

The short-term impact of urban air pollution on student achievement

El impacto a corto plazo de la contaminación del aire urbano en el rendimiento de los estudiantes

Jorge Rojas-Vallejos¹
Carmen Gloria Jiménez Bucarey²
Marcela Espinoza³
Luis Araya-Castillo⁴

DOI: 10.13043/DYS.87.1

Abstract

Goal: The relationship between air pollution and human health has been widely studied in medicine and economics. However, only a few recent articles have estimated the direct impact of air pollution on academic achievement. **Methodology:** We collect daily air pollution data from all monitoring stations in the city of Santiago, Chile, and combine it with academic results recorded by the National Learning Outcome Assessment System of (SIMCE in Spanish) for fourth grade students over the past decade, to create a unique pooled cross-sectional dataset. **Results:** Applying different econometric strategies, we find that certain air pollutants have a negative impact on the test scores

-
- 1 Assistant Professor, Facultad de Economía y Negocios, Universidad Andrés Bello, Chile. Email: jorge.rojas.v@unab.cl
 - 2 Vice-president of Quality Assurance, Universidad Andrés Bello, Chile. Email: carmen.jimenez@unab.cl
 - 3 Instructor, School of Medicine, Universidad de Santiago de Chile, Chile. Email: marcela.espinoza@usach.cl
 - 4 Associate Professor, Facultad de Economía y Negocios, Universidad Andrés Bello, Chile. Email: luis.araya@unab.cl

Este artículo fue recibido el 24 de octubre del 2019, revisado el 17 de abril del 2020 y finalmente aceptado el 22 de junio del 2020.

of the cohort. Our estimated effects are negative and statistically significant, ranging between 0.02 and 0.06 standard deviations. **Conclusions:** These results seem to be small, but are equivalent to around 30% of a successful educational policy that specifically targets performance. **Recommendations:** Public policy must consider the positive externalities of reducing air pollution.

Key words: government policy, academic performance, pollutants, Chile.

JEL Classification: C23, C55, I18, Q53.

Resumen

Objetivo: La relación entre la contaminación del aire y la salud humana se ha estudiado ampliamente en medicina y economía. Sin embargo, sólo unos pocos artículos recientes han estimado el impacto directo de la contaminación del aire en el rendimiento académico. **Metodología:** Recopilamos datos diarios de contaminación del aire de todas las estaciones de monitoreo en la ciudad de Santiago de Chile, y los combinamos con los datos de los resultados académicos registrados por el Sistema de Medición de la Calidad de la Educación (SIMCE) para estudiantes de cuarto grado durante la última década, para crear un conjunto agrupado único de datos transversales. **Resultados:** Aplicando diferentes estrategias econométricas, encontramos que ciertos contaminantes del aire tienen un impacto negativo en los puntajes de las pruebas de la cohorte. Nuestros efectos estimados son negativos y estadísticamente significativos, oscilando entre 0.02 y 0.06 desviaciones estándar. **Conclusiones:** Estos resultados pueden interpretarse como pequeños, pero equivalen al efecto de alrededor del 30% de una política educativa exitosa que apunta específicamente al desempeño. **Recomendaciones:** La política pública debe considerar las externalidades positivas de reducir la contaminación del aire.

Palabras clave: política pública, rendimiento escolar, contaminación del aire, Chile.

Clasificación JEL: C23, C55, I18, Q53.

Introduction

The effects of air quality on human health have been studied extensively by the medical community. There are epidemiological studies that determine the health effects in the short-, medium- and long term for different pollutants and population groups (Toro et al. 2015; Kampa and Castanas 2008). Some efforts

have been made to estimate the effect of pollution on social variables such as subjective utility (Levinson, 2012), school attendance (Currie *et al.* 2009; Ramson and Pope 1992), student performance (Mohai *et al.* 2011; Marcotte 2015) and overall well-being (Román-Collado and Jiménez de Reyna, 2019).

In this paper, we estimate the effects of air quality on student performance using a repeated pooled cross-section data structure rather than the typical cross-sectional data used in previous studies. We focus on fourth graders, because they tend to experience physical symptoms much more strongly than adults (Schwartz, 2004). Studies on attendance and air quality find that, even when they are below government standards, high levels of carbon monoxide (CO) have a significant negative impact on school attendance. However, there is limited information about the impact of air pollutants on student academic achievement, and this is subject to the usual external validity concerns (Lavy, Ebenstein and Roth, 2014).

Pollution is a serious environmental problem in Santiago, Chile, because the city is located in an enclosed valley with limited wind and rain. In addition, these types of valleys are subject to the effects of a thermal inversion layer, a well-known phenomenon that causes an 'atmospheric lid' to form above the valley floor, preventing the escape of pollutants. Coupled with high levels of industrial development and vehicle emissions, the inversion layer contributes to levels of pollution that often exceed the values suggested by the World Health Organization (WHO, 2016), the standards set by the US Environmental Protection Agency (EPA) (EPA, 2019), and the Chilean guidelines.

The medical literature emphasizes that pollutants may have negative health effects above a certain threshold (Bernstein *et al.* 2004). These thresholds depend upon many factors including age, gender, and nutrition. However, based on aggregate information, regulatory agencies have established some legal maximum thresholds for different periods of time. These regulations vary by country and over time.

