University of Hawai`i at Mānoa Department of Economics Working Paper Series

Saunders Hall 542, 2424 Maile Way,
Honolulu, HI 96822
Phone: (808) 956 -8496
www.economics.hawaii.edu

Working Paper No. 22-03

VOG: Using Volcanic Eruptions to Estimate the Impact of Air Pollution on Student Learning Outcomes

By
Timothy J. Halliday
Rachel Inafuku
Lester Lusher
Aureo de Paula

November 2022

VOG: Using Volcanic Eruptions to Estimate the Impact of Air Pollution on Student Learning Outcomes *

Timothy J. Halliday†
University of Hawai'i at Mānoa
IZA

Rachel Inafuku
University of Hawai'i at Mānoa

Lester Lusher
University of Hawai'i at Mānoa
IZA

Aureo de Paula
University College London

November 1, 2022

Abstract

This study pairs variation stemming from volcanic eruptions from Kilauea with the census of Hawai'i's public schools student test scores to estimate the impact of particulates and sulfur dioxide on student performance. We leverage spatial correlations in pollution in conjunction with proximity to Kilauea and wind direction to construct predictions of pollution exposure at each school. We precisely estimate that increased particulate pollution leads to a small but statistically significant drop in average test scores. Then, utilizing Hawai'i's rich diversity across schools in baseline exposure, we estimate sharp nonlinearities - schools with higher baseline levels of pollution experience larger decreases in test scores than schools with less pollution exposure on average. At levels of particulate pollution higher than six micrograms per cubic meter ($\mu g/m^3$), we estimate that a one standard deviation increase in $PM_{2.5}$ leads to a decline in test scores of 1.1 percent of a standard deviation. Lastly, we find that *within schools* the drop in test scores is concentrated among economically disadvantaged students. The effects of $PM_{2.5}$ on student test scores are larger by a factor of ten for the poorest pupils. Similarly, the effects of SO_2 are larger by a factor of six. We demonstrate that poor air quality disproportionately impacts the human capital accumulation of economically disadvantaged children.

*We thank seminar participants at the University of Hawaii Applied Micro Workshop for useful comments. We are especially indebted to Hawai'i P-20 for allowing us to access the data and for strong support of this research. This research is supported by the Institute of Education Sciences, U.S. Department of Education, through Grant 84.372A (Statewide Longitudinal Data Systems) to the Hawai'i State Department of Education.

†Corresponding Author. Address: 2424 Maile Way; 533 Saunders Hall; Honolulu, HI 96822. e-mail: halliday@hawaii.edu

Key Words: Vog, Particulates, Test Scores, Kriging, Environmental Justice
JEL Classification: I22, I24, Q52

1 Introduction

Researchers and policymakers have increasingly sought to understand the consequences of poor air quality. An abundance of evidence suggests pollution can have severe adverse effects on health, fertility, and mortality outcomes.[1] A smaller yet growing set of studies has identified labor productivity losses, where pollution harms workers across both physically demanding (e.g. fruit picking (Graff Zivin and Neidell, 2012) and pear-packing (Chang et al., 2016)) and mentally demanding (e.g. baseball umpiring (Archsmith et al., 2018)) occupations.[2]

However, despite some evidence of air pollution sharply reducing cognitive performance (Zhang et al., 2018), there are surprisingly few empirical investigations into how pollution affects student test scores. This is a pity as student scores not only are a marker of cognitive performance but they also often have long term consequences. A better understanding of this topic, therefore, has implications for how air quality impacts human capital acquisition and subsequent labor market outcomes. In this study, we investigate how air quality affects student performance on standardized tests.

An important feature of our study is that we pay close attention to how poor air quality affects poorer students within schools. It has long been understood that air pollution disproportionately impacts the poor and disadvantaged minorities in the United States despite recent progress (Currie et al., 2020). In this paper, we will shed light on how air pollution affects human capital acquisition differentially by socioeconomic status which is something that is not currently well understood.

There is a growing number of studies on the effects of specific pollutants (primarily par-

[1] Several studies have found a positive association between pollution and fertility abnormalities (Nieuwenhuijsen et al., 2014; Slama et al., 2013; Perin et al., 2010). See Carré et al. (2017) for a review of the literature. Burnett et al. (1999) and Koken et al. (2003) find increases in air pollution leads to an increase in cardiorespiratory hospitalizations. Linares et al. (2010) find children who attend schools closer to major air pollution sources were more likely to develop respiratory and lung abnormalities. Di et al. (2017) finds increases in pollution (even at levels below the national standard) were associated with an increase in mortality for US Medicare beneficiaries, especially amongst racial minority groups.

