72) | |
| | [132] | [75] | [75] | [75] | |
| Without Saber 11 data | 0.03 | 0.03 | 0.04 | 0.03 | 0.958 |
| | (0.17) | (0.16) | (0.20) | (0.16) | |
| | [132] | [75] | [75] | [75] | |
| Students (Grade 11) | 121.94 | 132.07 | 104.01 | 114.49 | 0.111 |
| | (79.39) | (93.24) | (63.59) | (69.33) | |
| | [132] | [75] | [75] | [75] | |

*Notes*: Columns 1-4 show the mean, standard deviation (in parentheses) and the number of observations (in square brackets) for each of the following groups: Control, schools where filters were installed, schools where air quality monitors were installed, and schools where both filters and air quality monitors were installed. Column 5 reports the p-value of whether the mean is the same across all groups. The equality test considers the randomization design (i.e., it includes strata fixed effects and clusters standard errors at the school level). All variables except age, number of students, pollution measures, and test scores are % with respect to the number of students in Grade 11. The p-value from a joint test that all coefficients are zero (from a multinomial logit with the treatment groups on the left-hand side and all the covariates on the right-hand side) is .354 for Panel A and .426 for Panel B. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.4: Student-level balance (2023)

| | Mean/SD for each group | | | | Treatment effect (vs control) | | | Comparing treatments | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Control (C) (1) | Filter (F) (2) | Monitor (M) (3) | Both (B) (4) | F vs C (2)vs(1) (5) | M vs C (3)vs(1) (6) | B vs C (4)vs(1) (7) | F vs M (3)vs(2) (8) | F vs B (4)vs(2) (9) | M vs B (4)vs(3) (10) |
| SES strata 1 or 2 | 0.60 | 0.63 | 0.57 | 0.60 | 0.02 | -0.02 | 0.01 | -0.04 | -0.01 | 0.03 |
| | (0.49) | (0.48) | (0.50) | (0.49) | (0.02) | (0.02) | (0.02) | (0.03) | (0.02) | (0.03) |
| | [16,096] | [9,905] | [7,801] | [8,587] | [26,001] | [23,897] | [24,683] | [17,706] | [18,492] | [16,388] |
| SES strata 3 or 4 | 0.30 | 0.29 | 0.35 | 0.31 | -0.01 | 0.02 | -0.01 | 0.03 | 0.00 | -0.03 |
| | (0.46) | (0.45) | (0.48) | (0.46) | (0.02) | (0.02) | (0.02) | (0.03) | (0.02) | (0.03) |
| | [16,096] | [9,905] | [7,801] | [8,587] | [26,001] | [23,897] | [24,683] | [17,706] | [18,492] | [16,388] |
| Female | 0.54 | 0.52 | 0.57 | 0.54 | -0.02 | 0.02 | -0.01 | $0.04^{***}$ | 0.01 | $-0.03^{*}$ |
| | (0.50) | (0.50) | (0.50) | (0.50) | (0.01) | (0.02) | (0.01) | (0.02) | (0.01) | (0.02) |
| | [16,096] | [9,905] | [7,801] | [8,587] | [26,001] | [23,897] | [24,683] | [17,706] | [18,492] | [16,388] |
| Age (exam date) | 17.55 | 17.39 | 17.48 | 17.47 | $-0.15^{*}$ | -0.06 | -0.07 | 0.10 | 0.08 | -0.02 |
| | (1.69) | (1.28) | (1.38) | (1.39) | (0.08) | (0.08) | (0.08) | (0.07) | (0.07) | (0.08) |
| | [15,737] | [9,782] | [7,703] | [8,466] | [25,519] | [23,440] | [24,203] | [17,485] | [18,248] | [16,169] |
| Mother some secondary education | 0.48 | 0.48 | 0.47 | 0.49 | 0.01 | -0.00 | $0.02^{*}$ | -0.01 | 0.01 | $0.02^{*}$ |
| | (0.50) | (0.50) | (0.50) | (0.50) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) | (0.01) |
| | [16,096] | [9,905] | [7,801] | [8,587] | [26,001] | [23,897] | [24,683] | [17,706] | [18,492] | [16,388] |

*Notes*: Columns 1–4 show the mean, the standard deviation (in parentheses), and the number of observations (in square brackets) for each of the experimental groups: Control, Filters (F), Monitors (M), and Both (B). Columns 5–7 report the difference between the three treatment groups and the control, the standard error of the difference (in parentheses), and the number of observations used for the estimation (in square brackets). Columns 8–10 show the difference between treatment groups, the standard error of the difference (in parentheses), and the number of observations used for the estimation (in square brackets). All the differences between groups (Columns 4–10) consider the randomization design (i.e., include strata fixed effects). Standard errors are clustered at the school level. The p-value from a joint test that all coefficients in Columns 1–4 are zero (from a multinomial logit with the treatment groups on the left-hand side and all the covariates on the right-hand side) is .354. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table A.5: School-level balance (2023)

| | Mean/SD for each group | | | | Treatment effect (vs control) | | | Comparing treatments | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Control (C) (1) | Filter (F) (2) | Monitor (M) (3) | Both (B) (4) | F vs C (2)vs(1) (5) | M vs C (3)vs(1) (6) | B vs C (4)vs(1) (7) | F vs M (3)vs(2) (8) | F vs B (4)vs(2) (9) | M vs B (4)vs(3) (10) |
| PM2.5 | 15.78 | 15.68 | 15.67 | 15.77 | -0.09 | -0.07 | 0.05 | 0.02 | 0.14 | 0.12 |
| | (2.04) | (2.06) | (2.09) | (2.00) | (0.11) | (0.10) | (0.11) | (0.11) | (0.12) | (0.10) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |
| PM10 | 31.80 | 31.71 | 31.25 | 31.76 | -0.04 | -0.49* | 0.14 | -0.44 | 0.18 | 0.62* |
| | (4.87) | (5.00) | (4.99) | (4.96) | (0.30) | (0.29) | (0.32) | (0.30) | (0.33) | (0.32) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |
| O3 | 27.16 | 26.69 | 27.49 | 26.93 | -0.50 | 0.14 | -0.16 | 0.64* | 0.34 | -0.30 |
| | (4.73) | (4.69) | (4.57) | (4.72) | (0.32) | (0.29) | (0.34) | (0.34) | (0.38) | (0.36) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |
| Without Saber 11 data | 0.03 | 0.03 | 0.04 | 0.03 | -0.00 | 0.01 | -0.00 | 0.01 | 0.00 | -0.01 |
| | (0.17) | (0.16) | (0.20) | (0.16) | (0.03) | (0.03) | (0.02) | (0.03) | (0.03) | (0.03) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |
| Students (Grade 11) | 121.94 | 132.07 | 104.01 | 114.49 | 11.15 | -17.34* | -6.32 | -28.49** | -17.47 | 11.02 |
| | (79.39) | (93.24) | (63.59) | (69.33) | (12.38) | (9.71) | (10.22) | (12.59) | (12.97) | (10.58) |
| | [132] | [75] | [75] | [75] | [207] | [207] | [207] | [150] | [150] | [150] |

*Notes*: Columns 1–4 show the mean, the standard deviation (in parentheses), and the number of observations (in square brackets) for each of the experimental groups: Control, Filters (F), Monitors (M), and Both (B). Columns 5–7 report the difference between the three treatment groups and the control, the standard error of the difference (in parentheses), and the number of observations used for the estimation (in square brackets). Columns 8–10 show the difference between treatment groups, the standard error of the difference (in parentheses), and the number of observations used for the estimation (in square brackets). All the differences between groups (Columns 4–10) consider the randomization design (i.e., include strata fixed effects). Standard errors are clustered at the school level. The p-value from a joint test that all coefficients in Columns 1–4 are zero (from a multinomial logit with the treatment groups on the left-hand side and all the covariates on the right-hand side) is .426. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table A.6: Balance across experimental groups (2022)

| | Control<br>(1) | Filter<br>(2) |
|---|---|---|
| **Panel A: Student level (2022)** | | |
| SES strata 1 or 2 | 0.62 | 0.03* |
| | (0.48) | (0.02) |
| | [23,110] | [41,035] |
| SES strata 3 or 4 | 0.31 | -0.02 |
| | (0.46) | (0.02) |
| | [23,110] | [41,035] |
| Female | 0.55 | -0.03** |
| | (0.50) | (0.01) |
| | [23,110] | [41,035] |
| Age (exam date) | 17.61 | -0.09* |
| | (1.52) | (0.05) |
| | [22,819] | [40,553] |
| Mother some secondary education | 0.49 | 0.01 |
| | (0.50) | (0.01) |
| | [23,110] | [41,035] |
| Global score | -0.02 | 0.05 |
| | (1.00) | (0.03) |
| | [23,110] | [41,035] |
| Reading score | -0.01 | 0.03 |
| | (1.01) | (0.03) |
| | [23,110] | [41,035] |
| Math score | -0.02 | 0.06* |
| | (1.00) | (0.03) |
| | [23,110] | [41,035] |
| **Panel B: School level (2022)** | | |
| PM2.5 | 17.50 | 0.06 |
| | (1.98) | (0.09) |
| | [207] | [357] |
| PM10 | 35.82 | 0.17 |
| | (5.61) | (0.25) |
| | [207] | [357] |
| O3 | 26.97 | -0.31 |
| | (4.30) | (0.28) |
| | [207] | [357] |
| Without Saber 11 data | 0.03 | -0.02 |
| | (0.17) | (0.02) |
| | [207] | [357] |
| Students (Grade 11) | 111.64 | 8.88 |
| | (68.92) | (7.83) |
| | [207] | [357] |