Particles less than 10 micrometers in diameter (PM_{10}) can potentially increase respiratory health risks and cardiovascular problems in the long run, including chronic diseases such as asthma or lung cancer. Fine particles ($PM_{2.5}$) are so small that they can get into the lungs, causing serious health problems in the long run and temporary symptoms such as irritation of the eyes, nose, and throat, coughing, chest tightness, shortness of breath, heart palpitations

and fatigue. Nitrogen dioxide (NO_2) can irritate the lungs and lower resistance to respiratory infections. Carbon monoxide (CO) is known to reduce oxygen in the bloodstream, while exposure to high levels of CO is associated with visual impairment, reduced work capacity, lowered cognitive abilities, and difficulty in performing complex tasks. CO in extremely high concentrations can cause death by asphyxiation. Exposure to ozone (O_3) for 6 to 7 hours, even at relatively low concentrations, may induce symptoms such as chest pain, coughing, nausea, and pulmonary congestion. In high concentrations methane (CH_4) weakens the central nervous system. Finally, there is sufficient evidence that short-term exposure to sulfur dioxide (SO_2) for periods ranging between 5 minutes and 24 hours may induce bronchoconstriction and increased asthma symptoms, particularly among children and the elderly.

We analyze the short-term impacts of air pollution on student academic achievement using observational data. As discussed above, pollutants produce short- and long-term negative effects on human health. Children tend to be particularly susceptible, and their experiences of temporary symptoms may disrupt usual activities (Schwartz, 2004). In this sense, pollution might impact the concentration of students while they are taking the test, significantly diminishing their ability to understand questions and search for solutions. Thus, this channel should be reflected in the coefficient estimates of pollution. Notice that we use observational data. This implies that the data is not obtained from lab conditions, so we cannot control any variable at our will. Hence, in this type of analysis there are several endogeneity concerns that should be taken into consideration. Simple cross-sectional comparisons of pollution and student achievement suffer from omitted variable bias (OVB). Air pollution is not randomly assigned, and several confounding factors (eg, urbanization, crime levels, school quality) that are correlated with air pollution and student achievement could bias the results. Furthermore, parents can choose where to live based on their preferences for clean air and school quality (Chay and Greenstone, 2005). We use two identification strategies to address these issues. Our first model uses school and year fixed effects. That is, we compare students within the same school over time, while controlling for constant unobservable school and year factors, as well as student, family, and school-observable characteristics. In our second strategy, we use this same model, but instrument the level of pollution by relative humidity levels. Relative humidity is an instrumental variable with some disadvantages because it is weakly correlated with some pollutants. Rain would have been a much stronger instrument because it strongly changes the pollution levels

in an exogenous way. However, there was no rainfall during the spring period analyzed, requiring relative humidity to be used instead. In general, we find that a one standard deviation increase in pollution levels decreases student performance in math and science. Finally, the effect on language exam scores is not statistically significant, despite having the expected negative sign.

The paper is organized as follows. The next section describes the relevant background literature and some specific characteristics of the city of Santiago. Section 3 presents the data, while Section 4 explains the econometric methodology. Section 5 describes and discusses the principal results of the paper. Section 6 concludes.

2. Background

Three different literature strands are relevant to this study. The first consists of literature describing the negative effects of different air pollutants on human health, as well as their transmission mechanisms. There is sufficient evidence of the association between SO_2 , NO_2 and CO exposures and respiratory and cardiovascular health problems. Similarly, short-term and long-term effects of particulate matter and ozone levels on human health have been found in different studies (Kampa and Castanas 2008; Bernstein *et al.* 2004; Brunekreef and Holgate 2002).

The second relevant line of research consists of studies of the relationship between pollution and school attendance. There is evidence of a monotonically decreasing pattern between CO levels and school attendance (Currie *et al.* 2009). Furthermore, an increase of $100\mu\text{g}/\text{m}^3$ in the 28-day moving average of PM_{10} translates to an increase in the school absence rate equal to approximately two percentage points (Ramson and Pope, 1992). These models control for climate variables such as temperature, snowfall, and variables indicating day of week, month of school year, and days preceding and following holidays and extended weekends.

The third related strand analyzes the relationship between differences in pollution across geographic areas and differences in average family income across these differences. Poorer students are exposed to higher levels of pollution than their higher-income counterparts (Levinson 2012; Chay and Greenstone 2005; Pastor, Morello-Frosch and Sadd 2006). These findings raise an important

endogeneity concern for our work, namely that differences in pollution levels across geographic areas might actually pick up differences in socioeconomic status in different student populations.

All these studies emphasize that pollution may have thresholds that mean that only once they are exceeded are its effects noticed. In our case, we are studying the effects of pollution on the score achieved by a given student on the day the test was taken. These standardized tests are taken by students corresponds in the Spring. This time of the year is one of low air pollution in Santiago and air quality might be considered "good" according to Chilean regulations. This might make it harder to detect statistically significant effects. Indeed, we fail to obtain statistically significant results for some pollutants. However, in the case of CO and NO_x, we do find a negative statistically significant coefficient estimate of the impact on achievement. This suggests that the threshold for both pollutants, which may harm student achievement and probably also their health, could be lower than the one imposed by the environmental agency. This result is in line with those reported in other articles (Currie et al. 2009; Levinson, 2012).