[2] Other studies examining labor productivity loss include He et al. (2019) (textile industry workers), Chang et al. (2019) (call center workers), and Lichter et al. (2017) (professional soccer players). For a comprehensive review of the literature on "non-health" effects of air pollution, we refer the reader to Aguilar-Gomez et al. (2022).

ticulates) on student test scores during high school.[3] Work from Ebenstein et al. (2016) uses Israeli data from 2000 to 2002 to find drops in high school exit exam test scores and worsened longer run outcomes in response to poor air quality. These effects were especially large for those of lower socioeconomic status. Next, Marcotte (2017) found decreased performance among kindergartners on testing days with worse pollen and fine airborne particulate matter. Similarly, Heissel et al. (2020) identify the effects of traffic pollution on student test scores and other shorter run outcomes in Florida. Finally, Carneiro et al. (2021) show that higher concentrations of particulates result in lower scores on college entry examinations in Brazil.[4] Despite this, there is still a dearth of studies on the topic, particularly, relative to the volume of research on the health impacts of poor air quality.

Importantly, work that demonstrates disproportionate effects of poor air quality on test scores by socioeconomic status is limited. Moreover, the results that we do have often conflict leaving the issue still up for debate. Ebenstein et al. (2016) finds that Israeli students of lower socioeconomic status experience larger declines in test scores due to higher pollution. They attribute this finding to higher rates of asthma among those of lower socioeconomic classes.[5] However, seminal work by Case et al. (2002) shows that asthma is more prevalent among children with richer parents in the United States. On the other hand, Heissel et al. (2020) show that economically disadvantaged students (as proxied by eligibility for free or reduced lunch) experience smaller impacts on test scores compared to their more advantaged peers. However, they also experience more absences and behavioral issues. Accordingly, there is not a clear consensus within the literature suggesting that poor air quality disproportionately impacts learning outcomes of poorer students.

The context of our study is the Hawaiian islands. This is a particularly advantageous setting for several reasons. A primary advantage comes from Hawai‘i's rich, plausibly exogenous vari-

[3] Bedi et al. (2021) also investigate the impacts of $PM_{2.5}$ on grammatical reasoning tests of university students in Brazil. They find evidence of adverse effects.

[4] Outside student performance outcomes, related studies from Currie et al. (2009), Liu and Salvo (2018), and Chen et al. (2018) find increased student absences in response to poor air quality.

[5] Marcotte (2017) also shows that the effects of particulate pollution are largest for asthmatic students.

ation in air quality. Despite its reputation for moderate climate, Hawai‘i can claim ten of the world's 14 classifications for climate zones (microclimates) - the only place in the world with such diversity in one small area.[6]

Hawai‘i provides a unique and powerful opportunity to estimate the effects of two pollutants, particulate matter ($PM_{2.5}$) and sulfur dioxide (SO_2), on cognitive performance. We do so using SO_2 emissions from Kilauea volcano which is located on the island of Hawai'i. These gaseous emissions eventually form particulate matter in the form of sulfate aerosols. This pollution is called vog and is similar to smog pollution in many cities. Because this species of particulates is high in sulfuric acid, they resemble particulates from sources that produce sulfate aerosols such as coal-fired power plants. Importantly, 8% of the world's population faces potential risks from volcanic eruptions and so, our estimates will have a direct bearing on these other settings (Choumert-Nkolo et al., 2021).

The emission of SO_2 from the Kilauea volcano represents a rare case of truly unpredictable variation in air pollution in the United States. Based on local wind conditions and whether the volcano is emitting, the air quality of Hawai‘i can shift from hazardous to pristine in a matter of hours across differing parts of the islands. Previous research has leveraged this high frequency variation on a day-to-day basis to find increased emergency room admissions due to respiratory reasons on days with higher pollution levels (Halliday et al., 2019).[7]

An additional advantage of our setting is that baseline pollution levels are far below Environmental Protection Agency (EPA) ambient air quality standards. Identifying and understanding the effects of pollution at lower baseline levels is important as this can help to inform and potentially update EPA standards. Moreover, lower pollution levels also better reflect modal households in the US. While prior literature has focused entirely on air quality within

[6]Source: Hawai‘i Magazine, *https://www.hawaiimagazine.com/content/hawaii-has-10-worlds-14-climate-zones-explorers-guide-each-them*, (*accessed 16 Sep. 2020*)