*Notes*: Column 1 reports the mean, standard deviation (in parentheses), and number of observations (in square brackets) of the control group. Column 2 shows the difference between the treatment and control groups, accounting for the randomization design (i.e., it includes strata fixed effects), the standard errors of the difference (clustered at the school level, in parentheses), and the number of observations (in square brackets). The p-value from a joint test that all coefficients are zero (from a regression with the treatment dummy on the left-hand side and all the covariates on the right-hand side) is .094 for Panel A and .718 for Panel B. $^{*}$ $p < 0.05$, $^{**}$ $p < 0.01$, $^{***}$ $p < 0.001$.

Table A.7: Balance across experimental groups (2023)

| | Control (1) | Filter (2) |
|---|---|---|
| **Panel A: Student level (2023)** | | |
| SES strata 1 or 2 | 0.59 | 0.02 |
| | (0.49) | (0.02) |
| | [23,897] | [42,389] |
| SES strata 3 or 4 | 0.32 | -0.01 |
| | (0.47) | (0.02) |
| | [23,897] | [42,389] |
| Female | 0.55 | -0.02** |
| | (0.50) | (0.01) |
| | [23,897] | [42,389] |
| Age (exam date) | 17.52 | -0.10* |
| | (1.59) | (0.06) |
| | [23,440] | [41,688] |
| Mother some secondary education | 0.47 | 0.01* |
| | (0.50) | (0.01) |
| | [23,897] | [42,389] |
| **Panel B: School level (2023)** | | |
| PM2.5 | 15.74 | 0.00 |
| | (2.06) | (0.08) |
| | [207] | [357] |
| PM10 | 31.60 | 0.22 |
| | (4.91) | (0.22) |
| | [207] | [357] |
| O3 | 27.28 | -0.38 |
| | (4.66) | (0.24) |
| | [207] | [357] |
| Without Saber 11 data | 0.03 | -0.01 |
| | (0.18) | (0.02) |
| | [207] | [357] |
| Students (Grade 11) | 115.44 | 8.69 |
| | (74.40) | (8.24) |
| | [207] | [357] |

*Notes*: Column 1 reports the mean, standard deviation (in parentheses), and number of observations (in square brackets) of the control group. Column 2 shows the difference between the treatment and control groups, accounting for the randomization design (i.e., it includes strata fixed effects), the standard errors of the difference (clustered at the school level, in parentheses), and the number of observations (in square brackets). The p-value from a joint test that all coefficients are zero (from a regression with the treatment dummy on the left-hand side and all the covariates on the right-hand side) is .126 for Panel A and .39 for Panel B. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table A.8: Effect on 2023 Saber tests for all treatment groups

| Dep. var. | Math | Spanish | Science | Social Sciences | English | Global |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Lagged scores** | | | | | | |
| Only Filter | 0.057** | 0.026 | 0.022 | 0.029* | 0.012 | 0.037* |
| | (0.025) | (0.020) | (0.023) | (0.017) | (0.024) | (0.022) |
| Only Monitor | -0.00086 | -0.0064 | -0.0055 | -0.010 | -0.022 | -0.0081 |
| | (0.023) | (0.019) | (0.020) | (0.018) | (0.021) | (0.020) |
| Filter and Monitor | 0.034 | 0.036* | 0.024 | 0.027 | 0.027 | 0.035* |
| | (0.021) | (0.019) | (0.019) | (0.019) | (0.024) | (0.020) |
| N. of obs. | 42,389 | 42,389 | 42,389 | 42,389 | 42,381 | 42,389 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.077 | 0.064 | 0.074 | 0.058 | 0.083 | 0.087 |
| **Panel B: Lagged scores + pollution** | | | | | | |
| Only Filter | 0.061** | 0.028 | 0.022 | 0.031* | 0.015 | 0.039* |
| | (0.024) | (0.019) | (0.022) | (0.016) | (0.024) | (0.021) |
| Only Monitor | 0.0019 | -0.00098 | -0.0039 | -0.0044 | -0.025 | -0.0043 |
| | (0.023) | (0.020) | (0.020) | (0.018) | (0.021) | (0.020) |
| Filter and Monitor | 0.033 | 0.036* | 0.024 | 0.027 | 0.026 | 0.034* |
| | (0.022) | (0.019) | (0.019) | (0.019) | (0.024) | (0.020) |
| N. of obs. | 42,389 | 42,389 | 42,389 | 42,389 | 42,381 | 42,389 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.078 | 0.065 | 0.074 | 0.058 | 0.083 | 0.088 |
| **Panel C: All controls** | | | | | | |
| Only Filter | 0.040* | 0.013 | 0.0025 | 0.016 | 0.0014 | 0.020 |
| | (0.021) | (0.017) | (0.020) | (0.016) | (0.021) | (0.019) |
| Only Monitor | -0.0054 | -0.0084 | -0.011 | -0.0090 | -0.032* | -0.012 |
| | (0.020) | (0.018) | (0.018) | (0.017) | (0.018) | (0.017) |
| Filter and Monitor | 0.039** | 0.041** | 0.029* | 0.031* | 0.031 | 0.040** |
| | (0.020) | (0.017) | (0.017) | (0.018) | (0.022) | (0.018) |
| N. of obs. | 42,389 | 42,389 | 42,389 | 42,389 | 42,381 | 42,389 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.16 | 0.11 | 0.14 | 0.095 | 0.14 | 0.16 |

*Notes*: All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Heterogeneity in treatment effects on the 2023 global score

| Covariate | High performing (1) | High pollution (2) | Female (3) | Age (4) | Low strata (5) |
|---|---|---|---|---|---|
| Treatment (filters) | 0.068* | 0.046** | 0.026 | 0.033** | 0.017 |
| | (0.0383) | (0.0203) | (0.0196) | (0.0140) | (0.0240) |
| Treatment (filters) × Covariate | -0.039 | -0.021 | 0.011 | -0.023 | 0.017 |
| | (0.0524) | (0.0317) | (0.0246) | (0.0171) | (0.0305) |
| Covariate | 0.42*** | 0.034 | -0.28*** | -0.10*** | 0.15*** |
| | (0.0379) | (0.0507) | (0.0183) | (0.0107) | (0.0213) |
| N. of obs. | 42,389 | 42,389 | 42,389 | 42,389 | 42,389 |
| N. Schools | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.055 | 0.087 | 0.10 | 0.11 | 0.092 |