Table (1) below presents the pollution standards for the United States and Chile. We see that although some standards are the same, this is not true for all pollutants. Furthermore, these regulations are revised constantly, and modified whenever new evidence suggests the need.

Table 1. Standards for Air Quality

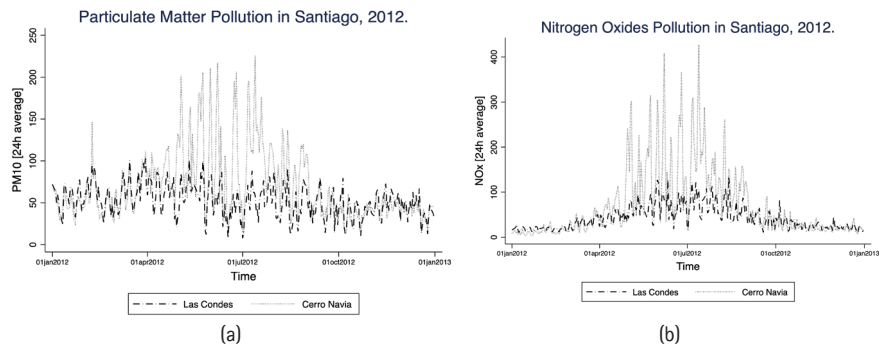
Pollutant	US Standards		Chilean Standards	
PM ₁₀ (µg/m³)	150 (24h)	50 (annual)	150 (24h)	50 (annual)
PM _{2.5} (µg/m³)	35 (24h)	12 (annual)	50 (24h)	20 (annual)
O ₃ (ppb)	75 (8h)	120 (1h) ^a	-	-
NO ₂ (ppb)	100 (1h)	53 (annual)	213 (1h)	53 (annual)
SO ₂ (ppb)	75 (1h)	140 (24h) ^b	750 (1h)	96 (24h)
CO (ppm)	9 (8h)	35 (1h)	9 (8h)	26 (1h)

Updated 2014 National Ambient Air Quality Standards (NAAQS).
Updated 2014 Chilean Regulation. Decrees 12/2010, 59/1998, 113/2002, 114/2002, 115/2002.
^a This standard from 1971 was revoked in 1997 because it would not provide additional public health protection.
^b This standard from 1971 was revoked in 2010 because it would not provide additional public health protection given a 1-hour standard at 75 ppb.

In the city of Santiago, the concentration of pollutants is determined by topographical and meteorological characteristics that produce a well-established pattern of air pollution. These geophysical conditions are more relevant than the location of pollution sources when it comes to defining the areas where the pollution will be concentrated. This means that air pollutants are redistributed by wind currents, and concentrated in specific geographical areas of the city (Schmitz 2005; Gramsch *et al.* 2006). Higher pollution levels coincide with poorer neighborhoods because of the patterns of the airstreams and the consistent social segregation that characterizes the city (Fernández and Wu 2016; Sabatini *et al.* 2009). Hence, they suffer the negative impacts of the pollution more than richer towns. This could imply some endogeneity issues because of sorting. In the methods, we discuss ways to address this question.

Fig. (1) illustrates that different pollutants do indeed tend to exhibit higher concentration levels in poorer towns than richer ones. The atmospheric science literature for the city suggests that this holds true for most areas.

Fig. 1. Different pollutant levels for two counties: Las Condes (high income) and Cerro Navia (low income).



Finally, the National Bureau of Economic Research published a paper whose conclusions are similar to ours concerning the causal relationship between air pollution and student academic performance (Lavy, Ebenstein and Roth 2014). The general methodology of this study is similar to ours, although the structure of their data is different. The authors find that air pollution has a negative effect on standardized test scores. We contribute to this literature by disaggregating the effects by subject, namely, math, science and language.

In addition, unlike the majority of studies, we provide a new case study for a city in a developing country.

3. Data

To analyze the relationship between student academic performance and air pollution, we collect data on standardized test scores regularly taken by fourth graders in the Chilean Educational System. These students are the youngest cohort to take standardized tests and are hence expected to be the most affected by air pollution. We also collect data on air quality and weather from all available monitoring stations in the city of Santiago. In addition to academic performance data, the dataset contains demographic information on every student. Using this information, we build a socioeconomic status (SES) index by means of factor analysis.

3.1 Student Achievement and Attendance Data

According to the Department of Education, in 2009 the Chilean educational system served approximately 3.5 million students in more than 11,000 schools. Ninety five percent of children aged 5 to 14 attend school, as did seventy five percent of 15–19 year-olds (OECD, 2012). One important characteristic of the Chilean educational system is that since the early 1980s, the administration of public schools has been shared between the public and private sectors. Local governments or municipalities administer public schools, which today represent forty-six percent of the schools in the country. Municipal schools are financed by the central student-based subsidy and cannot charge tuition. The private sector administers both publicly subsidized private schools (forty-nine percent of schools) and non-subsidized private schools (five percent). Publicly subsidized private schools receive the same student-based subsidy, but are allowed to charge tuition up to a certain amount. When subsidized schools charge tuition, their public subsidy is reduced but not by a commensurate amount. In this study, we include both municipal and publicly subsidized schools.

