Indoor Air Quality and Student Performance: Evidence from A Large Scale Field Study in Primary Schools

Nicolás Durán[1], Piet Eichholtz[1], Nils Kok[1], and Juan Palacios[2]

[1]Maastricht University, School of Business and Economics

[2]Massachusetts Institute of Technology, Center for Real Estate

February 21, 2021

Abstract

Governments devote a large share of public budgets to construct, repair and modernize school facilities. However, little is known about whether investments in the physical condition of schools translate into student achievements. In this study, we report the results of a large field study, providing quasi-experimental evidence on the implications for student performance of poor environmental conditions inside classrooms − key performance measure of school infrastructure, and a common indicator guiding investments in school facilities. We continuously monitor the environmental conditions (i.e. CO_2, fine particles, temperature, humidity) in the classrooms of 3,000 children over two school years, and link them to their scores in over 14,000 nationally standardized tests. Using a fixed-effects strategy, relying on within-pupil changes in environmental conditions, we find that exposure to poor indoor air quality during the school term preceding the test is associated with significant performance drops. We document that one standard deviation increase in average CO_2 during the school term leads to a 0.14 standard deviation drop in test scores. We document changes in teaching time as a potential mechanism of drops in performance. Classes exposed to poor indoor environmental conditions in a given school day tend to have significantly longer breaks, leading to a shorter time in the classroom. Our results add to the ongoing debate on the determinants of student human capital accumulation, highlighting the role of physical capital in affecting learning outcomes.

Keywords: Academic Performance, Human Capital, Indoor Air Quality, School Infrastructure, Education.

JEL classification: I21, Q53.

1 Introduction

Governments across the globe invest large amounts of capital in school facilities. In the U.S. alone, the 2020 “Reopen and Rebuild America’s Schools Act” allocates USD130 billion to the renovation, modernization and construction of schools across the country. These investments will mostly be devoted to improving schools that are in some state of disrepair. Indeed, results from a recent survey of 100,000 school districts by the Government Accountability Office (G.A.O.) show that half of those schools are poorly equipped or in poor physical condition, with heating, ventilation and air conditioning (HVAC) systems having the highest priority to receive investments (G.A.O., 2020).

Poor indoor air quality is a key risk factor for the health and performance of students and staff in schools. The airborne transmission of the SARS-COV-2 virus is further exacerbating the consequences of deficient school infrastructure, increasing the need for investments in school buildings. In response, governments in countries like Germany, the Netherlands or the US are increasing public spending to upgrade ventilation systems to reduce the risk of transmission in schools (BBC, 2020; Rijksoverheid, 2020). Will such investments also support the human capital accumulation of children? Numerous lab studies provide evidence on the detrimental consequences of exposures to sub optimal temperatures or to poorly ventilated rooms for human cognition and decision making (Seppanen et al., 2006; Fisk, 2017; Du et al., 2020). Current evidence is mainly based on lab studies or small-scale field interventions, where researchers employ artificial tasks to measure cognitive acuteness of subjects. The lack of large, long term studies hinders the evaluation of costly investments in school infrastructure, which will modify the environmental conditions in schools for a long term, and potentially affect differently children with abilities or socio−demographic background.

The main contribution of this paper is to provide evidence into the implications of indoor environmental conditions for academic achievement. Here, we report the results of a large field study for which we deploy a network of environmental sensors, continuously monitoring the indoor environmental conditions in 235 classrooms across 27 Dutch primary schools across two academic years. Each sensor collects high-frequency measurements on a range of indoor environmental variables – i.e. CO_2, fine particles, temperature, humidity, noise and light intensity. To estimate the causal impact of classroom conditions on the cognitive development of students, we relate daily measures of indoor air quality of the classroom of 3,000 primary school students aged 6 to 12 to their scores in 14,000 nationally standardized tests. Primary school students spend most of their schooldays in the same classroom, providing substantial exposure to the indoor climate conditions in that specific room. During the sample period, each student took an average of 9 standardized, national exams across a range of topics, including mathematics, spelling, reading and vocabulary. These exams are designed by a national examination center (i.e. not the teacher), with the aim of assessing the learning development of students throughout their primary school education. Student fixed effects regressions identify the impact of indoor environmental conditions during the school terms by leveraging within-student variation in classroom conditions over multiple test intakes.

The main results show that children who were exposed during the learning period to high concentrations of CO_2 perform worse in standardized tests. CO_2 is a widely used indicator by building scientists to measure how much fresh (outdoor) air is brought into a room, and by public officials to set guidelines and evaluate the performance of ventilation systems in public buildings. The presence of high levels of CO_2 concentration indicate deficient air exchange in the room. In our preferred specification, including a rich set of fixed effects, one standard deviation increase in CO_2 during the school term leads to a 0.14 standard deviation drop in test scores. We document non-linearities in these effects. The larger drops in test scores are driven by those classrooms falling in the percentiles of the distribution CO_2, reaching a 0.4 standard deviations drops in children performance in classrooms with an average CO_2 levels above 1,500 ppm. When broken down into specific test domains, we find that CO_2 hinders the learning of students in mathematics, and specially, reading. In addition, exposure to fine particles have significant detrimental effects on children's performance in mathematics. Our estimates show a detrimental impact of exposure to indoor air particles during the school term on the performance of students in mathematics by up to 0.5 standard deviations. In a secondary analysis, we study how detrimental indoor environmental conditions during a specific lesson affects the time spent in the subsequent break, as a potential mechanism through which poor indoor air quality might affect the final test scores. Our estimates suggest that after exposure to relatively high levels of CO_2 or temperature during the lesson, teachers and children take breaks that are longer by 15% to 20%.

This study is the first to show that exposure to poor environmental conditions *inside* the classroom can reduce the rate of human capital accumulation, which speaks to the long-standing debate on the relationship between investments in school infrastructure and academic achievements (see Hanushek (2003)). Existing evidence shows a positive impact of school construction projects in contexts where school facilities were either in extremely poor condition or just nonexistent, which suggests that new school construction projects generally are positively associated with student outcomes (Duflo, 2001; Aaronson and Mazumder, 2011; Neilson and Zimmerman, 2014). Similarly, another stream of quasi-experimental studies investigates the link between (general) school spending or school investment campaigns for school infrastructure and academic outcomes (Martorell et al., 2016; Jackson et al., 2016). Finally, Stafford (2015) provides evidence that public funding campaigns targeting mold reduction and ventilation improvements have a positive impact on student performance in elementary schools. This study departs from those studies by investigating actual indoor environmental conditions in the classroom, rather than just broad, monetary indicators of changes in school infrastructure, providing insights based on objective, high-frequency measures of indoor environmental quality. Our outcome-based approach of school quality should allow for more precise estimates that a purely input-based approach (see Hanushek (2003) for a discussion of misallocation of resources in school investments).

This paper also contributes to the sizeable literature exploring the role of environmental factors (i.e., air pollution and extreme heat) in cognitive performance and human capital development. Over the last decade, there have been a number of studies providing quasi-experimental evidence on the negative effects of exposure to extreme temperatures or ambient air pollution on human health and human capital accumulation (see, for example, Graff Zivin and Neidell (2013) and Graff Zivin and Neidell (2018)). Prolonged exposure to high levels of air pollution has been associated with numerous respiratory problems (e.g. asthma), ultimately leading to school absences (Currie et al., 2009; Currie and Walker, 2011; Knittel et al., 2016), and declines in infant mortality (Chay and Greenstone, 2003; Currie and Neidell, 2005).[1]

Beyond the health damage, there is increasing evidence on the harmful consequences of exposure to air pollution on the human brain and cognitive performance (Zhang et al., 2018). An increasing number of studies show that exposure to air pollution harms student performance. Numerous studies have linked local levels of air pollution on testing days (i.e. high levels of $PM_{2.5}$) to lower performance of young adults in high-stakes examinations (Ebenstein et al., 2016; Roth, 2018; Graff Zivin et al., 2020).[2] In the medium term, accumulated exposure to

[1] There are also numerous studies showing the effects of elevated concentrations of fine particles on mortality rates in adult populations (Liu et al., 2019). At the macro level, the impact of air pollution on human health is staggering: the World Health Organization (WHO) estimates seven million premature deaths due to poor air quality (WHO, 2014).

[2] Air pollution also affects labour market outcomes. In particular, the literature provides evidence of air pollution affecting the productivity of agricultural workers (Graff Zivin and Neidell, 2012), the productivity of factory workers (Chang et al., 2016), and soccer players (Lichter et al., 2017). Importantly, the effects of outdoor ambient air pollution also have implications for indoor labour, affecting call center productivity (Chang et al., 2019), trading activity (Meyer and Pagel, 2017), decision time and quality of judges (Kahn and Li, 2020), and the performance of chess players (Künn et al., 2019).

traffic or industry-induced pollution during an academic year is correlated with lower test scores in subsequent exams, and with behavioural incidents during high school (Persico and Venator, 2019).

Similarly, there is mounting quasi-experimental evidence that exposure to extreme temperatures affect academic performance and health. Exposure to extremely high outdoor temperatures (above 90F), during testing and learning periods of high-school students in the US have been associated with lower tests scores in U.S. schools (Park, 2020; Park et al., 2020b), and globally (Park et al., 2020a). Building infrastructure plays a key role in the adaptation to climate change, reducing the health and performance damaging consequences of ambient extreme temperatures (Auffhammer, 2018; Kahn et al., 2014). Recent studies show that the presence of air conditioning in schools plays a key moderating role in the impact of extreme temperatures on student performance. In particular, those students learning or taking a test in an air-conditioned school show no adverse performance when local temperatures get very high (Park, 2020; Park et al., 2020b).[3]

Our results contribute to the existing literature that highlights the role of environmental conditions on academic outcomes in multiple ways. First, the overwhelming majority of studies use *outdoor* climate measurements to assess students' exposure, often using data from air quality or weather stations located miles away from the schools where the pupils are learning and taking their tests. [4] We collect data on temperature, air quality and other environmental metrics inside the classrooms, where our subjects are learning and taking exams, overcoming the challenge of measurement error which could result from miss-assigning environmental conditions to individuals (Moretti and Neidell, 2011; Roth, 2018). Second, we provide evidence on the impact of environmental conditions on children in elementary schools, a cohort in which the implications of exposure to poor air quality or extreme temperatures are still largely unexplored. The current evidence mostly relies on samples of high-school or university students.[5] Our estimates are based on individual standardized tests taken by children twice a year throughout all elementary school years (age 6 to 12), a critical age range for cognitive and human capital development (Howard-Jones et al., 2012; Heckman, 2006).

The rest of the paper is organized as follows. In Section 2, we describe the study design

[3]There is also quasi-experimental evidence on the role of air conditioning on moderating the mortality impacts of extreme temperatures (Barreca et al., 2016).

[4]A notable exception is Roth (2018), who deploys indoor sensors to measure the level of air particles (PM_{10}) during the exams of university students.

[5]Persico and Venator (2019) is a notable exception, investigating the impact of proximity to industrial sites or busy highways on the performance of elementary school students, studying the learning performance of children from grade 5 ($10-11$ years old).

and descriptive statistics of the main variables in the study. Section 3 describes the empirical strategy used to link indoor environmental conditions to student academic performance. Section 4 present the estimation results, and Section 5 concludes.

2 Data

2.1 Monitoring Environmental Conditions in Classrooms

We use data from a large-scale network of sensors deployed in 235 classrooms across 27 schools in the south of the Netherlands, with an average enrolment of over 4,000 students during our sample period. Panel A in Table 1 displays the summary statistics of number of groups in each school, and the number of students per classroom in our sample. These 27 schools represent a random sample of a larger school board that has 47 schools under its management.[6] Figure 1 shows the location of each school, as well as the average household income at the household level. The area in which the schools are located is generally considered a low-SES part of the Netherlands, with median net household incomes varying from EUR21.9-25.6k, compared to a national median household income of €25.8k.