*Notes*: The outcome is the global score in all regressions. The covariate in Columns 1 and 2 is defined at the school level. In Column 1, the covariate is a dummy that equals 1 if the school's average score between 2019 and 2022 is above the median average score of schools in the sample. Likewise, in Column 2, the covariate is a dummy that equals one if the average $PM_{2.5}$ during school hours (8 am to 4 pm) in the 270 days before the exam is above the median for all schools. In Columns 3–5, the covariates are defined at the student level. We use sex, age (demeaned, so the interaction coefficient is interpreted as the marginal effect at the mean age), and whether the student lives in a household entitled to public utility subsidies. All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table A.10: Heterogeneity in treatment effects on the 2023 global score by school physical features

| Covariate | Vents | Rotative | Filter Exposure | Classroom Volume | Proxy Usage | Math Classroom |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Treatment | 0.0069 | -0.0059 | 0.016 | -0.0066 | -0.036 | -0.11 |
| | (0.0888) | (0.0933) | (0.0847) | (0.0849) | (0.0526) | (0.107) |
| Treatment x Covariate | -0.056 | 0.015 | -0.00052 | 0.00015 | 0.00039 | 0.059 |
| | (0.0473) | (0.0545) | (0.00213) | (0.000114) | (0.000598) | (0.0454) |
| Covariate | 0.0028 | -0.0044 | 0.00067 | -0.00023** | -0.00075 | -0.031 |
| | (0.0396) | (0.0497) | (0.00177) | (0.000107) | (0.000852) | (0.0388) |
| N. of obs. | 42389 | 42389 | 42389 | 42389 | 42389 | 42389 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | .088 | .087 | .087 | .088 | .087 | .077 |
| Mean of Covariate | .39 | .46 | 8.2 | 150 | 45 | .31 |

*Notes*: The outcome is the global score in all regressions. All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table A.11: Attrition in the pollution monitor data

| | Sensor-level panel | | | | Location-level panel | | | School-level panel | |
|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| Filter and Monitor | -.0062 | -.0062 | -.006 | -.033 | .0087 | .0087 | -.024 | .0091 | .0091 |
| | (.027) | (.027) | (.027) | (.03) | (.032) | (.032) | (.036) | (.041) | (.041) |
| Outside | | | | .032 | | | -.058* | | |
| | | | | (.029) | | | (.029) | | |
| Filter & Monitor × Outside | | | | .079* | | | .067* | | |
| | | | | (.042) | | | (.038) | | |
| Control mean | .41 | .41 | .41 | .41 | .46 | .46 | .46 | .51 | .51 |
| N. of obs. | 111,195 | 111,195 | 111,195 | 111,195 | 78,013 | 78,013 | 78,013 | 43,066 | 43,066 |

*Notes*: The outcome in all columns is whether the air pollution monitor data is missing. In Columns 1–4, the dataset is a monitor-level panel at the day level. In Columns 5–7, the data is a daily school-location-level panel: each school has two observations per day, one outdoors and one indoors. In Columns 8–9, the data is a school-level panel at the day level. All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table A.12: Effect of filters on students' exposure to $PM_{2.5}$

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Panel A: Outdoor** $PM_{2.5}$ | | | | | |
| Filter and Monitor | .57$^{*}$ | .47 | .42 | .41 | .39 |
| | (.33) | (.32) | (.32) | (.32) | (.36) |
| Control mean | 11 | 11 | 11 | 13 | 18 |
| N. of obs. | 808,273 | 808,269 | 808,189 | 233,908 | 156,174 |
| **Panel B: Indoor** $PM_{2.5}$ | | | | | |
| Filter and Monitor | -.4 | -.5$^{*}$ | -.53$^{**}$ | -.57$^{**}$ | -.99$^{**}$ |
| | (.27) | (.26) | (.25) | (.27) | (.43) |
| Control mean | 9.5 | 9.5 | 9.5 | 11 | 15 |
| N. of obs. | 1,405,455 | 1,405,455 | 1,405,447 | 410,059 | 205,102 |
| **Panel C: $\Delta$ Indoor–Outdoor** $PM_{2.5}$ | | | | | |
| Inside | -1.3$^{***}$ | -1.3$^{***}$ | -1.3$^{***}$ | -1.6$^{***}$ | -2.5$^{***}$ |
| | (.24) | (.24) | (.23) | (.28) | (.33) |
| Filter & Monitors $\times$ Indoors | -.9$^{**}$ | -.89$^{**}$ | -.8$^{**}$ | -.95$^{**}$ | -1.4$^{***}$ |
| | (.37) | (.37) | (.37) | (.43) | (.49) |
| Control mean | 11 | 11 | 11 | 13 | 18 |
| N. of obs. | 2,213,728 | 2,213,728 | 2,213,134 | 643,774 | 361,739 |
| Locality FE | X | X | | | |
| Hour FE | X | | | | |
| Date FE | X | | | | |
| Hour $\times$ Date FE | | X | | | |
| Locality $\times$ Date $\times$ Hour FE | | | X | X | X |
| Hours | All | All | All | 7am–4pm | 7am–4pm |
| Days | All | All | All | Mon–Fri | Mon–Fri |
| Outdoors pollution | All | All | All | All | $> 5$ |

*Notes*: The outcome in all regressions is the indoor readings of $PM_{2.5}$. In all regressions, we control for the outdoor readings of $PM_{2.5}$ (demeaned), both as measured by the monitors we installed in schools and by the city's monitors (SISAIRE) and their interaction with the treatment dummy (following J. Roth and Sant'Anna (2023); Lin (2013)). Column 5 defines outdoor pollution based on the monitors we installed. Column 6 defines outdoor pollution based on the city's monitors. All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

Table A.13: Effects of filters on students' exposure to $PM_{2.5}$ by schools' pollution levels

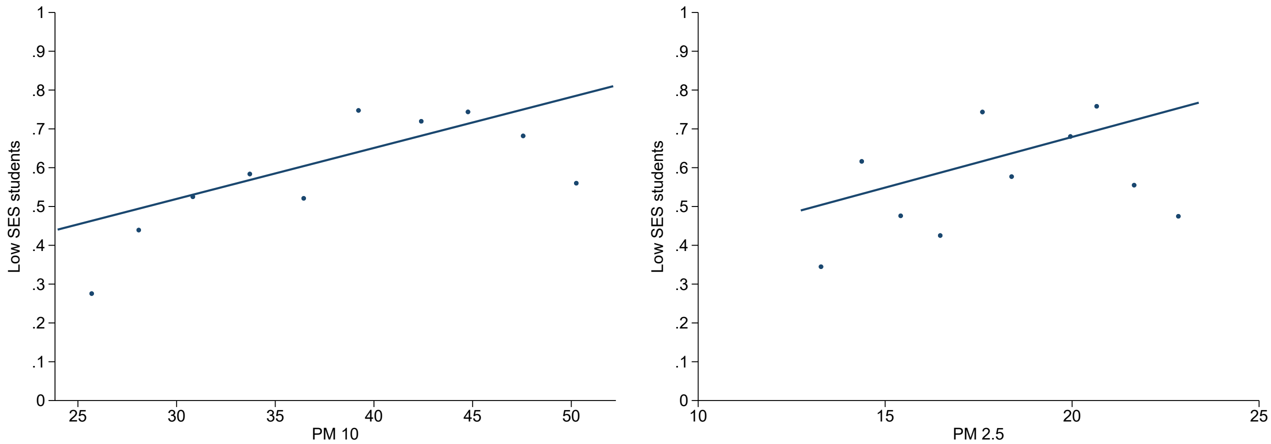
| | $PM_{2.5}$ (1) | (2) |
|---|---|---|
| Filter and Monitor | -.49 | -.58* |
| | (.34) | (.34) |
| PM25_ext_demean_NM × Dummy_NM | .083 | |
| | (.052) | |
| PM25_ext_demean_NM × Filter and Monitor × Dummy_NM | .013 | |
| | (.04) | |
| PM 2.5 (city sensors) | | .046* |
| | | (.026) |
| PM 2.5 (city sensors) × Filter and Monitor | | -.05* |
| | | (.027) |
| N. of obs. | 69,739 | 69,739 |
| Locality × Date × Hour FE | X | X |
| Hours | 7am–4pm | 7am–4pm |
| Days | Mon–Fri | Mon–Fri |
| Outdoor pollution | All | All |

*Notes*: The outcome in all regressions is the indoor readings of $PM_{2.5}$. In all regressions, we control for the outdoor readings of $PM_{2.5}$ (demeaned), both as measured by the monitors we installed in schools and by the city's monitors (SISAIRE) and their interaction with the treatment dummy (following J. Roth and Sant'Anna (2023); Lin (2013)). Column 1 defines outdoor pollution based on the monitors we installed. Column 2 defines outdoor pollution based on the city's monitors. All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

## A.2 Figures

Figure A.1: Pollution and SES across years

(a) $PM_{10}$ and proportion of low SES students (b) $PM_{2.5}$ and proportion of low SES students




*N*otes: These figures show the correlation between the average pollution and the likelihood that a student comes from a low SES background (measured by whether they live in a household entitled to subsidies for public utilities). The figures represent binscatter least squares estimations — following Cattaneo, Crump, Farrell, and Feng (2024) — using data from all students graduating from all schools in our experimental sample from 2021.