We use standardized student-level achievement scores from the National Learning Outcome Assessment (SIMCE in Spanish) for all fourth graders during the years 2002, and 2005 through 2012. This national assessment is equivalent to the National Assessment of Educational Progress (NAEP) in the United

States. We focus our analysis on all fourth graders enrolled in schools in Santiago. The SIMCE database also includes student demographic information, as well as family background characteristics. Given that the variables used to measure family characteristics are not always comparable across years, we estimated a composite score that measures the socioeconomic level of every student's family. We estimated this composite score by conducting a factor analysis using the father's and mother's educational level, household income, and the number of books at home. Student participation is mandatory in all subjects tested, which include *Mathematics* and *Language*, as well as a test that combines natural and social sciences that we call *Sciences*. Student achievement results are comparable across time and are reported on a scale of mean of approximately 250 points and a standard deviation of approximately 50 test points. For this study, we standardize test scores every year to a mean zero and a standard deviation of one, so that the interpretation of the estimates is that commonly used in the literature.

3.2 Pollution and Weather Data

We scrape the web for unique information on air quality from all eleven pollution-metering stations in Santiago, which have been monitoring air conditions for the past decades. These data are obtained from the National Information System on Air Quality (SINCA in Spanish). We focus our analysis on different pollutants that are measured by these stations, namely, PM_{10} , $PM_{2.5}$, CO , O_3 , CH_4 , (NO_x) , and SO_2 . Notice that NO_x is the sum of nitric oxide (NO) and nitrogen dioxide (NO_2).

These eleven stations are well-spaced across different counties of the city, namely Independencia, La Florida, Las Condes, Parque O'Higgins, Pudahuel, Cerrillos, El Bosque, Cerro Navia, Puente Alto, Talagante, and Quilicura. All the stations measure daily averages of the pollutants and start and stop collecting data on different dates. The only exceptions are Independencia and La Florida which do not measure CH_4 at all.

To estimate the average pollution levels for each school in Santiago, we gathered their geographical coordinates alongside those of all the monitoring stations located in the city. Geographical information for schools is provided by the Department of Education, while for the monitoring stations the data is taken from the SINCA website. We then estimate the distance between each school and each monitoring station. This figure is computed using the standard formula to measure Earth distances.

$$\begin{aligned} a &= \sin\left(\frac{\Delta\Phi}{2}\right) + \cos\Phi_1 \cdot \cos\Phi_2 \cdot \sin\left(\frac{\Delta\lambda}{2}\right)^2 \\ c &= 2 \cdot \arctan\left(\sqrt{a}, \sqrt{1-a}\right) \\ d &= R \cdot c \end{aligned} \tag{1}$$

where Φ is latitude, λ is longitude, 1 and 2 represent locations, Δ is the difference operator and R is Earth’s radius (mean radius = 6,371 km).

The next step is to compute the relevant pollution level for each school using two different methodologies: 1) a weighted average using the inverse of the distance between the school and each station with non-missing information; and 2) the average pollution level from the nearest station with non-missing information. We estimate the average pollution levels using the pollution level during the day of the test, as well as 2-day, 3-day, 7-day, 14-day, and 28-day averages previous to the day of the test. From 2002 to 2009, tests were conducted at the end of the first week of November, or the beginning of the second week. From 2010 onwards, tests took place during the third week of October. We use the exact dates for the estimation of the pollution averages in this study. We also collect daily information on temperature, relative humidity, and precipitation.

Table (2) presents descriptive statistics for pollutants across all years using the daily average metric.

Table 2. Descriptive Statistics for Daily Average Pollution

Pollutant	# Obs.	Mean	Std. Dev.	Min	Max
PM ₁₀ (μg/m³)	28,197	51.17	12.38	22.98	73.90
PM _{2.5} (μg/m³)	28,197	17.69	6.41	6.69	33.18
O ₃ (ppb)	28,197	19.71	1.87	15.05	26.09
SO ₂ (ppb)	28,197	2.33	1.02	0.61	5.58
CH ₄ (ppb)	21,931	1.31	0.57	0.01	1.98
CO (ppm)	28,197	0.32	0.08	0.15	0.73
NO (ppb)	28,197	11.81	4.95	4.05	25.57
NO _x (ppb)	28,197	27.14	8.11	10.74	48.43

Source: Authors’ computation based on Data from the National Air Quality Agency of Chile for the years of analysis.

The data show that pollutant levels for low- and high-income counties differ significantly. For instance, in Cerro Navia, a low-income county, NO_x levels exceeded the standard 22 times in 2012, while for Las Condes, a high-income county this did not happen once. Similar patterns are observed for other pollutants with the sole exception of O_3 , which is more concentrated in the higher parts of the city where Las Condes and most of the affluent towns are situated. This is explained by the fact that more sunlight is required to produce O_3 and these conditions are met in these more elevated localities.

Given the physical characteristics of the city, most of the towns in Santiago experience pollution levels that fall between these two extremes. In addition, the regional government has implemented policies that have noticeably reduced pollution levels over time. This provides sufficient variation in the data, a permitting us to expect some statistically significant results. Had these conditions not been present, the analysis would not have been possible.

4. Empirical Methods

Given that air pollution is not randomly assigned, and that parents may choose where to live based on air quality and other factors correlated with this and with student achievement (e.g., high urbanization, school quality, density, crime), simple cross-sectional comparisons of air pollution and student achievement cannot identify the causal effect that pollution may have on short-term student performance.