[7]Halliday et al. (2019) articulate the numerous advantages of using variation in vog to study the impact of pollution. For example, vog is emitted naturally, whereas the majority of the literature relies on variation in human activity (e.g. from cars, airplanes, factories) which may plausibly suffer from endogeneity biases. Another advantage comes from temporal variation: vog can vary on a day-to-day basis, whereas most other types of pollutants are highly serially correlated.

higher-baseline polluted environments, baseline pollutant levels in our study are comparable to pollution within the US. In 2019, the US average population-weighted concentration of particulate matter ($PM_{2.5}$) pollution was 7.65 micrograms per cubic meter (μg/m^3).[8,9] In our sample, Hawai'i island saw similar baseline $PM_{2.5}$ levels with an average of 8.27 $\mu g/m^3$ while baseline levels across other islands were in the three to four $\mu g/m^3$ range.

We pair this variation in particulates with the census of public school student test scores in the State of Hawai‘i. These data were obtained from the Hawai‘i P-20 Partnerships for Education (Hawai‘i P-20) initiative, a partnership between the University of Hawai‘i, the Executive Office of Early Learning, and the State of Hawai‘i Department of Education. Because we have a census, we will have enough power to detect even small impacts of air pollution. The data track students from elementary to middle and high-school from 2015 through 2018. Math and English literacy assessments are given in grades three through eight, and again in grade eleven. In total, the data include over 350,000 student-test observations. These data allow us to estimate day-of measures of air quality on student performance across varying ages, assessment types, and air quality conditions.

An important feature of this study is that we employ a technique from geostatistics called Kriging to predict pollution exposure at each school (Cressie, 1990; Montero et al., 2015). Specifically, we leverage information on the spatial correlation in pollution as well as the distance between and the relative locations of the pollution monitoring stations and the schools. We also exploit the fact that the presence of northeasterly winds, or trade winds, affects the spatial distribution of pollution in Hawaii. In general, trade winds lower pollution levels throughout most of the archipelago. One advantage of Kriging in our context is that it does not require that predictions are inside the simplex generated by the monitoring stations used in the prediction. This is desirable for us as many of the schools in our sample are located far away from the monitoring stations.

[8] For example, in Ebenstein et al. (2016) the average level of $PM_{2.5}$ on student test days was 21.05 μg/m^3.

[9] Source: US Environmental Protection Agency, *https://www.epa.gov/air-trends/particulate-matter-pm25-trends*, (*accessed 25 Apr. 2021*)

For the full sample of student test scores, we estimate a small but statistically significant impact of particulates on student test scores. A one standard deviation increase in $PM_{2.5}$ reduces test scores by 0.24 percent of a standard deviation. We then find that the effects are significantly tied to particulate levels at baseline. For example, schools with baseline $PM_{2.5}$ levels between three to six $\mu g/m^3$ see a drop in test scores of about 0.37 to 0.74 percent of a standard deviation for every standard deviation increase in $PM_{2.5}$. When subsetting our regression sample to schools with baseline $PM_{2.5}$ levels above six $\mu g/m^3$, we see reductions in the neighborhood of 1.1 percent of a standard deviation with respect to a one standard deviation increase in $PM_{2.5}$. Furthermore, our findings are concentrated amongst schools in south Hawaii, the region of the state that sees the highest level of pollution exposure on average. This suggests that the damages from pollutants increase precipitously with baseline exposure, yet are still present in relatively lower baseline environments.

The effects of SO_2 are more muted and nuanced. For the full sample, we do not find effects. However, we do estimate statistically significant effects on south Hawaii. These estimate also have a higher magnitude. A one standard deviation increase in SO_2 pollution decreases test scores in south Hawaii by 2.14 percent of a standard deviation. One important caveat with the results in south Hawaii (which includes Kilauea) is that we find that both $PM_{2.5}$ and SO_2 adversely impact student outcomes. Because both pollutants are highly correlated, we have not separately identified the effects on each pollutant.[10]

Lastly, and perhaps most importantly, we find that the effects of pollution are particularly concentrated among economically disadvantaged students. Students who are economically disadvantaged experience ten times the effect of $PM_{2.5}$ and six times the effect of SO_2 on exam scores when compared to their more advantaged counterparts. Interestingly, we find little difference in the effects of particulates across schools by the fraction of the school's students who

[10] We note that the inability to cleanly identify one pollutant from another is a common issue in the literature. In this regard, this paper is no different. Halliday et al. (2019) were able to cleanly identify the effects on particulates on emergency medical care. However, their design used data exclusively from Oahu where the only pollutant is $PM_{2.5}$. In the current setting, a large portion of the impacts occur close to Kilauea where SO_2 levels are extremely high as well.

were disadvantaged, suggesting that the economically disadvantaged student gap is not driven by differences in school resources (e.g. classroom air conditioning). We also note that this is less of an issue in Hawaii as there is one statewide school district in which schools are not funded by local property taxes. So, we conclude that disadvantaged students *within* the same school are significantly more harmed by pollution than their more advantaged counterparts. This result has obvious implications for environmental justice and our understanding of how environmental laws, regulations, and policies may disproportionately harm people from lower income and/or minority groups.