Figure A.2: Infographic about the filter

Figure A.3: Infographic about the air quality monitor

Figure A.4: Filter installations and exam date

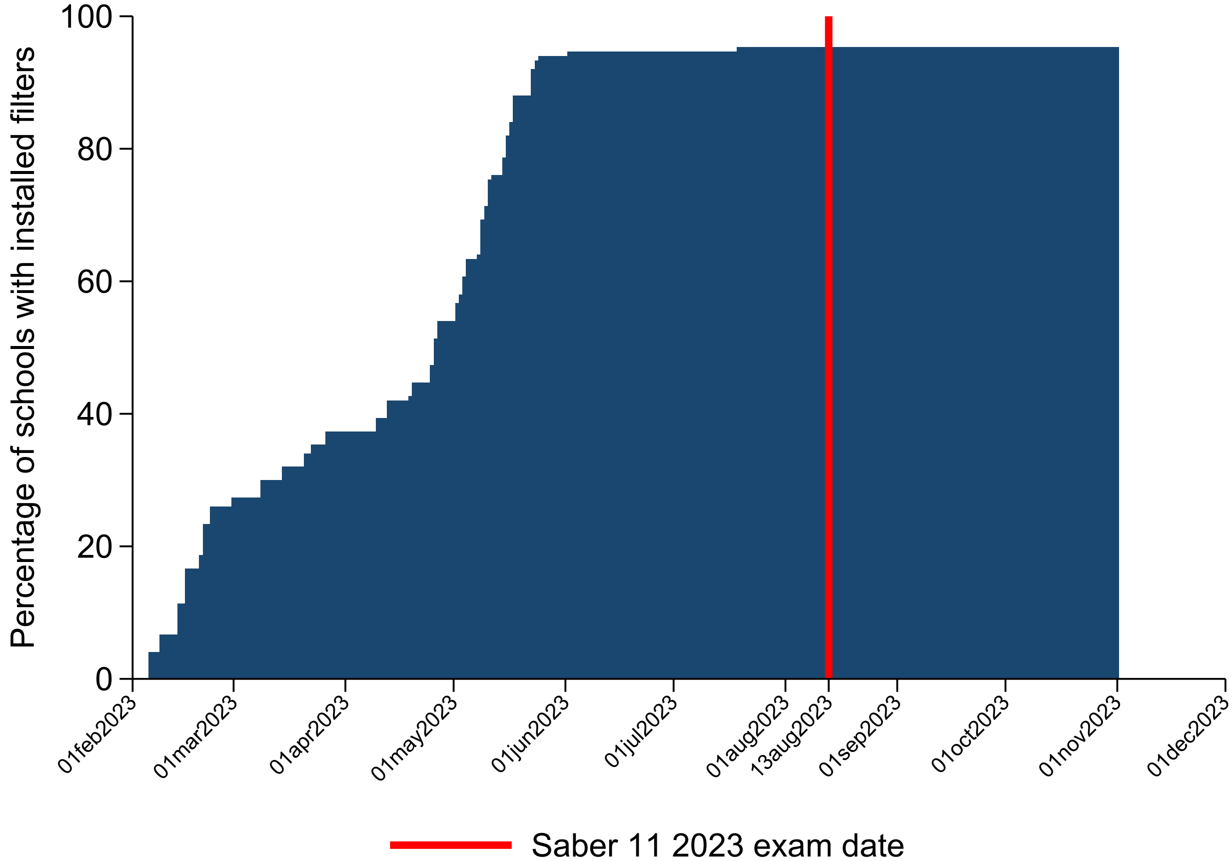


Figure A.5: Monitor installations and exam date

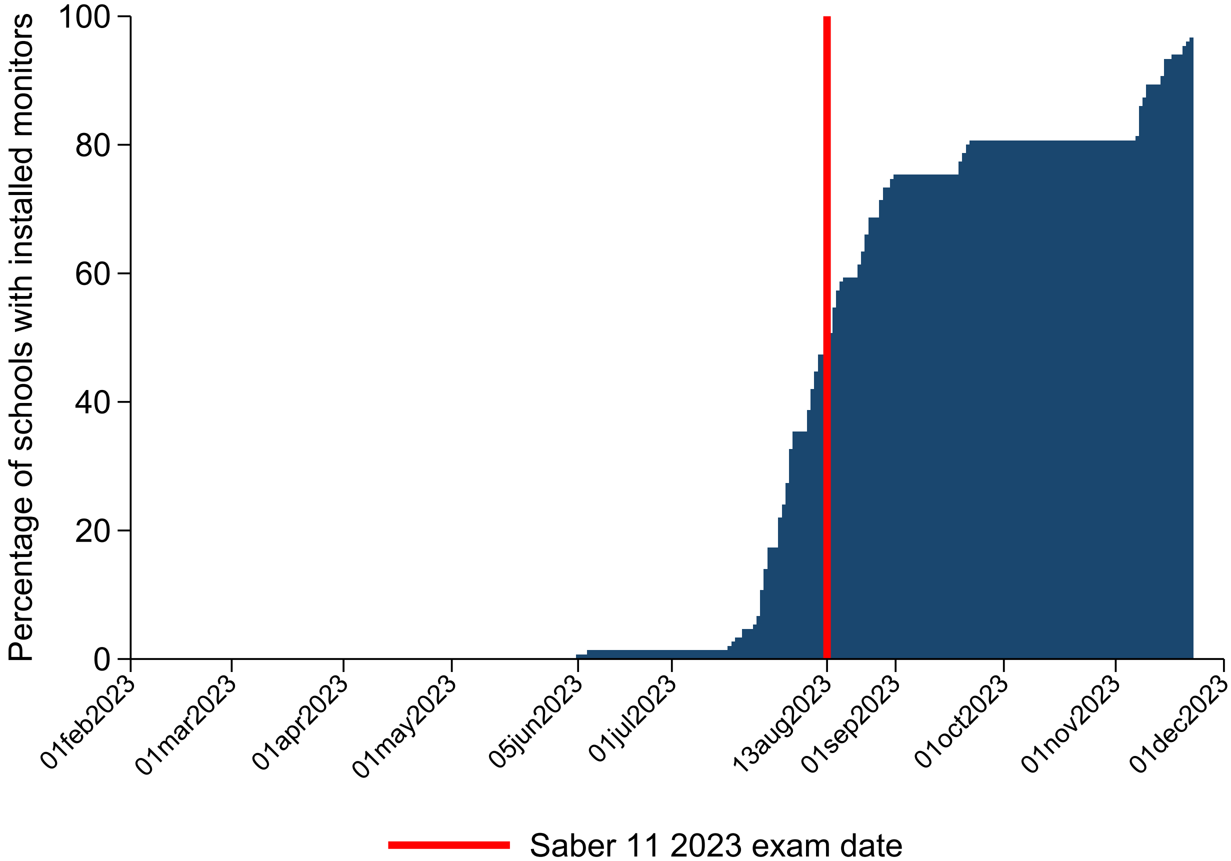


Figure A.6: Consort diagram

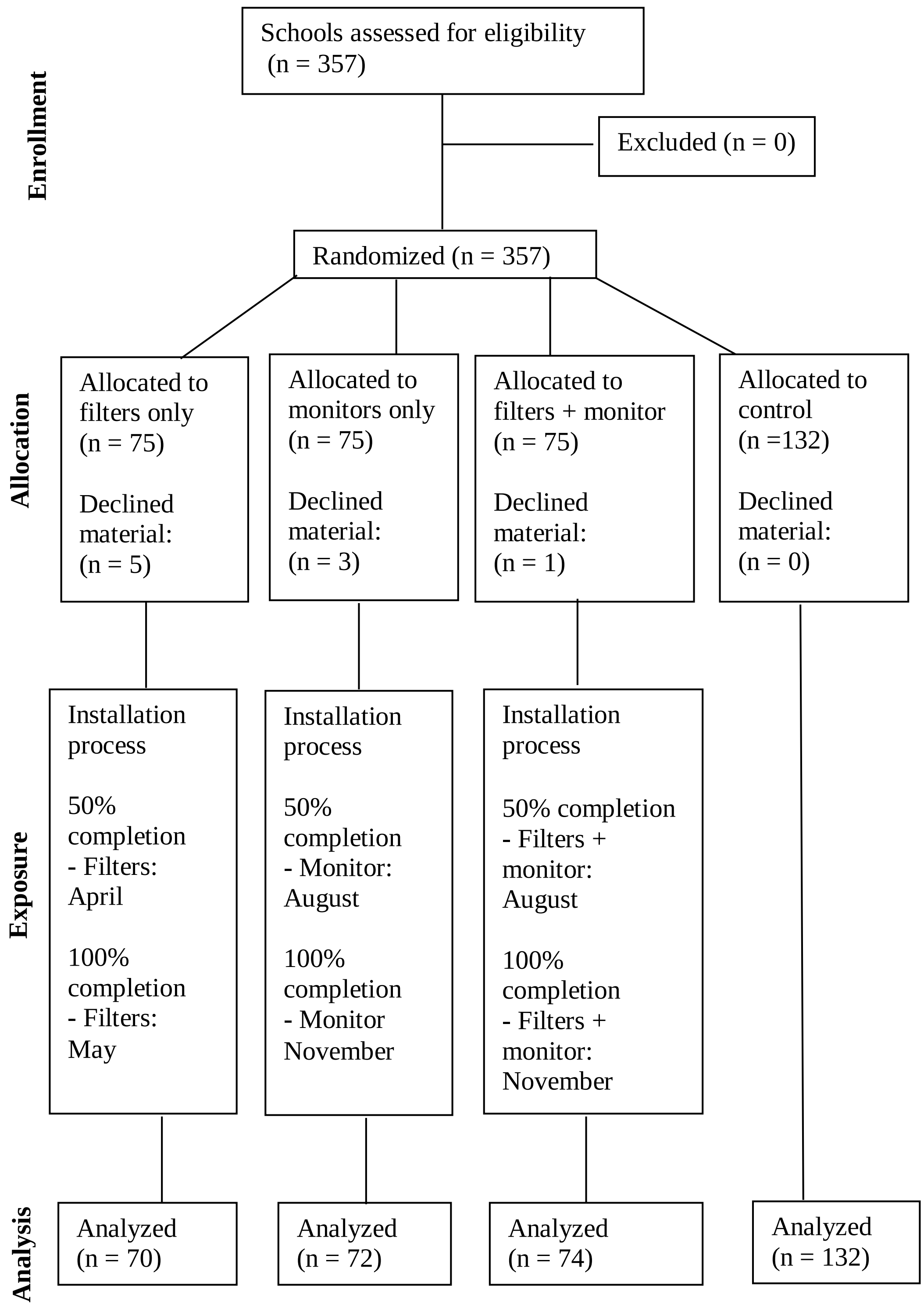
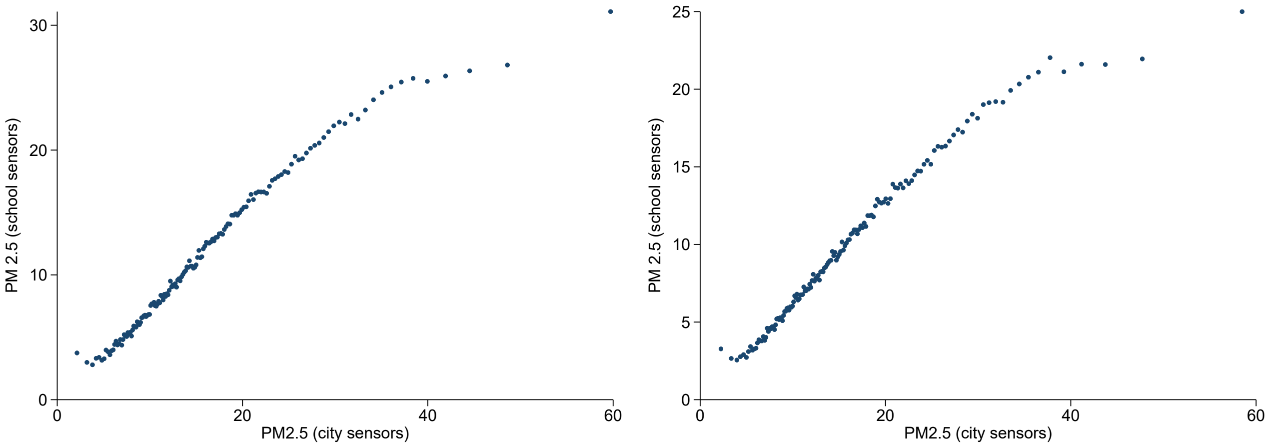


Figure A.7: Pollution readings: SISAIRE vs. school monitors