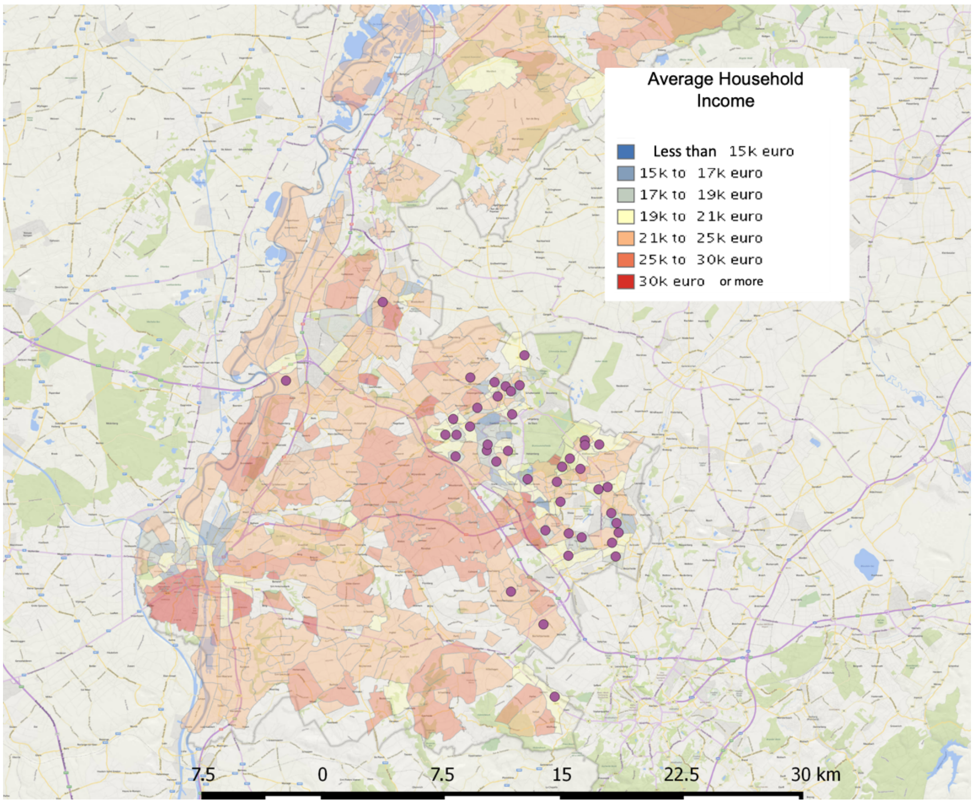
We use wall-mounted stationary sensors from the sensor company Aclima, Inc., to monitor the levels of CO_2 (ppm), fine particles (PM_{10}, counts/L), temperature (C), relative humidity (rH), light intensity (lux) and sound (dBA) in each classroom of our sample. The sensors capture raw data with a frequency of 1 to 30 seconds, transmitting all data to a cloud-based server. The data is aggregated at a minute level, and make available to the research team. We restrict the sample period to those official school days that the classroom is occupied, which can be inferred by the levels of CO_2 and noise in the room. The sensors monitor the levels of the six indoor environmental quality variables continuously throughout the year.

The deployment of sensors took place from January 2018 to December 2018. Figure 2 shows an example of a sensor installation. Each sensor is plugged into the wall for electricity and is connected to the local WiFi network for secure data transmission. [7]. Appendix Figure C.2 shows the daily statistics of sensor coverage per date as well as the time period covered by the sensors. The upper panel of that graph shows that the sensor coverage reaches full coverage in January 2019, upon completion of the sensor deployment, after which we rapidly deployed the full sensor network, which has been fully operational since that time.

[6]See Palacios et al. (2020) for the study protocol, including a detailed description of the sample and school typology, pre-analysis plan, and an extensive discussion of sensor placement and calibration.

[7]In some days there are sensors that do not deliver any data (typically the result of sensors that are unplugged during cleaning, etc.)

Figure 1: Map of School Locations and Average Income



Note: Figure 1 shows the location of each school, as well as the average household income in each zip code of the region in 2018.

Indoor environmental conditions vary across schools because of three main reasons: the existence and performance of the ventilation system at the school; the ambient temperature; and the interaction between ventilation type and temperature. The levels of CO_2 and to lesser extend PM_{10} are highly influence by the presence of ventilation systems in the school. In addition, mild temperatures allow the opening of windows, reducing the levels of indoor levels of CO_2. Figure 3 displays non-parametric comparisons between mechanically ventilated and conventional schools, over the distribution of ambient temperatures. Schools with non-mechanical ventilation show higher levels of maximum CO_2 concentrations, especially on relatively colder days. During warmer days, teachers can open windows to cool down the classroom, allowing CO_2 to leave the room. Variation in indoor temperature mostly depends on the local outdoor temperature.There are no schools with air-conditioning/cooling systems in our sample, as it is the case for the average school in the Netherlands. Finally, the levels of indoor PM_{10} are determined by the levels of outdoor PM_{10}, together with the presence of ventilation system. Those classrooms without ventilation system are more exposed to particles, given that their only mean of ventilate is opening windows.

We measure continuously indoor environmental conditions in all classrooms in the sample

Figure 2: Example school in the sample and sensor position in a classroom

Note: The picture in the top shows a of the school building in the sample. The picture in the bottom displays one of the indoor sensors in the sample. The sensors were deployed at breathing height (1.5 cm) in all classrooms in the sample.

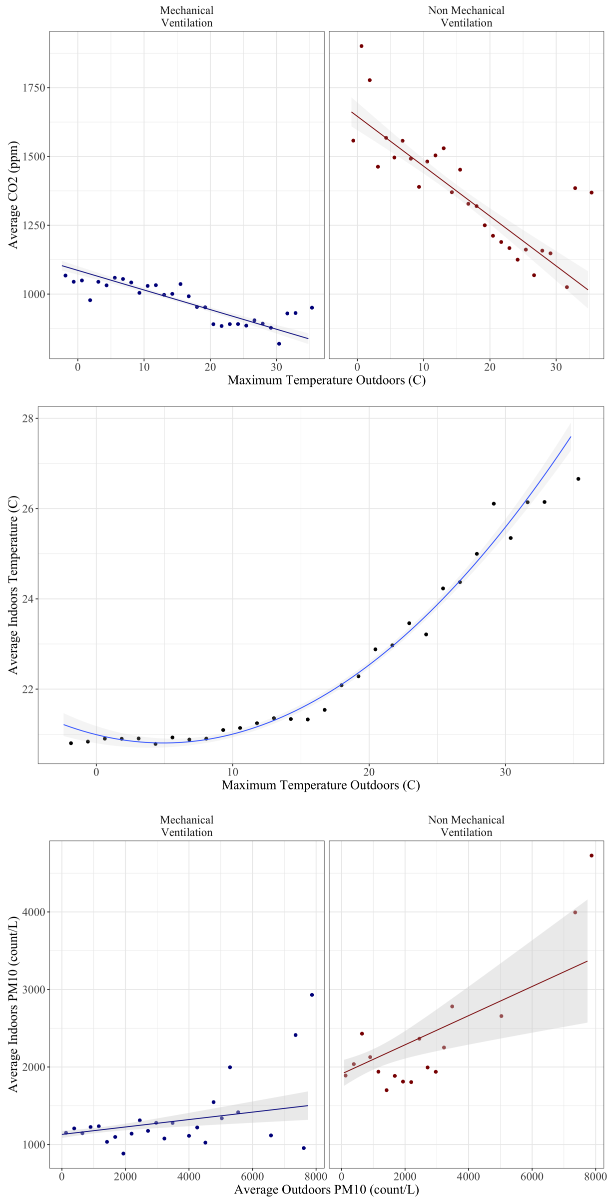
during every school day in the four months prior to a test taking place – the "Learning Period." For each school day and indoor environmental quality variable. We subsequently calculate the average levels over the entire learning period, as our main indicator of exposure during the school term. Panel B of Table 1 provides information on the indoor environmental quality variables obtained from the sensors during the times the children where inside the classroom.

2.2 Assessment of Children Exposure

Primary schools offer an ideal setting to study the impact of indoor environmental conditions on academic achievement. Students in primary schools spend most of their school day in the same classroom during the school year. This fixed allocation allows for individual students to be related very precisely to measurements of the indoor environmental conditions in their classroom. The allocation of students to classrooms in our sample of schools is predetermined by the school staff before the school year and exogenous to the kids.

In this study, we collected comprehensive information of each child in our sample, being able to identify the classroom assigned to the kid in each school term in our sample. In addition, we

Figure 3: Determinants of Indoor Environmental Conditions



Note: The figures display 3 bin-scatter plots linking outdoor (horizontal axis) and indoor conditions (vertical axis). Mechanical (non-mechanical) subplots describe the relationships for the sub-sample of schools with (without) ventilation system. The graph at the bottom plot PM_{10} counts/L indoors for the 5 schools closest located to Weerstation Parkstad (y-axis) and PM_{10} counts/L outdoors (x-axis) as measured by that weather station and discriminates schools by ventilation type.

are able assess daily exposure time more precisely, we compute the exact time (in minutes) that students spend inside the classroom during a school day. To determine whether the classroom is full or empty, we construct an algorithm that uses two regularities: (i) CO_2 increases (drops) after students enter (exit) an empty (full) classroom; (ii) students are noisy when they enter or exit the room.[8]. We use these two metrics to compute the number of entries and exits for each of the classrooms in our sample between 8am and 4pm, since Dutch primary schools do not start earlier than 8.30am, and end no later than 3.30pm.In Section C of the Appendix, we explain

[8]See Supplementary Figure C.3 for one example of the pattern of CO_2 and sound during a school day

Table 1: Summary Statistics

Panel A	Mean	Median	Std. Dev.	Min.	Max.
Groups per school	10	10	3	5	15
Students per group	22.3	22	5.6	8	35
Daily avg. minutes indoors	229.8	234.9	34.3	48.3	405
Daily avg. minutes outdoors	62.7	59.5	21.4	17.5	260
Panel B					
Daily max. CO_2	1,435.6	1,272.6	609.5	700.5	3,951.1
Daily avg. CO_2	1,087.7	998.9	368.9	599.6	2,725.4
Daily max. PM_{10}	2,319.1	2,160.2	1,036.5	326.7	6,164.5
Daily avg. PM_{10}	1,217.7	1,126.0	604.2	92.2	3,621.7
Daily max. temperature	22.3	22.3	0.9	19.4	25.9
Daily avg. temperature	21.2	21.2	0.9	17.7	24.3
Daily max. humidity	46.7	47.7	6.1	31.8	65.3
Daily avg. humidity	43.4	44.4	5.6	29.8	58.2

the algorithm in detail. Students in our sample leave the classroom during an school day for two main reasons: (1) breaks and (2) physical education, which is not part of the subjects that are part of the study.

Panel A in Table 1 displays the summary statistics of the predicted time spent indoors and outdoors during a school day in our sample. In our econometric analysis we include the total time a child spent in a classroom during the school term as a control. This allows to investigate the impact of indoor environmental conditions beyond changes in total teaching time during the learning period. In addition, we also estimate entry and exit to determine break lengths, which we use as a dependent variable in one of the secondary analyses.

2.3 Student Performance Data

In the Netherlands, student performance in elementary schools is tracked through semi-annual, nationally standardized tests, taking place halfway through the school year (January-February) and at the end of the school year (May-June). The tests cover a wide range of domains, including Mathematics, Reading, Spelling, and Vocabulary, and apply to students from kindergarten until 6th grade (in the Netherlands, grades correspond to groups, where Kindergarten is group 1, for 4 year-old kids, and 6th grade is group 8, for 12 year-old kids).