Moreover, Santiago is a highly economically segregated city. For historical reasons, low- and moderate-income families live on the west side of the city, while upper-middle income and rich families congregate in the east. The east side has an elevation averaging 900 meters above sea level, compared to around 550 meters on the west side. This is not a trivial observation since most pollutants have a heavier molecular weight than air, and hence tend to follow gravity. In addition, airstreams take pollutants to the lower section of the valley (Gramsch *et al.* 2006). As a result, low- and moderate-income areas have a higher concentration of air pollutants than upper-middle and rich areas.

To remove the influence of key confounding unobservable factors, we exploit the panel nature of the data by estimating the econometric model given in equation (2). We estimate all models for three different time periods: i) 2002

and 2005 through 2012; ii) 2006 to 2012; and iii) 2008 to 2012. We do this to take into consideration the fact that not all pollutants started being collected at the same time. We present the results using (i), that is, including all years.

$$A_{ijt}^s = \beta_0 + \delta * Pol_{jt}^p + X_{ijt} * \beta_1 + S_{jt} * \beta_2 + \eta_j + \lambda_t + \varepsilon_{ijt} \quad (2)$$

where A_{ijt}^s is the standardized test scores for student i , in school j at time t for subject s (math, language and sciences). The variable Pol captures the pollution level of the different pollutants (i.e., p). X is a vector of student demographic characteristics including family socioeconomic status (SES), and S includes school level characteristics. This model includes school and time fixed effects (η_j and λ_t , respectively). In simple words, it compares the achievement level of fourth graders at a given school across time, while controlling for the observable time-variant socioeconomic characteristics of the students and schools, as well as unobservable fixed characteristics. The results of this estimation are presented as Model B. We also estimate this equation without school fixed effects (Model A). See Rojas-Vallejos and Lastuka (2020) and references therein for a further discussion on the application of fixed effects in a panel context.

We then estimate a Two-Stage-Least-Square (2SLS) model where the instrument for the level of pollution is relative humidity, presenting these results as Model C. The ideal instrument is the level of precipitation, since rain clears the skies and lowers the levels of pollution exogenously. However, at the time of the test (i.e., mid-October and November) the level of precipitation in Santiago is zero for our dataset. These results must be interpreted with caution since the instrument is not very strong by looking the first stage output. Despite this, they do shed some light on the relationship being explored.

Our models provide consistent estimates of the effects of air pollution on student achievement as long as no other factors systematically affect both pollution and student achievement at the school-year level. One example that would violate this assumption would be a fire that affects air quality and student achievement at the same time. During this time period, we are not aware of any fires or large environmental disasters affecting Santiago. Another potential violation would be that certain areas of the city are undergoing high levels of development that have increased pollution levels and changed the demographic characteristics of the school population, factors that could affect final student performance levels. This is a potential threat to our models. We do our best to control for it by including the socioeconomic status of the students

attending each school. Results for our estimates, including attendance data, are not included in this study since our focus is on short- rather than long-run effects. Nevertheless, we obtain some insight into the longer-term effects of pollution on student performance by regressing the cumulative averages of pollution of different time periods on test scores.

5. Results and Discussion

In this section, we present our estimates of the short-term and cumulative effects of air pollution. Short-term effects are estimated by considering the pollution levels on the date the test was taken, while cumulative effects are estimated using the average pollution level for a number of days before the test and its date of application.

Panel (1) in Table (3) presents variants of equation (2) for three different pollutants (NO_x , NO, and CO) and math test results of fourth graders in Santiago for the years 2002, and 2005–2012. Each row represents a different estimating equation. Model A is constituted by OLS estimates that exclude fixed effects. Model B is the preferred model and shows the estimates for equation (2). Model C estimates a 2SLS equation model of (2). For each model we present the results using the weighted pollution level of all monitoring stations ("All" column), as well as the pollution level of the nearest station ("Nearest" column).

For the case of NO_x , our OLS results suggest that pollution levels have no impact on the math achievement levels of fourth graders in Santiago, raising the possibility of unobserved confounding variables. We find a negative and statistically significant association between NO_x and math test results in Model B. In terms of magnitude, the estimates imply that with a one standard deviation (sd) increase in pollution levels (i.e., 8.11 units in the case of NO_x) math test results decrease by about 0.06 sd and 0.008 sd including all stations and the nearest, respectively. These results are considered medium to low size effects on student performance. The results of Model C also show a negative impact of NO_x on math test results but are not statistically significantly different than zero. The results for NO are quite similar to those observed for NO_x .

In the case of CO, we find that the OLS results suggest a negative association between this pollutant and student performance. Model B, our preferred model, also suggests that CO has a negative impact on student achievement.

Model C shows statistically insignificant results. In terms of the magnitude of the effect in Model B, the results indicate that a one sd increase in CO (0.08 being the sd for this pollutant in the data: see Table (2)) decreases math test results by about 0.02 sd.

Table 3. The Impact of Air Pollution on Academic Performance

Panel 1: Pollution on Math, 1-day daily average.