(a) Outdoor school monitors vs. city monitors (b) Indoor school monitors vs. city monitors

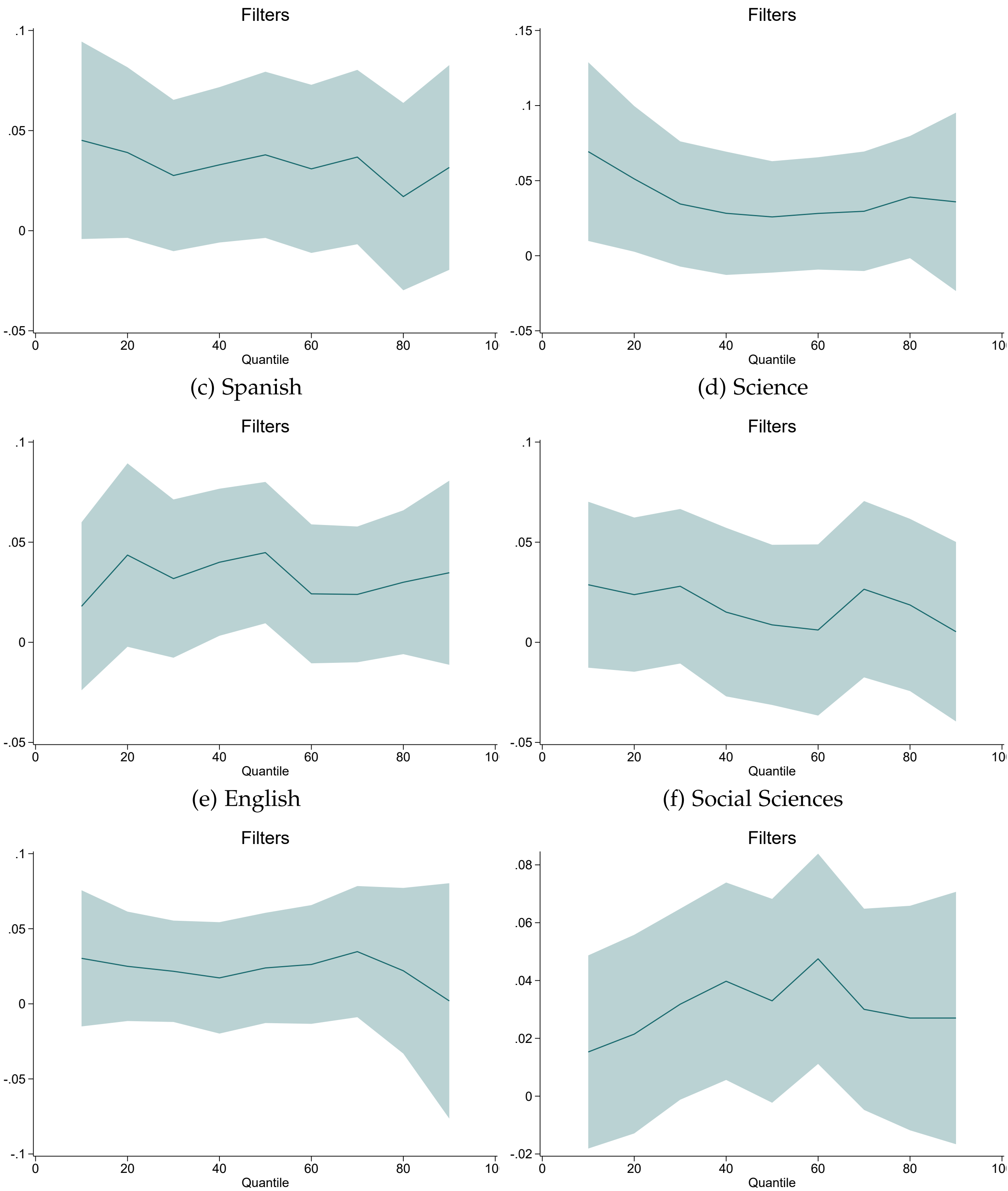



*N*otes: These figures represent binscatter least squares estimations—following Cattaneo et al. (2024)—using hourly data from the pollution monitors we installed in schools, as well as the network of 22 monitors managed by the government, known as SISAIRE. We use a simple interpolation (based on the quadratic distance to the monitor to obtain hourly pollution measurements for each school in the sample.