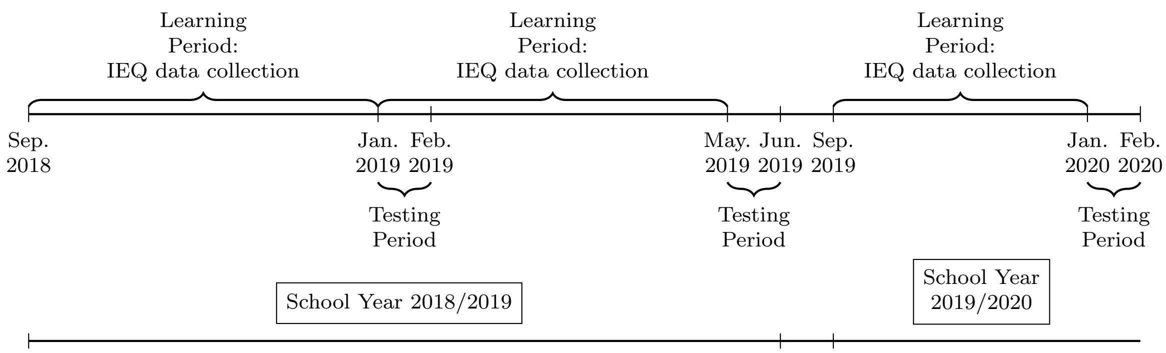
For each student in our sample, we obtained her scores in all tests over her entire primary school education through the *OnderwijsMonitor Limburg* – a collaboration between Maastricht

University and the elementary schools, school boards, municipalities in the province of Limburg (for more information, see Borghans et al. (2015a,0)). For the purpose of this study, we exclude testing data of Kindergarten (groups 1 and 2), given the relatively limited testing and comparability of test results to subsequent grades.

The tests are designed by a national examination center, administered at each individual school, and are graded by the teachers using a standardized grading scheme. The results are transformed into percentiles, but the norm for transformation is similar over time, such that results for the same student can be compared between periods. For each test period, we construct a comparable scale for each domain, standardizing the variable to have a mean of zero and a standard deviation of one within the relevant set of test scores. The main outcome variable is thus the standardized score for each student, in each test period, and domain (we only consider students that take the same test domain at least twice across different test periods).

We use the test-scores from three testing periods during two academic years: January-February 2019 and May-June 2019 (school year 2018/2019), and January-February 2020 (school year 2019/2020). Given the COVID-19-induced lockdown in Spring 2020, with children largely being outside of their regular classroom, we omit testing data from June 2020 in our data set (in addition, the learning effects of the lockdown may lead to noise in the estimation results, see Engzell et al. (2020)). The period preceding every testing period is considered the "Learning Period". Figure 4 shows a timeline for the data collection and data sets used in the study.

Figure 4: Study Timeline



Finally, to assess the direct effects of adverse indoor environmental quality conditions in the classroom at the time of taking tests, we collected the dates, start and end time for a sub-sample of tests. These data are gathered for the testing period January-February 2020 in collaboration with the janitors of each school, using a set of short surveys indicating test date, test domain, beginning and end time for each test in a given school on a voluntary basis.

3 Empirical Strategy

In our identification strategy, we exploit the variation over time in testing results for the same student, with varying levels of exposure to indoor environmental quality in the school term preceding the test (i.e. learning period). Using a fixed-effects approach, we remove the influence of confounding factors, driven by prior differences in student skills or socio-demographic background, classroom fixed infrastructure, test domain, and general changes in testing scores over time by estimating the following empirical model:

$$Score_{icdt} = B'IEQ_{itc} + \gamma_{id} + \lambda_{td} + \mu_c + \nu_\tau + \varepsilon_{it}, \tag{1}$$

where sub-index i indicates students, d test domains (i.e. subject), c classroom, and t the test period. Our main outcome variable is $Score_{icdt}$ which denotes the score for student i, classroom c, domain d, and test period t. γ_{id} is a fixed effect for each student by test domain, capturing idiosyncratic abilities of the pupil in that specific subject; λ_{dt} is a fixed effect indicating each domain by test period, controlling for common factors affecting all pupils taking the same subject in the same testing period. The classroom fixed effects μ_c control for all time-invariant characteristics of a classroom, such as views, or any relevant major (time invariant) teaching infrastructure in the room (e.g. digital boards). These capture confounding effects from the correlation among IEQ variables and between those and the time spent indoors. Finally, fixed effects for the test date ν_τ control for the outdoor environmental conditions in each specific day common to all students. Standard errors (ε_{ict}) are clustered at the classroom−period level to control for correlation among students within the same classroom taking tests in a given test period.

Our treatment variables are included in IEQ_{ict}, a vector including all the indoor environmental conditions. For each parameter within IEQ, we define cumulative exposure as the average mean values experienced during school days in the term prior to the test for all students i in classroom c taking the test in period t.

$$B'IEQ_{ict} = \beta_{CO_2}CO_{2ict} + \beta_{Temp}Temp_{ict} + \beta_{PM_{10}}PM_{10ict} + \beta_{Hum}Hum_{ict}. \tag{2}$$

where CO_2, $Temp_{ict}$, PM_{10ict}, Hum_{ict} describes the average levels of CO_2, temperature, fine particles and humidity in classroom c to which student i was exposed during the school term

t, respectively. Each variable enters the specification as the average daily levels over the term. Thus, the interpretation of the elements in vector B' are the changes in test scores associated with changes in indoor environmental conditions in the classroom where the student received the lessons in the school term preceding the test.

The identifying assumption of the analysis is that the variation in standardized test scores for each student is independent from other variables that might be correlated with indoor environmental quality. We assume that classroom equipment does not co-move over time with the indoor environmental quality in the classroom during our sample period. Given that teaching material and equipment is typically procured at the school level (if not the school board level), it is unlikely that classrooms vary widely on these dimensions.[9] In addition, we assume teachers do not sort into rooms based on indoor environmental quality, with high-quality teachers picking classrooms with healthier conditions. Our classroom fixed effects, partly control for this self-selection into classrooms, since those teachers that always teach in their favorite classroom will be part of the time non-varying aspects of the classroom, and therefore control by the classroom fixed effects. Moreover, it is the school principal who determines classroom allocation, rather than individual teachers.

4 Results

4.1 Indoor Environmental Quality and Learning Outcomes

4.1.1 Main Effects

We first analyze how changes in average exposure to indoor environmental conditions during the learning period affect the test scores of individual students.

Table 2 displays the estimated coefficients in Eq. 2. The coefficients are standardized to facilitate the comparisons across the different environmental factors. Poor air quality in the classroom during the school term lowers academic achievement. This result is robust to a variety of specifications. Table 2 presents the coefficients, introducing the set of fixed effects described in Eq. 2 sequentially. The first column in Table 2 includes only fixed-effects for the classroom, which we expand with school term, subject domain, student, period and domain, and student and domain fixed effects. Columns (5) and (6) incorporate the minutes spent that children spent during the learning period as a control. Column (6) uses only the minutes in which students

[9] A series of focus interviews between the research team and the board of the schools confirmed the lack of major changes in school furniture or equipment during our sample period.

are identified to be indoors during a given day to compute the average levels of IEQ variables over the school term. All specifications show that changes in CO_2 over time are associated with academic performance. In our preferred specification (Column (6)), being exposed to a one standard deviation increase in average CO_2 during the school term lowers academic achievement by 0.136 standard deviations.

Finally, the results indicate that the rest of environmental parameters considered in the study (i.e. PM_{10}, temperature, humidity) have, on average, no significant impact on academic achievement.

Table 2: Impact of Indoor Environmental Conditions on Student Achievement

	(1)	(2)	(3)	(4)	(5)	(6)
CO_2	-0.084*	-0.079*	-0.073***	-0.166***	-0.166***	-0.136***
	(0.046)	(0.044)	(0.017)	(0.043)	(0.043)	(0.029)
PM_{10}	0.011	0.037	-0.041	-0.012	-0.010	-0.010
	(0.056)	(0.053)	(0.074)	(0.063)	(0.064)	(0.082)
Temperature	0.006	-0.009	0.037	0.056	0.056	0.034
	(0.047)	(0.057)	(0.036)	(0.035)	(0.036)	(0.032)
Humidity	0.017	-0.071	0.114	0.035	0.032	0.035
	(0.019)	(0.090)	(0.089)	(0.071)	(0.076)	(0.090)
Mins. Indoors					0.012	0.008
					(0.028)	(0.033)
Obs.	14,349	14,349	14,349	14,349	14,349	14,349
R^2	0.055	0.066	0.585	0.855	0.855	0.854
Adj. R^2	0.047	0.051	0.481	0.698	0.698	0.698
Classroom FE	Y	Y	Y	Y	Y	Y
School term FE	N	Y	Y	Y	Y	Y
Subject Domain FE	N	Y	Y	Y	Y	Y
Student FE	N	N	Y	Y	Y	Y
Period by Domain FE	N	N	N	Y	Y	Y
Student by Domain FE	N	N	N	Y	Y	Y

Notes: The table presents the estimated coefficients described in Eq. 2. We standardized the coefficients to facilitate comparisons across environmental factors. Thus, all coefficients in the table describe how many standard deviations test scores ($Score_{icdt}$) will change per standard deviation increase in the corresponding environmental variable (IEQ_{ict}). Models (5) and (6) incorporate the minutes spent that children spent during the learning period as a control. Model (6) uses only the minutes that students are identified to be indoors during a given day to compute the average levels of IEQ variables over the school term. For a description of how we identify that the kids are in the classroom in a given minute, see Supplementary Section C. Standard errors in parentheses are clustered at the classroom and period levels to correct for correlation between scores within same group and between IEQ variables in the same classroom. Significance levels are *** 1%, ** 5%, * and 10%.

4.2 Non-linear effects

In this section we explore the presence of non-linearity in the estimated dose response function describing the impact of indoor environmental conditions on student performance.

We use a more flexible specification, that replaces the average exposure measures in our main specification with vectors of dummies, to explore non-linear effects over their distribution. Each dummy indicates a bin of the observed distribution for the full sample of the corresponding variable: CO_2, temperature, fine particles, and relative humidity. Each dummy equals 1 if, on average during the learning period, mean daily levels observed in the specific classroom where the pupil took lessons fall within the corresponding bin.[10] Dummies indicating each variable's third bin (values between the 40^{th} and 60^{th} percentiles) are left out, making observations surrounding the median percentile our reference level. Thus, the interpretation of the elements in vector B' related to a specific IEQ variable measure is relative to the medians observed of its average levels during the learning period in standard deviations. In all specifications displayed in this section, we use the full set of fixed effects.

Figure 5 present the results, depicting the coefficients for each bin value of the distribution of average daily levels of the indoor environmental quality variables. [11]

The results show that students exposed to relatively high levels of daily CO_2 concentrations performed considerably worse than those students with exposure to the lowest levels of CO_2 concentration in the sample. The effect of CO_2 concentration on learning outcomes are mainly driven by those children exposed to high levels. The figure shows an increase in the drop in test scores as we move to the higher values of CO_2. The economic effects of a detrimental indoor environment are quite substantial: in the extreme, children exposed to concentrations above the 80^{th} percentile ($>$ 1,500 ppm) the estimated impact of CO_2 reaches 0.4 of a standard deviation drop in test scores. In addition, we find evidence indicating that low average temperatures ($<$ 20.5C) also affect learning outcomes, with the magnitudes of the effect are smaller (less than half of the effects documented for CO_2 concentrations). Exposure to fine particles (PM_{10}) is insignificant for any concentration levels. However, the effect increases in the expected direction with higher concentrations.