Pollutant	Model A		Model B		Model C	
	All	Nearest	All	Nearest	All	Nearest
<i>NO_x</i>	-0.001 (0.005)	0.000 (0.001)	-0.008*** (0.003)	-0.001*** (0.000)	-0.016 (0.012)	-0.002 (0.002)
<i>NO</i>	-0.005 (0.006)	-0.001* (0.001)	-0.006* (0.004)	-0.001* (0.001)	-0.024 (0.018)	-0.004* (0.003)
<i>CO</i>	-0.522** (0.221)	-0.478*** (0.097)	-0.231** (0.113)	-0.189** (0.085)	0.217 (0.164)	0.345 (0.218)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes
N	670,000	670,000	670,000	670,000	670,000	670,000

Panel 2: Pollution on Sciences, 1-day daily average.

Pollutant	Model A		Model B		Model C	
	All	Nearest	All	Nearest	All	Nearest
<i>NO_x</i>	-0.001 (0.004)	0.000 (0.001)	-0.008*** (0.002)	-0.001*** (0.000)	-0.039*** (0.012)	-0.005*** (0.002)
<i>NO</i>	-0.003 (0.005)	-0.001 (0.001)	-0.004 (0.003)	-0.001 (0.001)	-0.005 (0.019)	-0.001 (0.003)
<i>CO</i>	-0.415** (0.192)	-0.078*** (0.027)	-0.144 (0.102)	-0.029** (0.015)	0.523*** (0.163)	0.433*** (0.112)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes
N	670,000	670,000	670,000	670,000	670,000	670,000

Panel 3: The Pollution on Language, 1-day daily average.

Pollutant	Model A		Model B		Model C	
	All	Nearest	All	Nearest	All	Nearest
NO_x	-0.002 (0.004)	0.000 (0.001)	-0.006*** (0.002)	-0.001** (0.000)	-0.039*** (0.012)	-0.005*** (0.002)
NO	-0.003 (0.005)	-0.001 (0.001)	-0.004 (0.003)	-0.001 (0.001)	-0.005 (0.019)	-0.001 (0.003)
CO	-0.554*** (0.183)	-0.036*** (0.007)	-0.300*** (0.107)	-0.229** (0.145)	0.046 (0.172)	0.043 (0.112)
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes
N	670,000	670,000	670,000	670,000	670,000	670,000

*** $p < 0.01$, ** $p < 0.05$, $p < 0.10^*$. Clustered standard errors are in parentheses. All models control for student and school socioeconomic level.

Panel (2) in Table (3) presents the same types of results as Panel (1), but this time for science. The patterns in Models A and B are quite similar to the ones observed for math. The only difference is that the results for the 2SLS (Model C) for the case of NO_x are now statistically significant and larger in magnitude. We also observe a perverse sign for CO in the case of Model C. This could be explained by the weakness of the instrument. Let us stress once again that we are using relative humidity as an instrument since there is no rainfall for the application dates in this sample. Relative humidity as an instrument is relatively weak for different pollutants.

Panel (3) in Table (3) presents the language test results. Observe that the patterns are again consistent with the ones observed for math and sciences. However, the magnitudes of the effects are somewhat smaller, with the exception of those associated with CO.

Nevertheless, there is sufficient evidence to argue that, compared with math and sciences, pollution seems to have almost no negative effect on language performance. This result might be expected if we consider that language is a more intuitive skill than math and sciences, and is also more frequently used in daily life. Furthermore, medical research shows that different parts of the

brain deal with different cognitive skills. Notice that the impacts of pollution levels seem to be robust to whether we consider the average of all monitoring stations or the value of the nearest. The estimates change in magnitude, but the signs and significance levels remain similar in most cases.

It is worth pointing out that on the dates the tests were administered pollution levels were relatively low. However, some pollutants do still have negative and significant effects on students' achievements, while it appears that the levels of PM_{10} are so low that we fail to detect significant effects. For these periods, PM_{10} is on average around $50\mu g/m^3$ per day with a standard deviation of about $10\mu g/m^3$, while the regulation states that the maximum cannot exceed $150\mu g/m^3$ in 24 hours. Thus, pollution remains considerably below the threshold that is considered to have negative health effects in Chile.

Thus, as the medical research shows, at higher pollution levels, health effects are intensified. Hence, in the months of June and July when pollution reaches its highest levels, exceeding the threshold at least 20% of the time, the impact on learning and performance may be quite significant. June and July are months that form part of the academic calendar in Chile, so students attend classes regularly during these months.

Next, we briefly discuss the impact of other relevant variables on student performance. In Panel (1) in Table (4) we show that socioeconomic status (SES) plays an important role in student performance. A one sd increase in socioeconomic status will increase student achievement by around 0.35 sd. This is quite high considering that policies exclusively targeting student performance achieve improvements of around 0.1 sd. We can also see that whether the school is private (benchmark group), public or subsidized (similar to charter schools in the United States) also affects student academic performance. The data tell us that students from public schools tend to do a lot worse than otherwise similar students from private schools, while in the case of subsidized schools this reduced academic performance is less marked but still present.

We find important negative effects of certain pollutants on students' academic performance (0.02–0.06 sd), equivalent to around 30% of a successful educational policy in developing countries [18]. Reducing pollution levels has many benefits, and this research finds evidence that it would likely increase student performance, thereby reducing Santiago's achievement gap. This holds

Table 4. Pollution, Demographics and Cumulative Effects

Panel 1: Socioeconomic Status Impacting Test Scores.