Figure A.8: Classroom with slits

Figure A.9: Quantile treatment effects (2023)

(a) Global (b) Math




(c) Spanish (d) Science

(e) English (f) Social Sciences

*N*otes: These figures present unconditional quantile treatment effects (Firpo et al., 2009) and their 95% confidence intervals estimated using the RIF package in Stata (Rios-Avila, 2020).

Figure A.10: $PM_{2.5}$ readings from the monitors we installed in schools over time

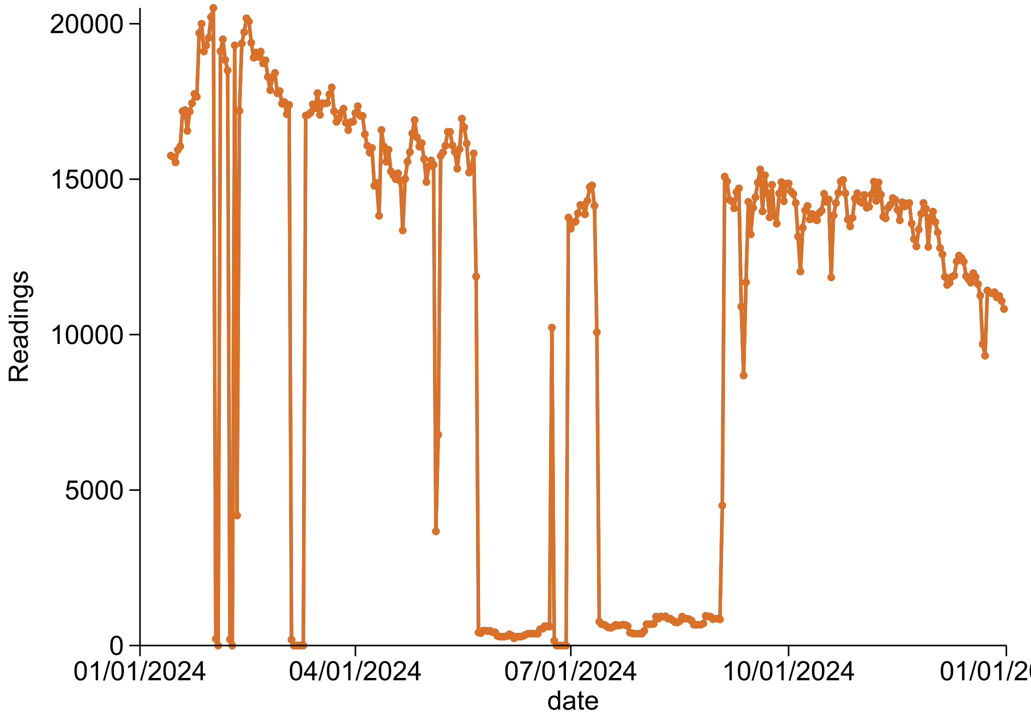
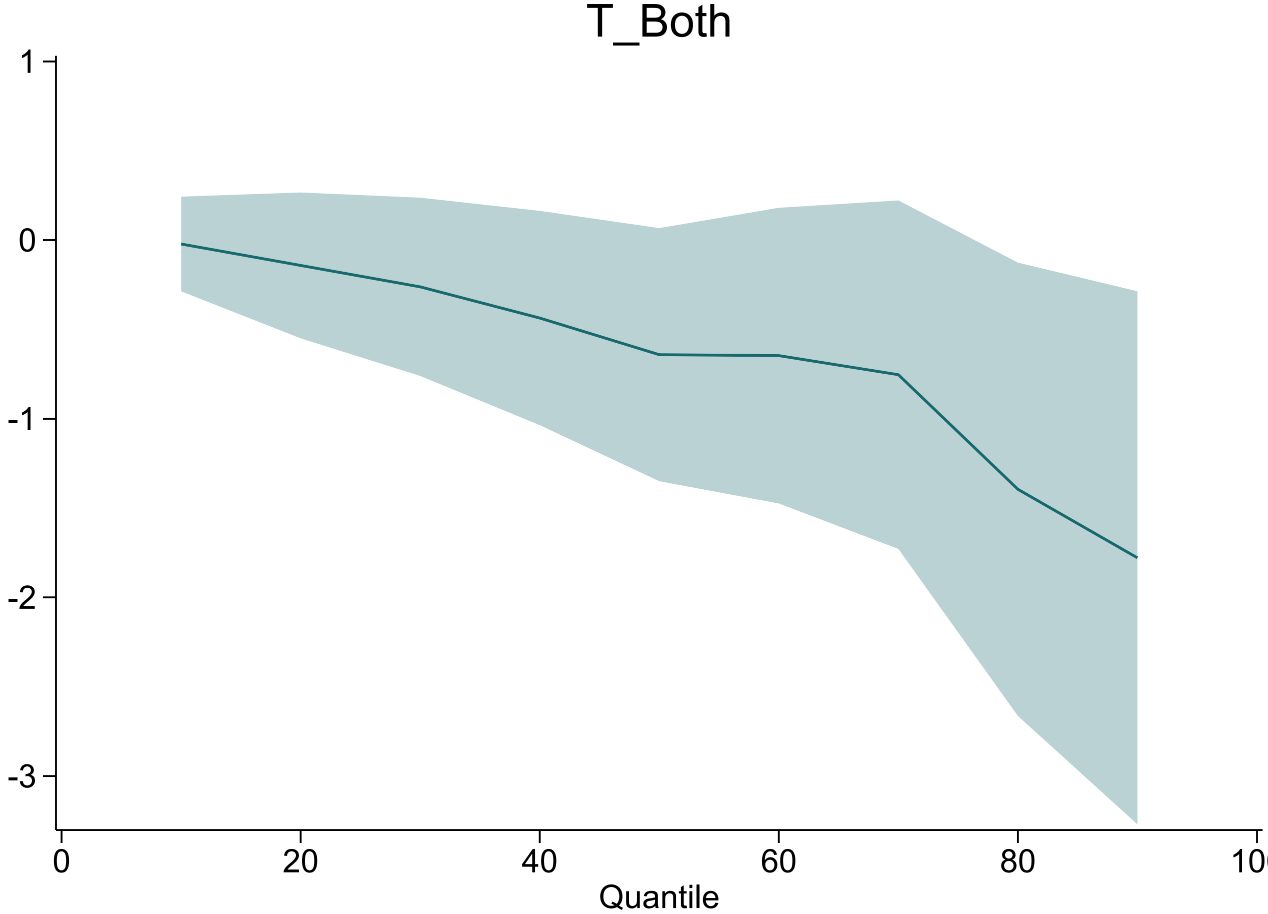
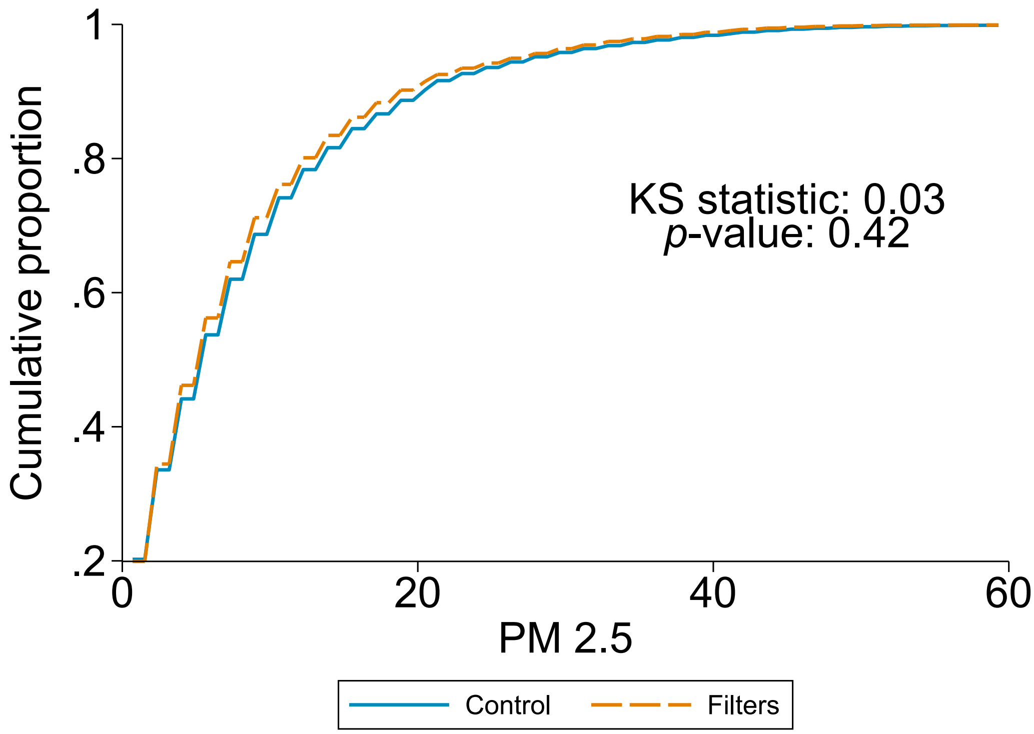