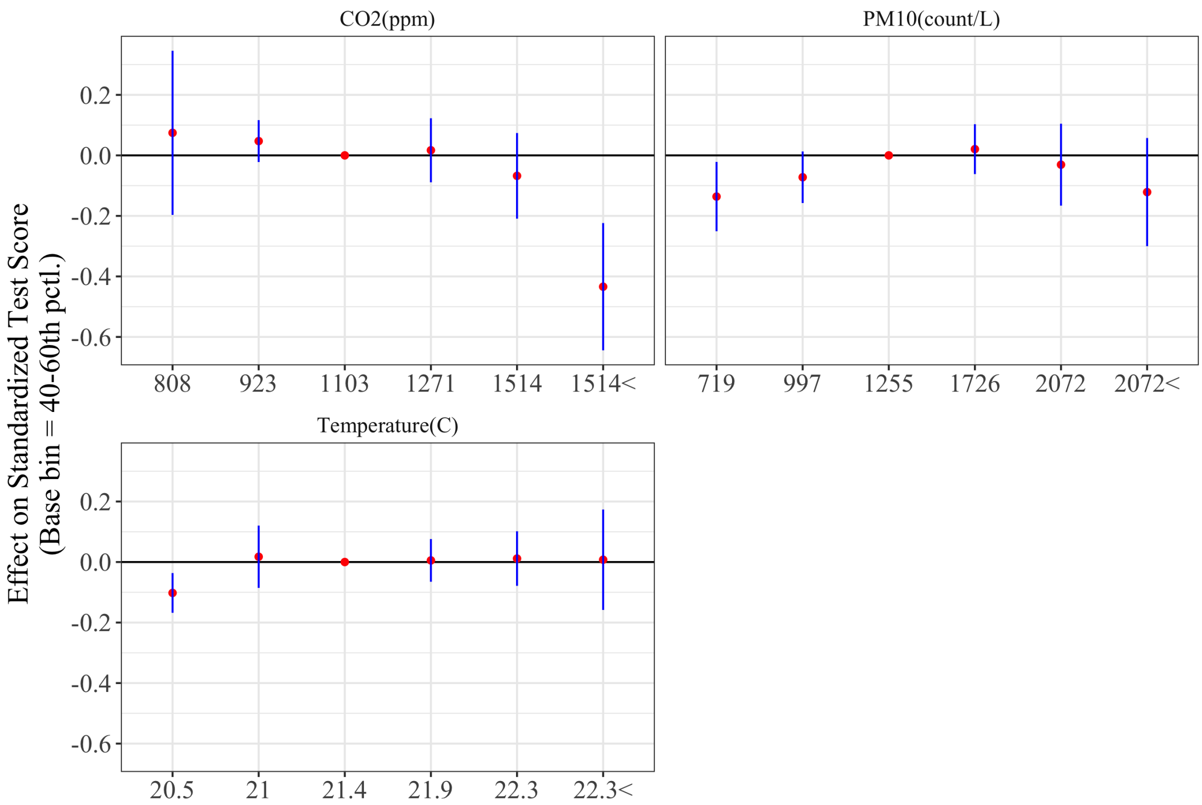
4.2.1 Effects by Test Domain

This section explores differences in the damage of poor indoor environmental conditions across different test domains – Mathematics, Reading and Spelling. Figure 6 presents the subgroup analysis by test domain to explore which specific domain determines the aggregate test results.

[10] The first four bins represent 20 percentile increments of the observed distribution, and the last bin represents a 10 percentile increment.

[11] These results are also presented in Section B of the Appendix in Tables B.2 and B.3 with additional results for inclusion of one IEQ variable at the time. Supplementary figure A.2 includes the results using maximum daily levels instead of average daily levels.

Figure 5: Average levels of IEQ observed during learning period and standard scores



Note: This figure plots point estimates (red dots) and 95% confidence levels (blue bars) for all coefficients corresponding to each IEQ variable (daily average values) in Equation (1) with standard errors clustered at the classroom and period level. Red dot without error bar describes the reference level. First four values on the x-axes give upper bound levels of 20 percentiles wide bins, and last value gives the 90^{th} percentile.

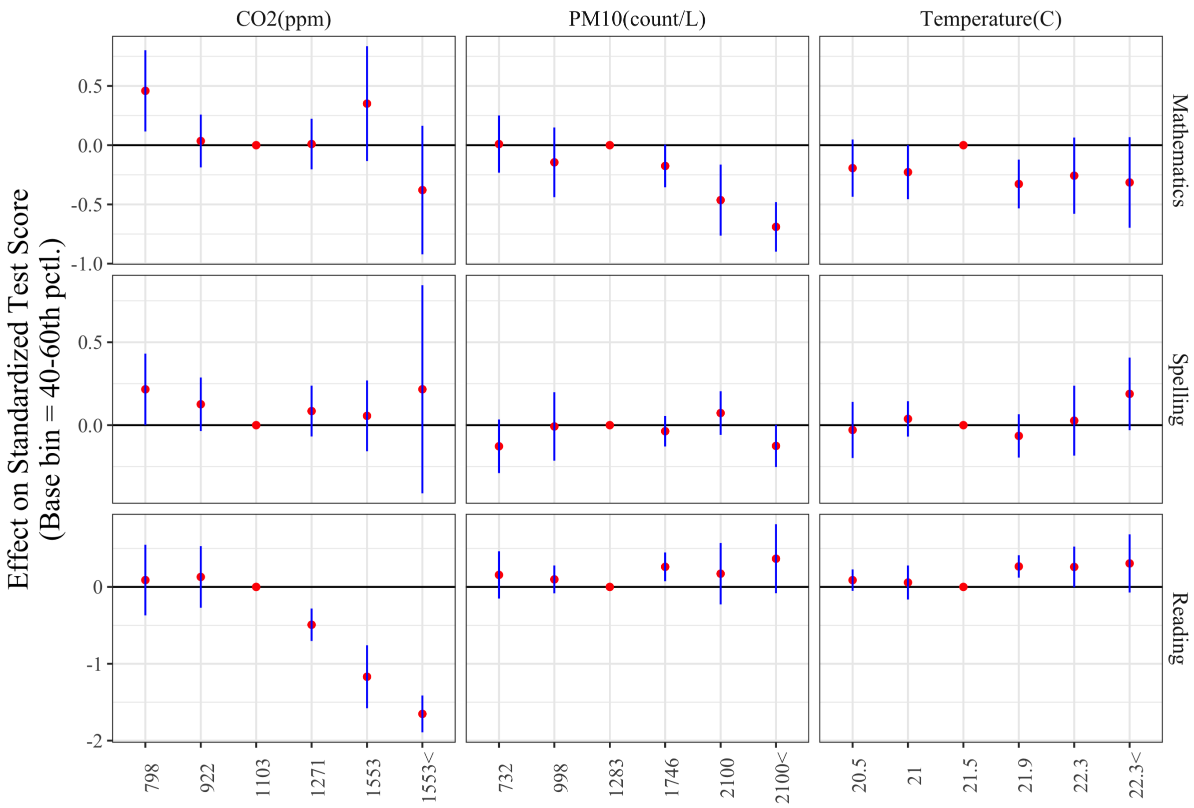
The figure shows that the negative effects we observe for CO_2 concentration on learning outcomes are mainly driven by the tests in mathematics and specially reading. In addition, we find evidence that fine particles (PM_{10}) mostly affects test scores in math exams. Results for temperature are noisier, but suggest that sub optimal temperatures harm the performance of students in mathematics, with a smaller coefficients, and only marginally significant ($p-value < 0.10$).

4.3 Indoor Environmental Quality and Test Results

In addition to affecting cognitive development of children over a prolonged period, the acute exposure to air pollution or extreme temperatures at the moment of taking a test have been documented to have an impact on test results Roth (2018); Park et al. (2020b)).

In this section, we estimate the effect of indoor environmental quality on standardized test scores, using high-frequency measurements during the test period. For a subsample of schools and classrooms, we gathered data on the exact times at which children were taking each of the

Figure 6: Average levels of IEQ observed during learning period and test scores by subject domain



Note: This figure plots point estimates (red dots) and 95% confidence levels (blue bars) for for all coefficients corresponding to each IEQ variable (daily average values) and for each test domain except for Vocabulary. Red dot without error bar describes the reference level. Standard errors are clustered at the classroom and period levels. First four values on the x-axes give the upper bound levels of 20 percentiles wide bins, and last value gives the $90^{th} percentile$.

four tests, analyzing the effect of maximum and average indoor environmental quality conditions on their scores. [12] For each group, we observe students taking multiple test during the period January-February 2020.

We identify the impact of environmental conditions on test scores using student and test domain fixed effects, hence controlling for idiosyncratic abilities. Moreover, we also introduce fixed effects for the time at which the test started controlling for possible confounding effects from e.g. having a test early in the morning or immediately after lunch. In our preferred specification we also control for the duration of the test to control for confounding effects such as fatigue or higher CO_2 levels from being longer inside the classroom. We use the following empirical model

[12] We were not able to obtain these data for all schools. The sample for which we do have these data covers 12 schools, 44 groups, 353 children, and 734 tests.

to link the contemporaneous indoor environmental conditions to tests scores:

$$Score_{id} = B' IEQ_{id} + TestLength_{dt} + \gamma_i + \lambda_d + \mu_h + \varepsilon_{id}, \tag{3}$$

where i indexes student, d test domain, and h the hour of the day the test started. $Test\_Length_{dt}$ describes the duration of the test. In line with our non-linear specification, the components of the vector IEQ_{id} are introduced in this specification as a set of dummies dividing the distribution in six equal parts. Student and test domain fixed effects are given by γ_i and λ_t, respectively, while μ_h are fixed effects for the starting time of the test (1 hour wide bins between 8:30 and 13:30). Each IEQ variable enters as the average level observed during the test. Table 3 give results for this analysis.

Table 3: Test Scores and IEQ conditions during test

	(1)	(2)	(3)	(4)	(5)
CO_2	-0.039	-0.029	-0.047	-0.016	0.010
	(0.079)	(0.060)	(0.048)	(0.058)	(0.051)
PM_{10}	0.025	0.057	0.053	0.022	-0.012
	(0.123)	(0.111)	(0.081)	(0.097)	(0.109)
Temperature	-0.032	-0.075	-0.202**	-0.267**	-0.263***
	(0.044)	(0.058)	(0.085)	(0.109)	(0.099)
Humidity	-0.014	-0.062	-0.033	-0.026	0.001
	(0.053)	(0.081)	(0.079)	(0.081)	(0.082)
Duration of exam					0.163
(in mins)					(0.179)
Obs.	734	734	734	734	734
R^2	0.091	0.094	0.102	0.703	0.705
Adj. R^2	0.065	0.066	0.069	0.415	0.417
Group FE	Y	Y	Y	Y	Y
Domain FE	N	Y	Y	Y	Y
Start Time FE	N	N	Y	Y	Y
Student FE	N	N	N	Y	Y

Notes: This table reports results for a linear model for standardized test scores on average of IEQ variables as measured at the time that students were taking the tests. The coefficients associated with environmental factors are standardized to allow for comparisons across them. Standard errors (in parentheses) are clustered at the group level. Significance levels are *** 1%, ** 5%, and * 10%.

Our results suggest that, during the test, the important factor in indoor environmental conditions limiting students' achievement is temperature. A one standard deviation increase in the average temperature while taking the test is associated to a 0.26 standard deviation loss in test scores. Our results provide supportive evidence on the effects of indoor temperatures for a raft of papers that use outdoor temperature to assess the effect on test scores (e.g. Park et al. (2020b)).

The remaining IEQ variables show no significant effects on test scores. In particular, the effects of contemporaneous CO_2 concentrations during the test period on test scores lack statistical significance. This contrasts with recent evidence from lab experiments, showing that elevated levels of CO_2 have a strong impact on cognitive ability (Allen et al., 2016). However, it is important to note the small sample available for the analysis, and therefore limited power of the tests.

4.4 Time in the Classroom

In this section, we investigate changes in teaching time as a potential mechanism for the impacts of exposure to poor environmental conditions during the school term on tests scores documented in this paper. In response to poor environmental conditions, teachers may decide to let students leave the room earlier if they become restless, or give the students (and themselves) a longer break. Thus, bad indoor environmental quality could crowd out learning activities in the classroom.

Most of the variation in the time that children spend outside of the classroom during breaks at any given day is predetermined by school schedules. There are multiple scheduled breaks every day that every group in a school has to take, such as the lunch break, where teachers have no discretion in deciding either the starting time or its duration. However, there are other breaks in our sample where teachers have discretion on start and length, such as the mid morning break. We focus our analysis on these breaks, seeking to explain how variation in indoor environmental conditions explains time spent outside of the classroom.