	Model A		Model B		Model C	
	All	Nearest	All	Nearest	All	Nearest
<i>NO_x</i>	-0.001 (0.004)	0.000 (0.001)	-0.008*** (0.002)	-0.001*** (0.000)	-0.039*** (0.012)	-0.005*** (0.002)
<i>SES</i>	0.422*** (0.007)	0.422*** (0.007)	0.266*** (0.003)	0.266*** (0.003)	0.266*** (0.002)	0.266*** (0.002)
<i>Pub.School</i>	-0.334*** (0.030)	-0.335*** (0.030)				
<i>Sub.School</i>	-0.042** (0.021)	-0.043** (0.021)				
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
School FE	No	No	Yes	Yes	Yes	Yes
N	670,000	670,000	670,000	670,000	670,000	670,000

Panel 2: Cumulative effect of *NO_x* for Sciences using all stations.

Period	Model A	Model B	Model C
<i>NO_x</i> , 1-day	-0.001 (0.004)	-0.008*** (0.002)	-0.039*** (0.012)
<i>NO_x</i> , 2-day	-0.003 (0.005)	-0.011*** (0.003)	-0.043*** (0.014)
<i>NO_x</i> , 3-day	-0.003 (0.006)	-0.014*** (0.003)	-0.042*** (0.012)
<i>NO_x</i> , 7-day	0.001 (0.007)	-0.013*** (0.004)	-0.023*** (0.007)
<i>NO_x</i> , 14-day	0.000 (0.006)	-0.014*** (0.004)	-0.013*** (0.005)
<i>NO_x</i> , 28-day	0.003 (0.006)	-0.007 (0.005)	-0.011*** (0.005)
Year FE	Yes	Yes	Yes
School FE	No	Yes	Yes
N	670,000	670,000	670,000

*** $p < 0.01$, ** $p < 0.05$, $p < 0.10^*$. Clustered standard errors are in parentheses. All models control for student and school socioeconomic level.

for the city of Santiago given the specific geographical and environmental characteristics discussed above.

Panel (2) in Table (4) shows a non-monotonic behavior of the impact of pollution (in this case NO_x) on student achievement. However, this adverse effect seems to be persistent and does not go away. This could play a significant role in the accumulation of human capital over the long run. Suppose that achievement is negatively affected every year by 0.2 sd, then a student's years spent at primary and secondary school, the effect could reach a value between 0.24 sd and 0.28 sd. This type of behavior is similar for other pollutants. The results for math are similar, and for language are mostly non-significant.

6. Conclusion and Policy Recommendations

The main results of this study suggest that air pollution has a negative impact on student academic performance. The magnitude of this impact seems small. However, it represents around 30% of a successful intervention targeting educational achievement in developing countries (JPAL, 2019). Furthermore, there could be important cumulative effects in cases where students are exposed to a combination of different pollutants.

Obtaining convincing estimates of the causal impact of air pollution on student performance is a difficult endeavor. There are two major obstacles. The first is the presence of confounding factors resulting from residential sorting that is correlated with both air and school quality. The second is that there is a paucity of air pollution measurements and of student test score data readily available. By merging administrative student achievement records with data that we scraped from the web, we were able to create a cross-sectional pooled dataset that allows us to control for school- and time-unobservable constant characteristics that could bias our estimates. Nonetheless, we provide sufficient evidence and robustness checks to suggest that air pollution has a negative and statistically significant impact on student achievement, a relation that is particularly robust in the cases of math and science.

Notice that not all pollutants have a statistically significant negative sign. One reason could be the low levels of pollution during November in the city of Santiago. Thus, air quality standards for pollutants such as CO and NO_x , which

show a negative effect, should be revised. In addition, given the specific environmental conditions of Santiago, poor neighborhoods tend to have lower air quality levels. We rarely think of air quality distribution as an inequality concern in education, but our results suggest that pollution could be an important factor in the academic achievement gap observed across students of different socioeconomic backgrounds.

All this suggests two straightforward policy recommendations. First, the air quality standards for CO and NO_x should be revised. The data show that even when concentration levels of these pollutants are low, during November (Spring in Chile), they may have a negative impact on student achievement. In addition, these pollutants reach higher levels during winter and tend to be present at higher levels in poorer towns. Thus, there are some inequality concerns that should be addressed. One short-term policy for this would be to create a compensation scheme to benefit schools located in those towns. These benefits might consist of providing complementary health insurance or improving the quality of the food supply for these students. In the longer run, sources of pollution should be moved away from all schools, in a joint effort of the private sector and the government. In the medium- and long-term improving the academic performance of thousands of students will pay off in terms of increased tax revenues for government and growing sales for companies.

Finally, future research on this topic should aim to understand the long-run effects of pollution on student achievement and, consequently, on wages. The challenge to exploring this question lies in the difficulty of obtaining a panel dataset containing information on pollution, education and wages for different cohorts of people.

Acknowledgements

The authors did not receive funding from any institution to carry out this study. We thank Amy Lastuka of the Center for Sustainable Energy (CSE) in San Diego, CA, United States, for her constructive suggestions. We also thank two anonymous reviewers whose comments have greatly improved this manuscript. Any remaining errors and omissions are our own.