Figure A.11: Quantile treatment effects on $PM_{2.5}$ levels



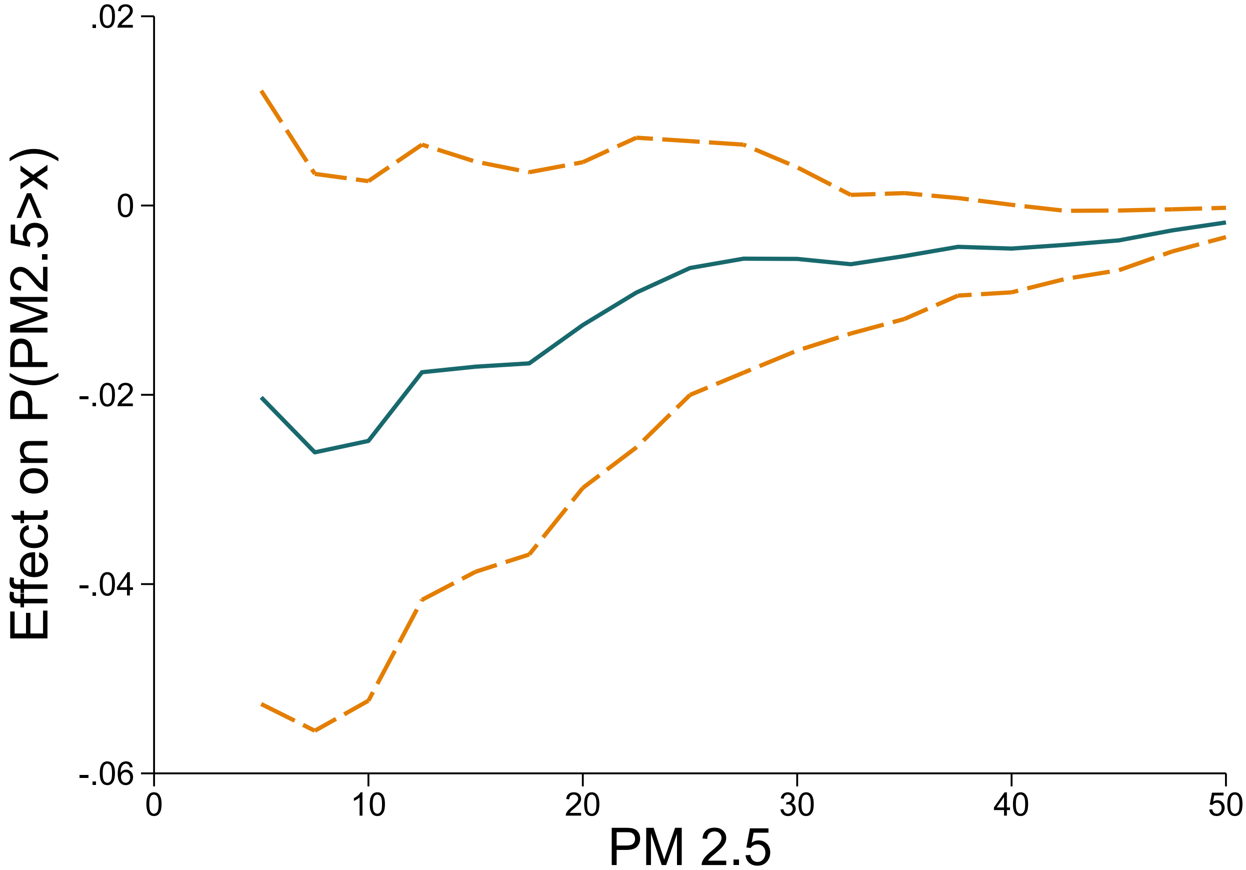

*N*otes: These figures present unconditional quantile treatment effects (Firpo et al., 2009) on indoor $PM_{2.5}$ levels and their 95% confidence intervals estimated using the RIF package in Stata (Rios-Avila, 2020).

Figure A.12: Indoor $PM_{2.5}$ distribution across schools with and without filters




*Note*: This figure presents the cumulative distribution of $PM_{2.5}$ by treatment status (with or without filters, among schools with monitors) (Abadie, 2002) (and the implementation of Abdulkadiroğlu, Pathak, and Walters (2018)). All estimates control for outdoor pollution readings, hour, day, and locality fixed effects. This figure only includes data for school hours (weekdays, 7 am–4 pm). All models use a Gaussian kernel and Silverman (2018)'s rule of thumb bandwidth. KS statistics are the maximum differences in CDFs.P-valuess are estimated via bootstrap.

Figure A.13: Distribution regression for indoor $PM_{2.5}$ levels




*Notes*: This figure presents distribution regression estimators. For each value of $PM_{2.5}$, it displays the effect on the probability that the indoor pollution level is above this value.

# B Results for 2024

## B.1 Test scores

Table B.1: Effect of filters on test scores

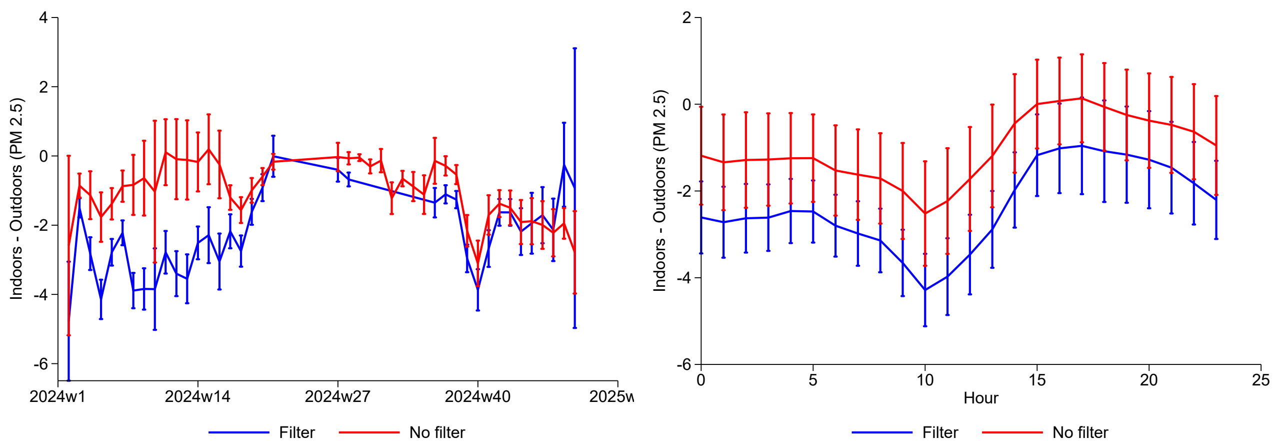
| Dep. var. | Math | Spanish | Science | Social Sciences | English | Global |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| **Panel A: Lagged scores** | | | | | | |
| Treatment (filters) | -0.0052 | -0.019 | -0.0057 | -0.019 | -0.011 | -0.014 |
| | (0.0183) | (0.0164) | (0.0185) | (0.0172) | (0.0171) | (0.0180) |
| N. of obs. | 42,491 | 42,491 | 42,491 | 42,491 | 42,003 | 42,491 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.070 | 0.061 | 0.066 | 0.054 | 0.086 | 0.082 |
| **Panel B: Lagged scores + pollution** | | | | | | |
| Treatment (filters) | -0.0046 | -0.022 | -0.0039 | -0.021 | -0.0059 | -0.014 |
| | (0.0173) | (0.0159) | (0.0182) | (0.0166) | (0.0167) | (0.0173) |
| N. of obs. | 42,491 | 42,491 | 42,491 | 42,491 | 42,003 | 42,491 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.071 | 0.061 | 0.066 | 0.054 | 0.087 | 0.083 |
| **Panel C: All controls** | | | | | | |
| Treatment (filters) | -0.0030 | -0.021 | -0.0026 | -0.021 | -0.0024 | -0.012 |
| | (0.0151) | (0.0142) | (0.0170) | (0.0165) | (0.0162) | (0.0158) |
| N. of obs. | 42,491 | 42,491 | 42,491 | 42,491 | 42,003 | 42,491 |
| N. Schools | 346 | 346 | 346 | 346 | 346 | 346 |
| $R^2$ | 0.19 | 0.12 | 0.15 | 0.096 | 0.14 | 0.17 |