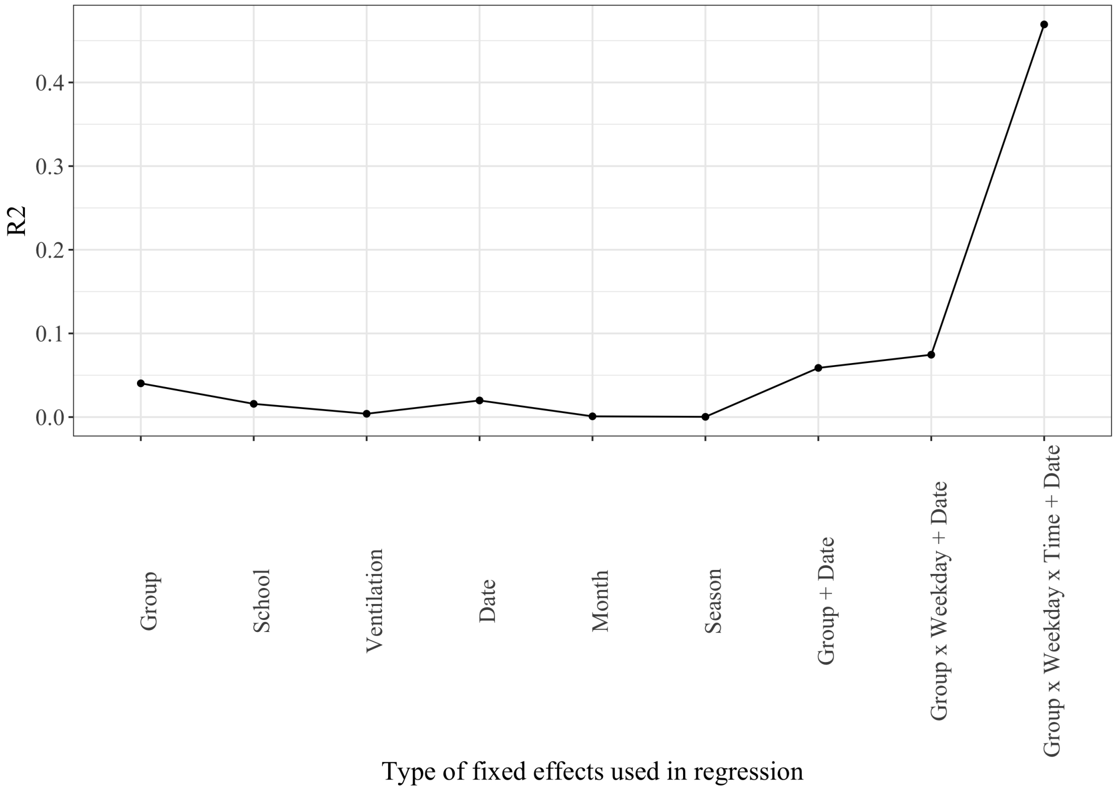
Figure 7 shows the predictive power of different sets of fixed effects for the explaining time in breaks. Almost half of the variation in break length is explained by observing which classroom is taking a break and when– the group in a given school, day of the week, and during what time of the day (time is defined as half hour bins during the school day).

In the analysis, we link the environmental conditions in the hour immediately before the beginning of a break, with the length of the break using the following specification:

$$\log(Break\ Length_{gdh}) = IEQ_{gdh} + Length\_Lesson_{gdh} + \mu_{gw(d)h} + \alpha_{dh} + \varepsilon_{gdh} \tag{4}$$

where the subindex g indicate school-group, d date, and the time of the day where the break takes place h. In addition, $w(d)$ indicates the weekday corresponding to the specific date d. $Break\ Length_{gdh}$ gives the total minutes that children spent outside during the break. Again,

Figure 7: Variation in time spent at breaks explained with fixed effects

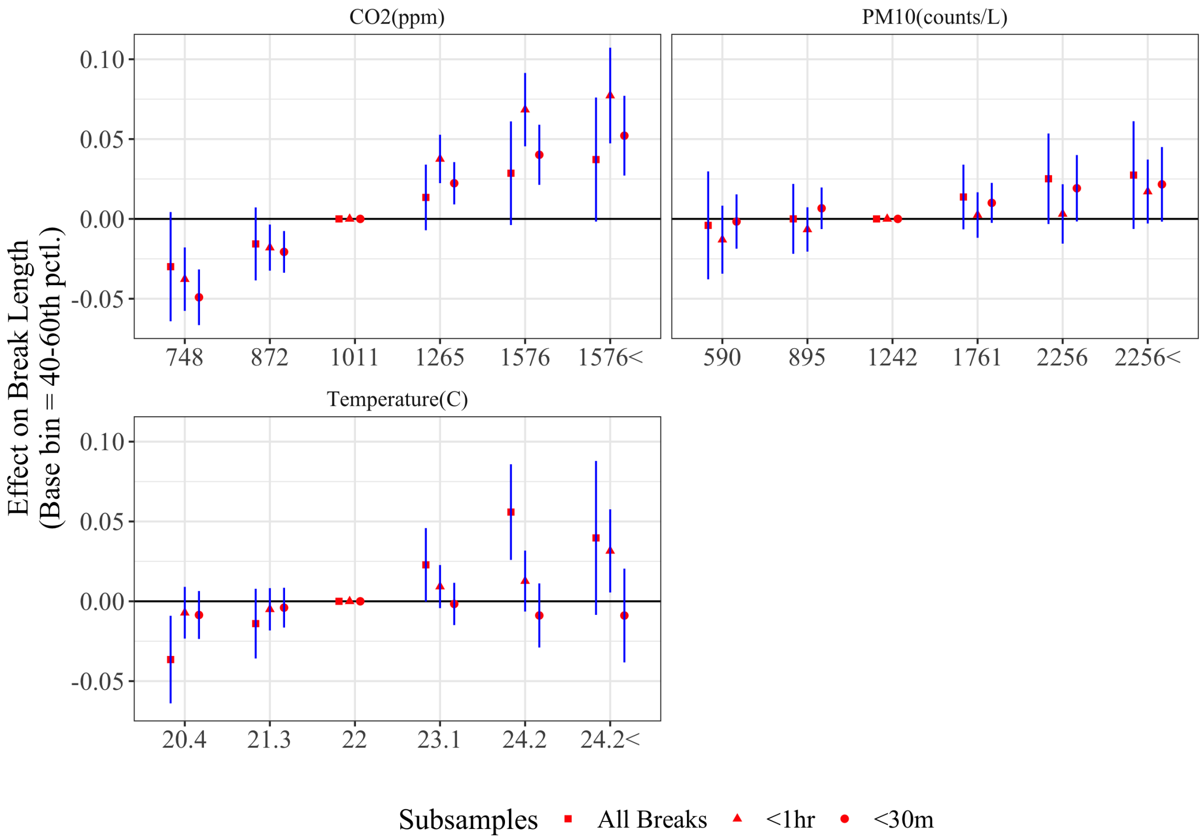


$Length\_Lesson_{gdh}$ gives the length of the lesson immediately previous to the break controlling for having longer breaks because of longer lessons; $\mu_{gw(d)h}$ are fixed effects for each school-group by day of the week and by half hour bin which capture the usual lengths of breaks scheduled for specific hours in specific days of the week; and α_{dh} are the fixed effects associated with the time and date, controlling for events taking place across all schools during specific hours of specific dates.

Figure 8 displays the results for the βs estimates with their 95% confidence intervals. Results show that pupils exposed to average levels of CO_2 higher than 1,200 ppm subsequently have breaks that are between 4% to 8% longer (depending on the subsample analyzed) than those taken by pupils exposed to average concentrations between 900 and 1000 ppm (reference level). Similarly, students exposed to CO_2 concentrations below 750 ppm take breaks that are 4% to 5% shorter than the same reference group.

In addition, higher temperatures during the lesson also seem to increase break lengths. For longer breaks, temperatures above 23°C induce breaks that are 5% longer on average, while the effect is similar for more than 24°C when breaks are shorter than an hour. No significant effects are found for temperature on breaks that are shorter than half an hour. When adding the results of temperature and CO_2, results suggest that students exposed to a combination of high temperature and elevated CO_2 concentrations spend around 4 minutes less in the classroom in a given day. [13] Particulate matter has a lower impact and only marginally significant, although there is a clear positive trend in the effect as concentration increases, with effects on the break length around 2.5% at more than 2,200 PM_{10}/L.

Figure 8: Average levels of IEQ observed during lesson and length of subsequent break.



Note:This figure plots point estimates (red dots) and 95% confidence levels (blue bars) for for all coefficients corresponding to each IEQ variable (average during lesson before break) and for subsamples of shorter breaks (i.e. 1 hour and 30 minutes long). Red dots without error bars describe the reference level. Estimation results are also included in Appendix Tables B.4 and B.5. Standard errors are clustered at the classroom and date levels. First four values on the x-axes give the upper bound levels of 20 percentiles wide bins, and last value gives the $90^{th} percentile.$

5 Conclusion and Discussion

This paper provides the first quasi-experimental evidence on how environmental conditions in the classroom affect learning outcomes. We design and implement a large field study, deploying an indoor sensing network in 27 elementary schools, continuously monitoring the indoor environmental conditions in 139 individual classrooms, covering some 3,000 children aged 6 to

[13] As an upper bound, since the academic year in the Netherlands amounts to 215 school days or 30 weeks, students in the chronically worst CO_2 conditions receive up to 882 minutes of in-class lessons as compared to students that are exposed to the lowest CO_2 concentrations and temperature, during one academic year.

12. We document large variation in indoor environmental condition across schools, classrooms and over time. Such variance cannot be explained (add here percentage) by classroom or school characteristics, or ambient weather conditions – strengthening the need to measure those factors indoors, and on a continuous basis. Using within-child variation in exposure to environmental conditions, we document that when children are systematically exposed to poor air quality during a prolonged learning period, their cognitive development is slowed down, as measured by the improvement on standardized tests made by the same student. Effects are largest for exposure to elevated CO_2 levels: above 1,500 parts per million, test scores are lower by 0.40 standard deviations. Besides the effects associated with medium-term exposure, we also find that relatively high temperatures during tests lead to detrimental outcomes of pupils taking these tests. Finally, our findings document behavioral reactions to adverse climate conditions, that might act as a mediating channel – classes that have been exposed to adverse indoor climate conditions during a part of the day have longer subsequent breaks.

To put our finding into perspective, we can compare them to the effects documented for other school interventions, for example, the Active Living Program (Golsteyn et al., 2020) or the Early Training Program (Anderson, 2008). Of course, these intervention programs are not directly comparable, given the non-experimental setting of this paper, but the comparison of average treatment effects is nonetheless helpful in understand the magnitude of the documented results. The effects of the Active Living Program, an experiment aimed at stimulating physical activity on Dutch primary schools, showed a negative effect of 0.06 standard deviations, using a comparable sample and setting. Clearly, the effects of *passive* exposure to poor environmental conditions in the classroom trump the negative effects of a program aimed at stimulating physical activity (see Golsteyn et al. (2020)). Anderson (2008) documents average impact outcomes across programs of 0.28 standard deviations, while Deming (2009) documents an average impact of 0.23 standard deviations for the Head Start program. Although the setting of these programs is fundamentally different than ours, these simple comparisons provide some indication that the economic magnitude of the findings in this paper is quite substantial.

This study highlights the understudied role that school facility conditions play in generating educational outcomes. It also highlights the extent to which disparities in such physical environments contribute to inequality in educational outcomes. The results allow us to estimate the benefits of potential public investment, for example in school HVAC systems, that may help reduce such gaps. A variety of recent papers have examined connections between investment in school facility and academic achievements, but the mechanisms explaining that connection remain a matter of speculation. In addition, this paper complements the growing literature on the impacts of pollution, temperature, and academic performance. The existing evidence depends on measurements of outdoor temperature or ambient air quality to assess the relationship between environmental conditions and student performance, relying on strong assumptions to extrapolate outdoor measurements to indoor exposure. The use of sensors inside the classrooms where children are learning and being tested allows us to assess exposure to specific pollutants with an unprecedented accuracy.