References

1. Bernstein J.A., Alexis N., Barnes C., Bernstein I.L., Nel A., Peden D., et al. (2004). Health effects of air pollution. *Journal of Allergy and Clinical Immunology*, 114(5), 1116-1123. <https://doi.org/10.1016/j.jaci.2004.08.030>
2. Brunekreef B., & Holgate S.T. (2002). Air pollution and health. *The Lancet*, 360(9341), 1233-1242. [https://doi.org/10.1016/S0140-6736\(02\)11274-8](https://doi.org/10.1016/S0140-6736(02)11274-8)
3. Chay K.Y., & Greenstone M. (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy*, 113(2), 376-424. <https://doi.org/10.1086/427462>
4. Currie J., Hanushek E.A., Kahn E.M., Neidell M., & Rivkin S.G. (2009). Does pollution increase school absences? *Review of Economics and Statistics*, 91(4), 682-694. <https://doi.org/10.1162/rest.91.4.682>
5. EPA. (2019). Ozone Pollution: US Environmental Protection Agency. [Available from: <https://www.epa.gov/ozone-pollution>].
6. Fernández, I.C., & Wu, J. (2016). Assessing environmental inequalities in the city of Santiago (Chile) with a hierarchical multiscale approach. *Applied Geography*, 74, 160-169. <https://doi.org/10.1016/j.apgeog.2016.07.012>
7. Gramsch E., Cereceda-Balic F., Oyola P., & von Baer D. (2006). Examination of pollution trends in Santiago de Chile with cluster analysis of PM10 and ozone data. *Atmospheric Environment*, 40(28), 5464-5475. <https://doi.org/10.1016/j.atmosenv.2006.03.062>
8. JPAL. (2019). Tailoring instruction to students' learning levels to increase learning: J-PAL Abdul Latif Jameel Poverty Action Lab. [Available from: <https://www.povertyactionlab.org/policy-insight/tailoring-instruction-students-learning-levels-increase-learning>].
9. Kampa M., & Castanas E. (2008). Human health effects of air pollution. *Environmental Pollution*, 151(2), 362-367. <https://doi.org/10.1016/j.envpol.2007.06.012>

10. Lavy V., Ebenstein A., & Roth S. (2014). The impact of short term exposure to ambient air pollution on cognitive performance and human capital formation. National Bureau of Economic Research. [Available from: <http://www.nber.org/papers/w20648>
11. Levinson A. (2012). Valuing public goods using happiness data: The case of air quality. *Journal of Public Economics*, 96(9-10), 869-880. <https://doi.org/10.1016/j.jpubeco.2012.06.007>
12. Marcotte D.E. (2015). Allergy test: Seasonal allergens and performance in school. *Journal of Health Economics*, 40, 132-140. <https://doi.org/10.1016/j.jhealeco.2015.01.002>
13. Mohai P., Kweon B-S., Lee S., & Ard K. (2011). Air pollution around schools is linked to poorer student health and academic performance. *Health Affairs*, 30(5), 852-862. <https://doi.org/10.1377/hlthaff.2011.0077>
14. OECD. (2012). Equity and Quality in Education: Supporting Disadvantaged Students and Schools: Organization for Economic Cooperation and Development. [Available from: <https://www.oecd.org/education/school/50293148.pdf>.
15. Pastor M., Morello-Frosch R., & Sadd J.L. (2006). Breathless: schools, air toxics, and environmental justice in California. *Policy Studies Journal*, 34(3), 337-362. <https://doi.org/10.1111/j.1541-0072.2006.00176.x>
16. Sabatini, F., Wormald, G., Sierralta, C., & Peters, P.A. (2009). Residential segregation in Santiago: Scaled-related effects and trends, 1992-2002. In Roberts, B.R., and Wilson, R.H. (eds) *Urban Segregation and Governance in the Americas*. Palgrave Macmillan, New York. https://doi.org/10.1057/9780230620841_7
17. Ransom M.R., & Pope C.A. (1992). Elementary school absences and PM10 pollution in Utah Valley. *Environmental Research*, 58(1-2), 204-219. [https://doi.org/10.1016/S0013-9351\(05\)80216-6](https://doi.org/10.1016/S0013-9351(05)80216-6)
18. Rojas-Vallejos, J., & Lastuka, A. (2020). The income inequality and carbon emissions trade-off revisited. *Energy Policy*, 139, 111302. <https://doi.org/10.1016/j.enpol.2020.111302>

19. Román-Collado, R., & Jiménez de Reyna, J. (2019). The economic benefits of fulfilling the World Health Organization's limits for particulates: A case study in Algeciras Bay (Spain). *Journal of the Air & Waste Management Association*, 69(4), 438-449. <https://doi.org/10.1080/10962247.2018.1544178>
20. Schmitz, R. (2005). Modelling of air pollution dispersion in Santiago de Chile. *Atmospheric Environment*, 39(11), 2035-2047. <https://doi.org/10.1016/j.atmosenv.2004.12.033>
21. Schwartz J. (2004). Air pollution and children's health. *Pediatrics*, 113(Supplement 3), 1037-1043. <https://doi.org/10.1542/peds.113.4.S1.1037>
22. Toro R., Córdova A., Canales M., & Mardones P. (2015). Trends and threshold exceedances analysis of airborne pollen concentrations in Metropolitan Santiago Chile. *PloS One*, 10(5), 1-21. <https://doi.org/10.1371/journal.pone.0123077>
23. WHO. (2016). Ambient Outdoor Air Quality and Health: World Health Organization. [Available from: <http://www.who.int/mediacentre/factsheets/fs313/en>].