*Notes*: All regressions consider the randomization design (i.e., include strata fixed effects). Panel A controls for average school-level lagged scores (we do not have individual student-lagged scores). Panel B also controls for the average pollution levels of $PM_{2.5}$, $PM_{10}$, $NO_2$, $CO$, and $O_3$ during school hours (7 am–4 pm) in the 270 days (roughly the length of the school year) and 90 days (roughly the length of exposure to the filters) preceding the exam. Panel C controls for lagged scores, pollution levels, age, sex, mother's level of education, and socio-economic strata. Standard errors, clustered at the school level, are in parentheses. $^{*}$ $p < 0.10$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

## B.2 Air quality

Figure B.1: Difference between indoor and outdoor readings in schools with and without filters

(a) Differences by weeks (b) Differences by time of the day




*N*otes: These figures display the average difference between indoor and outdoor readings in each school, depending on whether they have filters. Figure 3a averages the data across weeks, while Figure 3b averages the data across hours of the day.

Figure B.2: Percentage of Monitors with data transmission by day in 2024

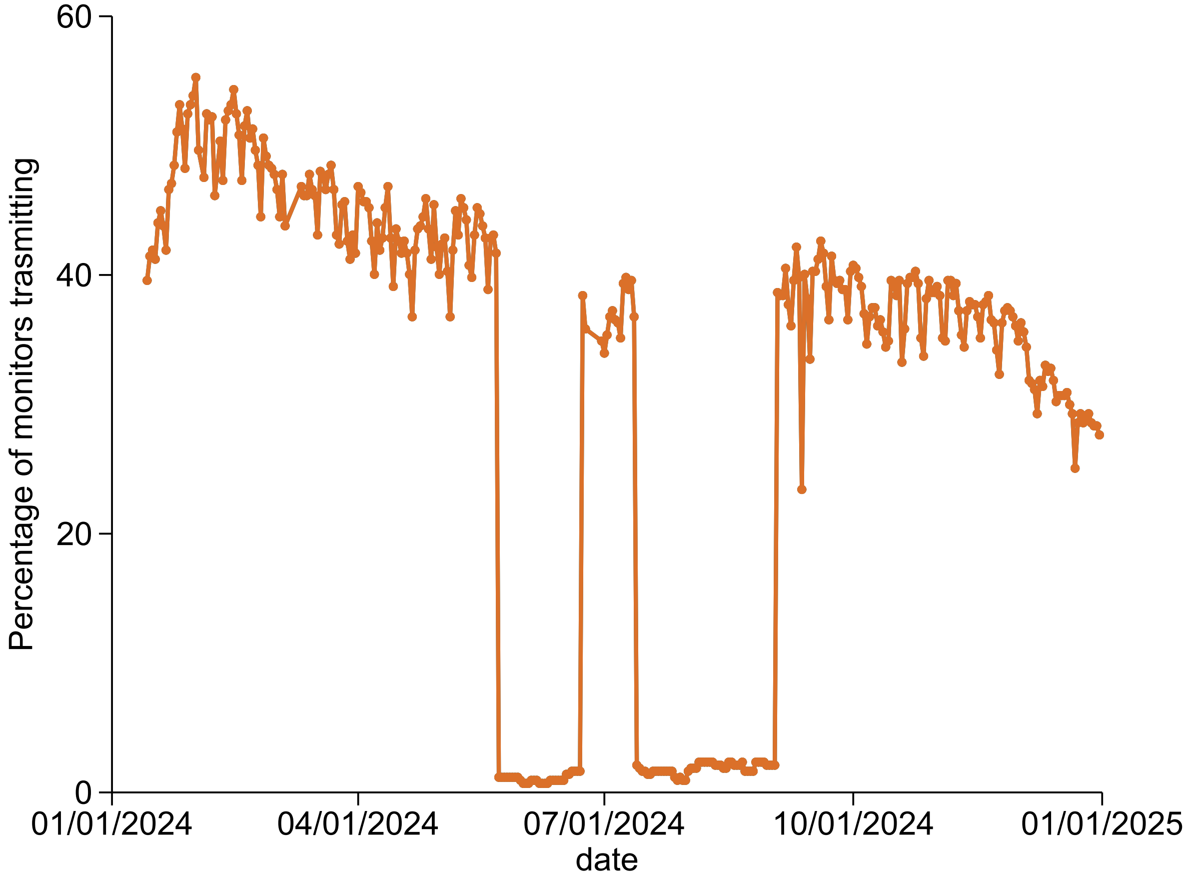


Table B.2: Effect of filters on students' exposure to $PM_{2.5}$

| Dep. var. | $PM_{2.5}$ Indoors pollution ($\mu g/m^3$) | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| Filter and Monitor | -.94** | -.86** | -.8** | -.66* | -1.2*** |
| | (.36) | (.34) | (.33) | (.36) | (.45) |
| Control mean | 15 | 15 | 15 | 16 | 20 |
| N. of obs. | 2,000,016 | 2,000,013 | 1,999,964 | 609,559 | 331,699 |
| Locality FE | X | X | | | |
| Hour FE | X | | | | |
| Date FE | X | | | | |
| Hour × Date FE | | X | | | |
| Locality × Date × Hour FE | | | X | X | X |
| Hours | All | All | All | 7am–4pm | 7am–4pm |
| Days | All | All | All | Mon–Fri | Mon–Fri |
| Outdoors pollution | All | All | All | All | $> 5$ |

*Notes*: The outcome in all regressions is the indoor readings of $PM_{2.5}$. All regressions control for outdoor readings of $PM_{2.5}$ (demeaned), both as measured by the monitors we installed in schools and the city monitors and their interaction with the treatment dummy (following J. Roth and Sant'Anna (2023); Lin (2013)). All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## B.3 Effect of air quality on test scores

Table B.3: Effect on 2024 Saber tests

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **Panel A: First Stage** | | | | | | |
| Dep. var. | $PM_{2.5}$ | | | | | |
| Treatment (filters) | -0.96** | | | | | |
| | (0.461) | | | | | |
| N. of obs. | 12,235 | | | | | |
| Only monitors mean | 17.2 | | | | | |
| N. Schools | 109 | | | | | |
| $R^2$ | 0.31 | | | | | |
| **Panel B: Reduced Form** | | | | | | |
| Dep. var. | Math | Spanish | Science | Social Sciences | English | Global |
| Treatment (filters) | -0.0011 | -0.0014 | -0.0100 | 0.021 | 0.023 | 0.0049 |
| | (0.0350) | (0.0288) | (0.0305) | (0.0297) | (0.0284) | (0.0309) |
| N. of obs. | 12,235 | 12,235 | 12,235 | 12,235 | 12,086 | 12,235 |
| N. Schools | 109 | 109 | 109 | 109 | 109 | 109 |
| $R^2$ | 0.068 | 0.054 | 0.066 | 0.052 | 0.075 | 0.077 |
| **Panel C: IV Estimates** | | | | | | |
| Dep. var. | Math | Spanish | Science | Social Sciences | English | Global |
| Average PM 2.5 (indoors) | 0.0012 | 0.0014 | 0.010 | -0.022 | -0.024 | -0.0051 |
| | (0.0363) | (0.0299) | (0.0322) | (0.0327) | (0.0312) | (0.0321) |
| N. of obs. | 12,235 | 12,235 | 12,235 | 12,235 | 12,086 | 12,235 |
| N. Schools | 109 | 109 | 109 | 109 | 109 | 109 |
| $R^2$ | 0.067 | 0.054 | 0.065 | 0.051 | 0.072 | 0.077 |

*Notes*: All regressions consider the randomization design (i.e., include strata fixed effects). Standard errors, clustered at the school level, are in parentheses. The p-value of joint F-test for the first stage is 4.318000000000001. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.