Our results also yield several important policy lessons. The airborne transmission of the SARS-COV-2 virus has elevated the salience of indoor environmental quality as an important factor to prevent the spread of the disease. Many countries have begun preparing major investment outlays to improve ventilation, via modernization or installation of HVAC systems, or upgrading the standards of ventilation in buildings. Schools buildings are among the major targets in many nations' building portfolios, due to the high density of children in classrooms, and the long need for disrepair or installation of schools' air treatment systems. Our results suggest that upgrades in indoor air quality in schools is relevant besides reduction of the spread of viral diseases, such as COVID-19, and supports children's cognitive capacities, precisely at a critical age for human capital accumulation and skill formation.

Next Steps: This is an ongoing study, that will be extended in multiple ways. First, our monitoring is still in process. The sensors will collect data for two extra academic years, doubling the statistical power of the tests presented in the current version, and allowing for a more precise estimation of each parameter's coefficient. In the short term, the research team is undertaking a new exam data collection of the exact time, location and dates of all tests taking place in February 2021. In this version of the paper, we have not yet included heterogeneous effects across different individual characteristics, but we intend to do so in the next version of the paper. In particular, we are planning to explore two key dimensions. First, our cohort is at a critical stage of brain development. The literature shows that children between ages 6 and 12 have rapidly evolving brains, developing from extreme plasticity to faster learning ability. It is therefore crucial to study the salience of indoor environmental quality for learning outcomes across the age groups covered in our sample. Second, there are important inequality dimensions that are important to understand for their social justice implications. In particular, we will test whether children who are already weaker in terms of socio-economic background and school performance than other children are more affected by poor environmental conditions in the classroom – thereby casting the existing inequality in concrete. That is why we will compare learning effects across groups of children with different ex-ante cognitive performance and socio-economic characteristics (e.g. household income). Finally, the high-frequency character of our environmental measures, at a minute frequency, allows to incorporate in the analysis different moments of the distribution of environmental conditions (e.g. intra-day variance), and decompose the effect measures of environmental conditions over the school term (e.g. last month vs first month of the term) and over the school day.

References

Aaronson, D. and Mazumder, B. (2011). The impact of rosenwald schools on black achievement. Journal of Political Economy, 119(5):821–888.

Allen, J. G., MacNaughton, P., Satish, U., Santanam, S., Vallarino, J., and Spengler, J. D. (2016). Associations of cognitive function scores with carbon dioxide, ventilation, and volatile organic compound exposures in office workers: a controlled exposure study of green and conventional office environments. Environmental health perspectives, 124(6):805–812.

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

Auffhammer, M. (2018). Quantifying economic damages from climate change. Journal of Economic Perspectives, 32(4):33–52.

Barreca, A., Clay, K., Deschenes, O., Greenstone, M., and Shapiro, J. S. (2016). Adapting to climate change: The remarkable decline in the us temperature-mortality relationship over the twentieth century. Journal of Political Economy, 124(1):105–159.

BBC (2020). Coronavirus: Germany improves ventilation to chase away covid.

Borghans, L., Golsteyn, B. H., and Stenberg, A. (2015a). Does expert advice improve educational choice? PloS one, 10(12):e0145378.

Borghans, L., Golsteyn, B. H., and Zölitz, U. (2015b). Parental preferences for primary school characteristics. The BE Journal of Economic Analysis & Policy, 15(1):85–117.

Chang, T., Graff Zivin, J., Gross, T., and Neidell, M. (2016). Particulate pollution and the productivity of pear packers. American Economic Journal: Economic Policy, 8(3):141–69.

Chang, T. Y., Graff Zivin, J., Gross, T., and Neidell, M. (2019). The effect of pollution on worker productivity: evidence from call center workers in china. American Economic Journal: Applied Economics, 11(1):151–72.

Chay, K. Y. and Greenstone, M. (2003). The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. The quarterly journal of economics, 118(3):1121–1167.

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

Currie, J. and Neidell, M. (2005). Air pollution and infant health: what can we learn from california's recent experience? The Quarterly Journal of Economics, 120(3):1003–1030.

Currie, J. and Walker, R. (2011). Traffic congestion and infant health: Evidence from e-zpass. American Economic Journal: Applied Economics, 3(1):65–90.

Deming, D. (2009). Early childhood intervention and life-cycle skill development: Evidence from head start. American Economic Journal: Applied Economics, 1(3):111–34.

Du, B., Tandoc, M. C., Mack, M. L., and Siegel, J. A. (2020). Indoor co2 concentrations and cognitive function: A critical review. Indoor air, 30(6):1067–1082.

Duflo, E. (2001). Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. American economic review, 91(4):795–813.

Ebenstein, A., Lavy, V., and Roth, S. (2016). The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution. American Economic Journal: Applied Economics, 8(4):36–65.

Engzell, P., Frey, A., and Verhagen, M. D. (2020). Learning inequality during the covid-19 pandemic.

Fisk, W. J. (2017). The ventilation problem in schools: literature review. Indoor Air, 27(6):1039–1051.

G.A.O. (2020). K-12 education. school districts frequently identified multiple building systems needing updates or replacement.

Golsteyn, B. H., Jansen, M. W., Van Kann, D. H., and Verhagen, A. M. (2020). Does stimulating physical activity affect school performance? Journal of Policy Analysis and Management, 39(1):64–95.

Graff Zivin, J. and Neidell, M. (2012). The impact of pollution on worker productivity. American Economic Review, 102(7):3652–73.

Graff Zivin, J. and Neidell, M. (2013). Environment, health, and human capital. Journal of Economic Literature, 51(3):689–730.

Graff Zivin, J. and Neidell, M. (2018). Air pollution's hidden impacts. Science, 359(6371):39–40.

Graff Zivin, J., Song, Y., Tang, Q., and Zhang, P. (2020). Temperature and high-stakes cognitive performance: evidence from the national college entrance examination in china. Journal of Environmental Economics and Management, 104:102365.

Hanushek, E. A. (2003). The failure of input-based schooling policies. The economic journal, 113(485):F64–F98.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. Science, 312(5782):1900–1902.

Howard-Jones, P. A., Washbrook, E. V., and Meadows, S. (2012). The timing of educational investment: A neuroscientific perspective. Developmental Cognitive Neuroscience, 2:S18–S29.

Jackson, C. K., Johnson, R. C., and Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. The Quarterly Journal of Economics, 131(1):157–218.

Kahn, M. E., Kok, N., and Quigley, J. M. (2014). Carbon emissions from the commercial building sector: The role of climate, quality, and incentives. Journal of Public Economics, 113:1–12.

Kahn, M. E. and Li, P. (2020). Air pollution lowers high skill public sector worker productivity in china. Environmental Research Letters, 15(8):084003.

Knittel, C. R., Miller, D. L., and Sanders, N. J. (2016). Caution, drivers! children present: Traffic, pollution, and infant health. Review of Economics and Statistics, 98(2):350–366.

Künn, S., Palacios Temprano, J. F., and Pestel, N. (2019). Indoor air quality and cognitive performance. IZA Discussion Paper Series, (12632).

Lichter, A., Pestel, N., and Sommer, E. (2017). Productivity effects of air pollution: Evidence from professional soccer. Labour Economics, 48:54–66.

Liu, C., Chen, R., Sera, F., Vicedo-Cabrera, A. M., Guo, Y., Tong, S., Coelho, M. S., Saldiva, P. H., Lavigne, E., Matus, P., et al. (2019). Ambient particulate air pollution and daily mortality in 652 cities. New England Journal of Medicine, 381(8):705–715.

Martorell, P., Stange, K., and McFarlin Jr, I. (2016). Investing in schools: capital spending, facility conditions, and student achievement. Journal of Public Economics, 140:13–29.

Meyer, S. and Pagel, M. (2017). Fresh air eases work–the effect of air quality on individual investor activity. Technical report, National Bureau of Economic Research.

Moretti, E. and Neidell, M. (2011). Pollution, health, and avoidance behavior evidence from the ports of los angeles. Journal of human Resources, 46(1):154–175.

Neilson, C. A. and Zimmerman, S. D. (2014). The effect of school construction on test scores, school enrollment, and home prices. Journal of Public Economics, 120:18–31.

Palacios, J., Eichholtz, P., Willeboordse, M., and Kok, N. (2020). Indoor environmental quality and learning outcomes: protocol on large-scale sensor deployment in schools. BMJ open, 10(3):e031233.

Park, R. J. (2020). Hot temperature and high stakes performance. Journal of Human Resources.

Park, R. J., Behrer, A. P., and Goodman, J. (2020a). Learning is inhibited by heat exposure, both internationally and within the united states. Nature human behaviour, pages 1–9.

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

Persico, C. L. and Venator, J. (2019). The effects of local industrial pollution on students and schools. Journal of Human Resources, pages 0518–9511R2.

Rijksoverheid (2020). 360 miljoen voor verbetering ventilatie in scholen.

Roth, S. (2018). The effect of indoor air pollution on cognitive performance: Evidence from the uk.

Seppanen, O., Fisk, W. J., and Lei, Q. (2006). Effect of temperature on task performance in office environment.

Stafford, T. M. (2015). Indoor air quality and academic performance. Journal of Environmental Economics and Management, 70:34–50.

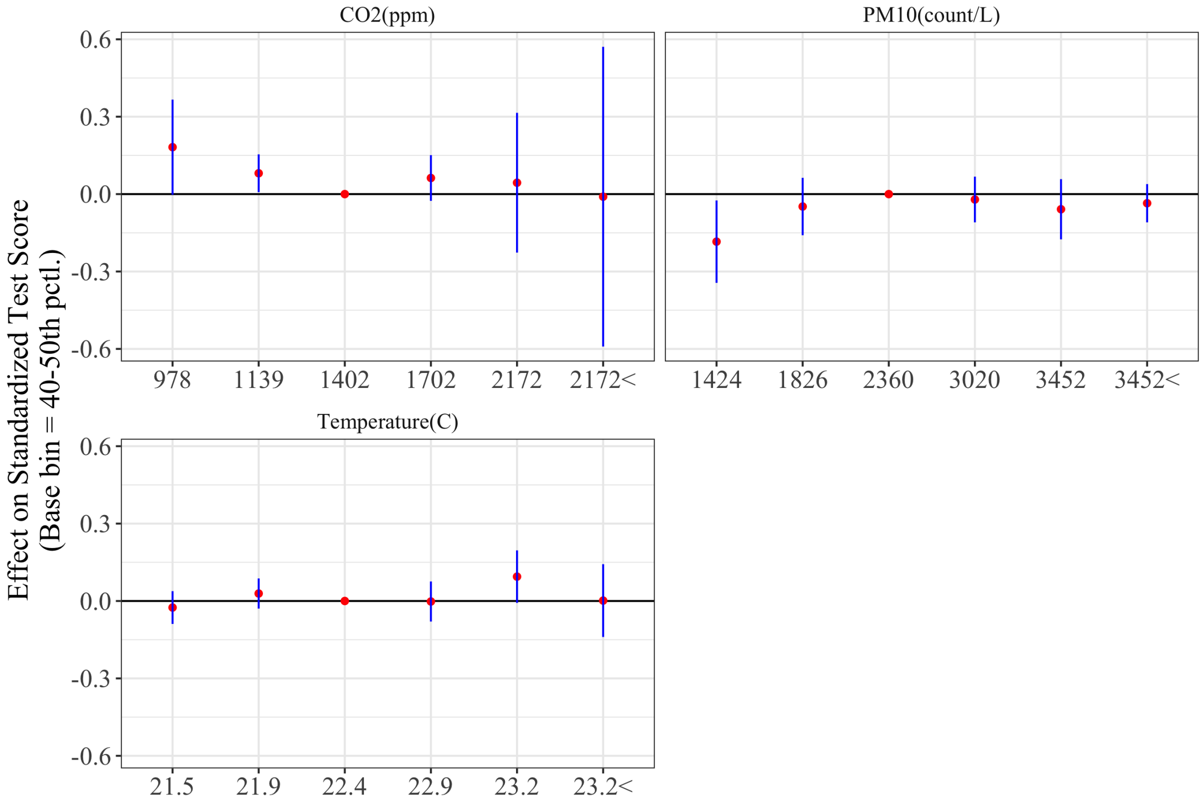
WHO (2014). 7 million premature deaths annually linked to air pollution. World Health Organization, Geneva, Switzerland.

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

Appendix

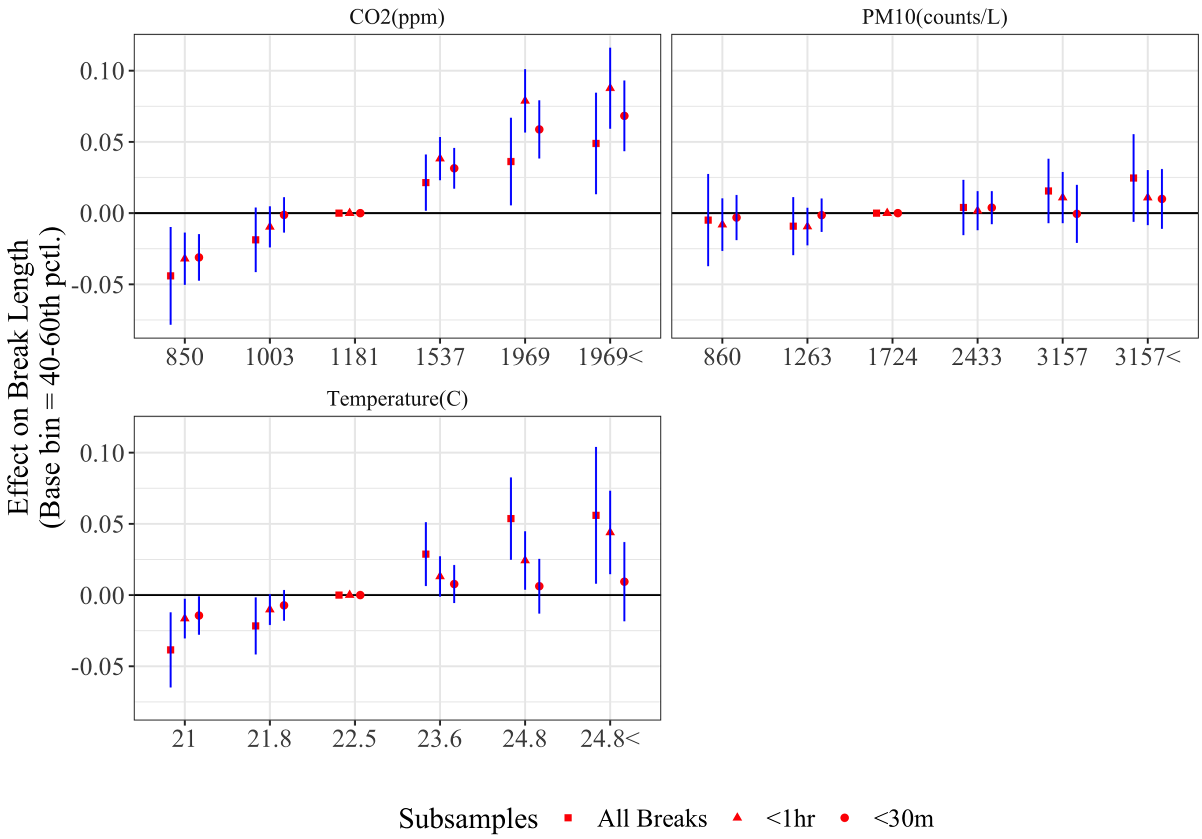
A Analysis Using Daily Maximum Levels

Figure A.2: Maximum levels of IEQ observed during learning period and standard scores



Note: This figure plots point estimates (red dots) and 95% confidence levels (blue bars) for all coefficients corresponding to each IEQ variable (daily maximum values) in Equation (1) with standard errors clustered at the classroom and period level. First four values on the x-axes give the upper bound levels of 20 percentiles wide bins and last value gives 90th percentile.

Figure A.3: Maximum levels of IEQ observed during lesson and length of subsequent break.



Note: This figure plots point estimates (red dots) and 95% confidence levels (blue bars) for for all coefficients corresponding to each IEQ variable (maximum during lesson before break) and for subsamples of shorter breaks (i.e. 1 hour and 30 minutes long). Standard errors are clustered at the classroom and date levels. First four values on the x-axes give the upper bound levels of 20 percentiles wide bins, and last value gives the $90^{th} percentile$.

B Regression tables in main text

Table B.2: Standardized scores and average levels of IEQ variables during learning period

	(1)	(2)	(3)	(4)	(5)
CO2 >0.9	-0.314***				-0.434***
	(0.104)				(0.141)
CO2 0.9	-0.040				-0.068
	(0.081)				(0.081)
CO2 0.2	0.107				0.075
	(0.108)				(0.112)
PM10 >0.9		-0.097**			-0.121**
		(0.046)			(0.056)
PM10 0.9		0.054			-0.031
		(0.058)			(0.072)
PM10 0.2		-0.137***			-0.136*
		(0.050)			(0.074)
Temp. >0.9			0.017		0.008
			(0.080)		(0.078)
Temp. 0.9			0.027		0.012
			(0.078)		(0.071)
Temp. 0.2			-0.064		-0.102*
			(0.055)		(0.053)
Hum. >0.9				-0.031	-0.060
				(0.098)	(0.091)
Hum. 0.9				0.058	-0.054
				(0.076)	(0.088)
Hum. 0.2				-0.005	0.034
				(0.076)	(0.069)
Obs.	14,349	14,349	14,349	14,349	14,349
R^2	0.855	0.855	0.855	0.855	0.856
Adjusted R^2	0.699	0.699	0.698	0.698	0.700
Classroom FE	Y	Y	Y	Y	Y
Group FE	Y	Y	Y	Y	Y
Grade FE	Y	Y	Y	Y	Y
Period by Domain FE	Y	Y	Y	Y	Y
Student by Domain FE	Y	Y	Y	Y	Y

Notes: This table reports results for the model including all binned IEQ variables (only lower and two highest bins shown) for daily averages observed during learning period as in Equation (1). The model includes fixed effects for student by domain, period by domain, and classroom, and incorporate controls for average minutes spent indoor during learning period. All standard errors (in parentheses) are clustered at the classroom and period level. Significance levels are $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.3: Maximum IEQ conditions during learning period and tests scores

	(1)	(2)	(3)	(4)	(5)
CO2 >0.9	-0.059				-0.010
	(0.164)				(0.296)
CO2 0.9	-0.031				0.044
	(0.064)				(0.138)
CO2 0.2	0.169**				0.182*
	(0.081)				(0.094)
PM10 >0.9		-0.043			-0.035
		(0.035)			(0.038)
PM10 0.9		-0.032			-0.059
		(0.047)			(0.060)
PM10 0.2		-0.171*			-0.184**
		(0.096)			(0.081)
Temp. >0.9			0.074		0.002
			(0.092)		(0.072)
Temp. 0.9			0.085*		0.094*
			(0.045)		(0.052)
Temp. 0.2			-0.010		-0.025
			(0.024)		(0.032)
Hum. >0.9				0.061	0.024
				(0.115)	(0.107)
Hum. 0.9				0.017	-0.062
				(0.081)	(0.075)
Hum. 0.2				0.052	0.152*
				(0.088)	(0.083)
Obs.	14,349	14,349	14,349	14,349	14,349
R^2	0.855	0.855	0.855	0.855	0.856
Adj. R^2	0.698	0.698	0.698	0.699	0.699
Classroom FE	Y	Y	Y	Y	Y
Group FE	Y	Y	Y	Y	Y
Grade FE	Y	Y	Y	Y	Y
Period by Domain FE	Y	Y	Y	Y	Y
Student by Domain FE	Y	Y	Y	Y	Y

Notes: This table reports results for the model including all binned IEQ variables (only lower and two highest bins shown) for daily maximums observed during learning period as in Equation (1). The model includes fixed effects for student by domain, period by domain, and classroom, and incorporate controls for average minutes spent indoor during learning period. All standard errors (in parentheses) are clustered at the classroom and period level. Significance levels are $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.4: Average IEQ conditions during lesson and break length

Percentile	(1)	(2)	(3)	(4)	(5)	< 1h (6)	< 30m (7)
CO2 >0.9	0.032				0.037*	0.077***	0.052***
	(0.021)				(0.020)	(0.015)	(0.013)
CO2 0.9	0.025				0.029*	0.068***	0.040***
	(0.017)				(0.017)	(0.012)	(0.010)
CO2 0.2	-0.028				-0.030*	-0.038***	-0.049***
	(0.018)				(0.017)	(0.010)	(0.009)
PM10 >0.9		0.019			0.027	0.017*	0.022*
		(0.017)			(0.017)	(0.010)	(0.012)
PM10 0.9		0.022			0.025*	0.003	0.019*
		(0.014)			(0.014)	(0.009)	(0.011)
PM10 0.2		-0.003			-0.004	-0.013	-0.002
		(0.018)			(0.017)	(0.011)	(0.009)
Temp. >0.9			0.036		0.040	0.032**	-0.009
			(0.027)		(0.025)	(0.013)	(0.015)
Temp. 0.9			0.053***		0.056***	0.013	-0.009
			(0.016)		(0.015)	(0.010)	(0.010)
Temp. 0.2			-0.037***		-0.036***	-0.007	-0.009
			(0.014)		(0.014)	(0.008)	(0.008)
Hum. >0.9				0.044*	0.042*	0.022	0.003
				(0.027)	(0.025)	(0.017)	(0.015)
Hum. 0.9				0.016	0.016	0.001	-0.002
				(0.019)	(0.019)	(0.013)	(0.012)
Hum. 0.2				0.002	0.005	0.010	0.016
				(0.020)	(0.020)	(0.013)	(0.013)
Obs.	58,219	58,219	58,219	58,219	58,219	36,382	18,312
R^2	0.429	0.429	0.429	0.429	0.430	0.686	0.646
Adj. R^2	0.256	0.256	0.256	0.256	0.257	0.540	0.402
Classroom by Weekday by Hour bin FE	Y	Y	Y	Y	Y	Y	Y
Date by Hour bin FE	Y	Y	Y	Y	Y	Y	Y

Notes: This table reports results for the model of average IEQ variables as measured during the previous lesson on subsequent break length. Columns (1)-(4) incorporate each IEQ variable individually and column (5) includes all IEQ variables. Columns (6) and (7) report results for subsamples of shorter breaks, i.e. one hour and thirty minutes long. All models incorporate the full set of fixed effects as in Equation (4). Standard errors (in parentheses) are clustered at the group level. Significance levels are $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B.5: Maximum IEQ conditions during lesson and break length

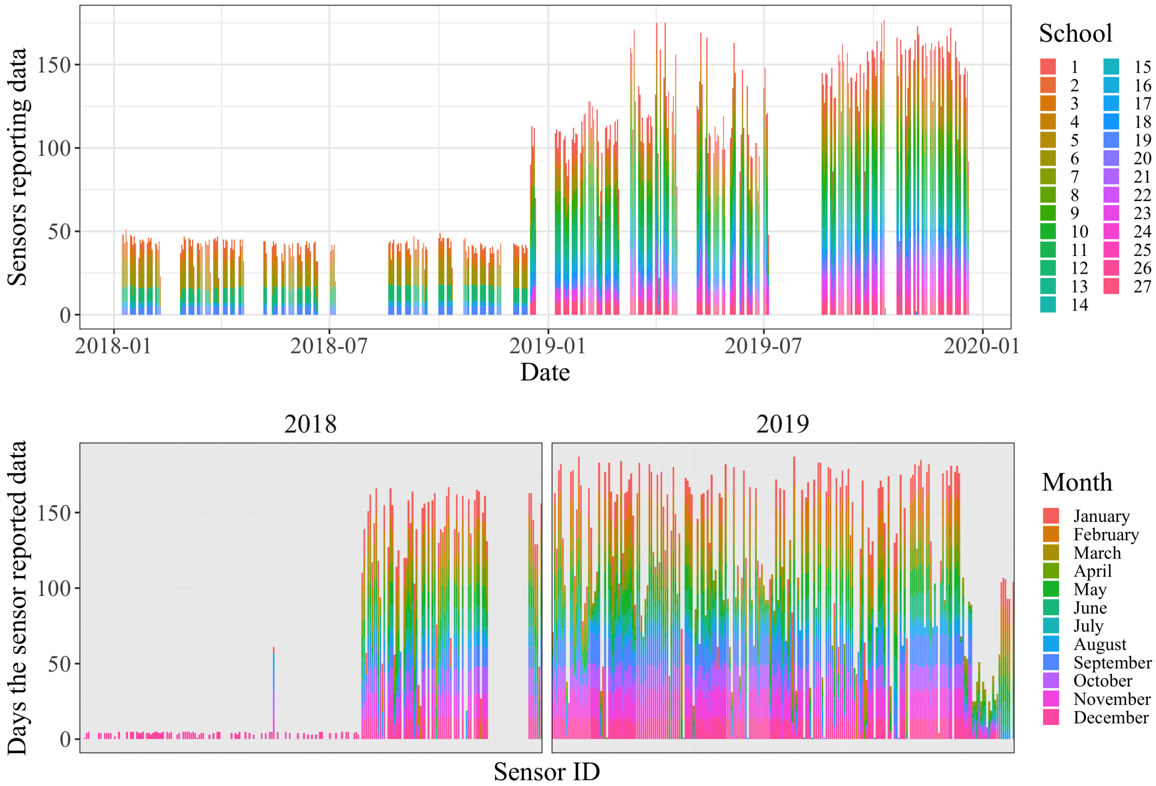
						< 1h	< 30m
Percentile	(1)	(2)	(3)	(4)	(5)	(6)	(7)
CO2 >0.9	0.045**				0.049***	0.088***	0.068***
	(0.020)				(0.018)	(0.015)	(0.013)
CO2 0.9	0.033**				0.036**	0.079***	0.059***
	(0.017)				(0.016)	(0.011)	(0.010)
CO2 0.2	-0.041**				-0.044**	-0.032***	-0.031***
	(0.017)				(0.018)	(0.009)	(0.008)
PM10 >0.9		0.018			0.025	0.011	0.010
		(0.016)			(0.016)	(0.010)	(0.011)
PM10 0.9		0.013			0.016	0.011	-0.001
		(0.012)			(0.012)	(0.009)	(0.010)
PM10 0.2		-0.009			-0.005	-0.008	-0.003
		(0.017)			(0.017)	(0.009)	(0.008)
Temp. >0.9			0.058**		0.056**	0.044***	0.009
			(0.026)		(0.025)	(0.015)	(0.014)
Temp. 0.9			0.053***		0.054***	0.024**	0.006
			(0.014)		(0.015)	(0.010)	(0.010)
Temp. 0.2			-0.040***		-0.038***	-0.016**	-0.014**
			(0.014)		(0.013)	(0.007)	(0.007)
Hum. >0.9				0.034	0.028	0.022	0.014
				(0.024)	(0.024)	(0.017)	(0.014)
Hum. 0.9				0.022	0.020	-0.002	0.000
				(0.017)	(0.018)	(0.012)	(0.011)
Hum. 0.2				0.001	0.006	0.010	0.011
				(0.018)	(0.017)	(0.012)	(0.011)
Observations	58,219	58,219	58,219	58,219	58,219	36,382	18,312
R^2	0.429	0.429	0.429	0.429	0.430	0.686	0.646
Adjusted R^2	0.257	0.257	0.256	0.257	0.257	0.541	0.402
Classroom by Weekday by Hour bin FE	Y	Y	Y	Y	Y	Y	Y
Date by Hour bin FE	Y	Y	Y	Y	Y	Y	Y

Notes: This table reports results for the model of maximum IEQ variables as measured during the previous lesson on subsequent break length. Columns (1)-(4) incorporate each IEQ variable individually and column (5) includes all IEQ variables. All models incorporate the full set of fixed effects as in Equation (4). Standard errors (in parentheses) are clustered at the group level. Significance levels are $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$.

C Data Coverage and Algorithm to Detect Occupancy

C.1 Data Coverage

Figure C.2: Coverage of Sensors by Date and of Dates by Sensor



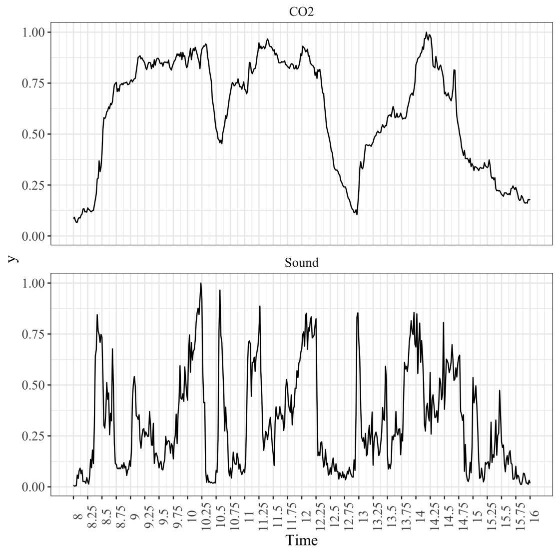
C.2 Algorithm to Detect Occupancy

The algorithm used to determine entry and exit in the classroom searches for increases in CO_2 concentration that are sustained in time while looking for a spike in the sound to detect the exact time when the children have entered the classroom. Sustained increases in CO_2 show that the room is no longer empty and that the door is closed (the rate of CO_2 generation is higher than the rate of air exchange), while a spike in sound indicates that students have entered the classroom and are (in the process of) sitting down. To detect when the classroom is empty, the algorithm searches for a sustained decrease in CO_2 concentration (the rate of CO_2 generation is now lower than the rate of air exchange, as the door is opened and there are fewer students inside the classroom), while also looking for a spike in sound (students make noise when exiting the classroom).

As described in the main text, the algorithm makes use of regularities in the behavior of both CO_2 and sound in the event of children entering or exiting the classroom. Those regularities are easily spotted in Figure C.3. The graph plots how CO_2 concentration and sound decibels move between 8am and 4pm (school hours). One can identify how accumulation of CO_2 starts and sound spikes at the morning entry between 8:15am and 8:30am (8.25 to 8.5 in the graph). For exits, the opposite occurs for CO_2, while sound also spikes as is evident during the first break at 10:15am (10.25 in the graph).

Using these regularities, the algorithm first detects all series of j consecutive minutes showing a CO_2 increase (decrease) during the school day (8am-4pm). After these series are found, we label their first minute as a candidate entry or exit if at any of those minutes in the series, we observe a spike in sound above a threshold s. Once all entry and exit candidates are labelled, the algorithm orders them by time in decreasing order, and retains the first of all consecutive entries before and exit occurs, and the first of all consecutive exits before an entry takes place, such that in order to get an exit, an entry must have been labeled before and vice-versa (except for the first entry and last exit of the day, of course).

Figure C.3: Example of observed levels of CO_2 and Sound (normalized) across a school day

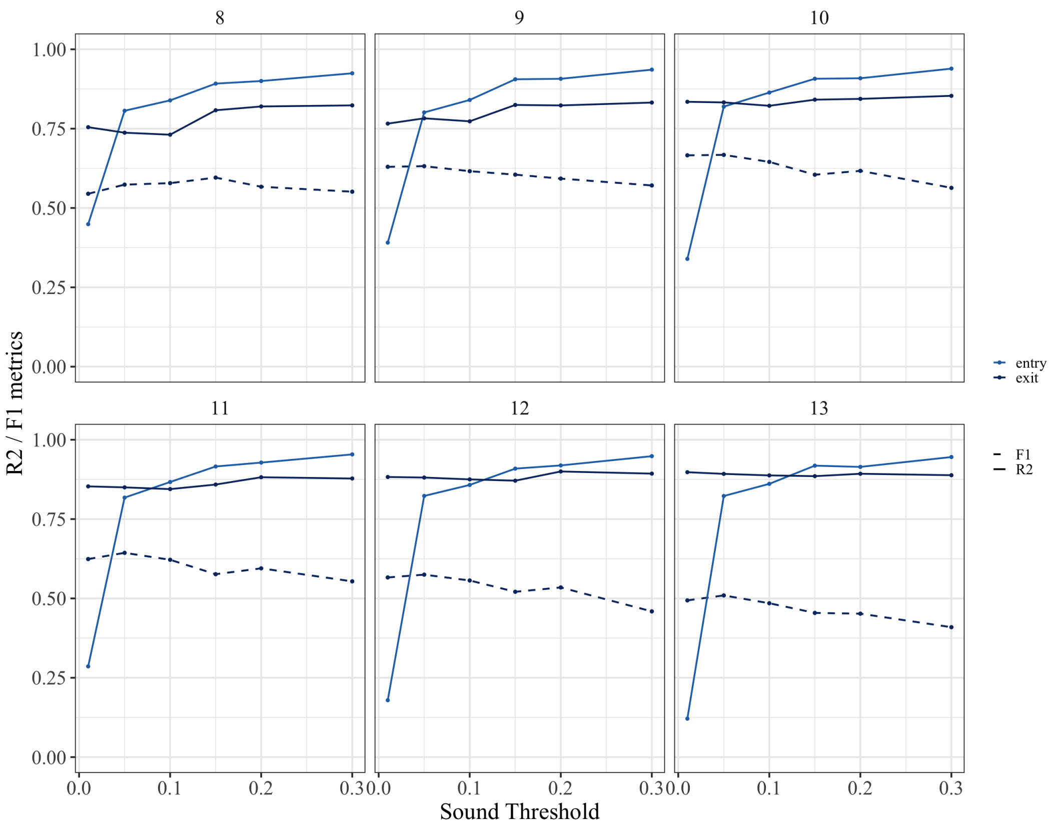


Note: This graphs describe how CO_2 (above) and sound (below) move along a school day (8am to 4pm) inside a particular classroom. It clearly shows how CO_2 starts accumulating and sound spikes when children enter the room and the opposite happens when they exit.

We assess the algorithm's sensitivity and accuracy relative to different values of j and s to determine which give the optimal result. For this purpose, we labeled the observed entries and exits during the school day for nearly 500 graphs of CO_2 and sound series in different days and schools (randomly chosen with monthly stratification). We then compare those labels to the algorithm predictions.

To measure the algorithm's performance we use two metrics: an F1 indicator and the algorithms R^2. The F1, widely used in machine learning contexts, takes the geometric mean of two ratios: (i) the number of correctly predicted entries over all predicted entries; and (ii) the number of correctly predicted entries over all observed entries. This indicator gives a sense of the algorithm's sensitivity as it assesses the proportions of false entry/exit predictions and those of unpredicted but observed entry/exit. However, this measure is silent on the algorithm's accuracy in predicting the exact time at which entry and exits took place. Hence, we assess the algorithm's prediction accuracy using an R^2 coefficient, a well known indicator to assess predictive power. Figure C.4 shows the resulting F1 and R^2 for $j = 7, 8, 9, 10, 11, 12$ minutes and for $s = 0.01, 0.05, 0.1, 0.2, 0.3$ normalized dBA.

Figure C.4: Algorithm performance (F1 and R^2)



The combination of both indicator values suggests that $j = 10$ and $s = 0.05$ predict entries and exits both most accurately as well as most frequently. The highest point achieved by the F1 indicator is at this point (upper right plot), while the highest R2 for both, entries and exits, is also achieved at the same point. We therefore construct our data set on indoor environmental quality using these parameters in the algorithm to predict when students are inside or outside of the classroom.