2 Data and Background

2.1 Student Learning Outcome Measurements

We measure student learning outcomes using data from the Hawaii P-20 Partnerships for Education (Hawaii P-20). The Hawaii P-20 manages student level data collected through the Data eXchange Partnership (DXP), a collaboration between five of Hawaii's state agencies (Department of Health, Department of Labor and Industrial Relations, Department of Education, Department of Human Services, and University of Hawaii). The data from the DXP consists of all students in Hawaii's public school system spanning elementary through secondary education. The data includes education performance measures as well as demographic characteristics of the student. Test score data come from the Smarter Balanced Assessment (SBA). The SBA is an annual assessment of college and career readiness that includes modules on math and English literacy. It has been administered to students in grades three through eight and grade ten since 2015.[11] We standardize test scores to a mean of zero and a standard deviation of one at the grade-module-year level.

[11] Prior to 2015, the DOE administered the Hawaii State Reading and Math Assessment (HSA) to measure student performance. The HSA was administered to students in grades three through eight and ten. Though our data includes test scores from the HSA, we do not have data on test dates, and thus we strictly focus on utilizing data from the SBA.

Though the SBA is mandatory for all public school students, test dates are unique at the student level. Each school year, the DOE provides a one to three month testing window within which schools are required to administer the modules. Each school is then individually responsible for determining the exact date that students take their assessments. Importantly, schools determine their exam dates at the beginning of the school year, well before the school can forecast potential weather conditions or vog levels on the exam date. Schools typically have students within the same grade take the same module on the same date. The two modules (math and reading) are always taken on separate dates. In some circumstances, school faculty are authorized to have some students take the exams earlier or later than their peers. Since students with cognitive disabilities are subject to alternative assessments, we drop them from our sample.

2.2 Air Quality Measurements

We employ data on particulates ($PM_{2.5}$) and sulfur dioxide (SO_2) obtained from the State of Hawaii Department of Health (DOH).[12] Particulates are measured in micrograms per cubic meter ($\mu g/m^3$). $PM_{2.5}$ measures particulates that are 2.5 micrometers in diameter or smaller. SO_2 is measured in parts per billion (ppb). The DOH reports measures of each pollutant at hourly frequencies. For our analysis, we aggregate the pollutant measures from each DOH monitoring station to 24-hour averages and merge these data with the Hawaii P-20 data using the date that students took their assessments.

2.3 Summary Statistics

Table 1 displays summary statistics from the P-20 student data. At the student-year level, the mean month of the math and reading exam is just over four, indicating that students tend to take both of their assessments in April. About half of student-years in the public school system come from economically disadvantaged families and around six percent received English

[12] We do not use data on PM_{10} as the state only had three stations monitoring it.

Table 1: Summary Statistics (Hawaii P-20 Student Measures)

	Mean	Std. Dev.
Panel A: Student-Year Level Statistics		
Month of Math Exam	4.43	0.69
Month of Reading Exam	4.12	0.66
Economically Disadvantaged	0.49	0.50
Received English Language Services	0.06	0.24
Panel B: Student Level Statistics		
Female	0.50	0.50
Asian (Non-Filipino)	0.16	0.37
Filipino	0.24	0.43
Native Hawaiian	0.24	0.43
Pacific Islander	0.09	0.29
White	0.18	0.39
Other Ethnicity	0.08	0.27
Unique Individuals	153,448	
Schools	282	
Years	2015 - 2018	

Notes: Data on student summary statistics comes from the Hawaii P-20 Partnerships for Education. Economically disadvantaged students refer to those who are enrolled in free lunch programs through the Department of Education. Those who received English language services are students who enrolled in the State of Hawaii Department of Education's English Learner Program for the academic school year.

language services. Table 1 also reveals Hawai'i's ethnically diverse population. Nearly a quarter of students identify as Native Hawaiian and another quarter identify as Filipino. Another 16% of students identify as non-Filipino Asian, 9% as Pacific Islander, 18% as White, and 8% identify as another ethnicity. The data include 153,473 unique individuals enrolled across 282 schools.

Summary statistics for pollution are presented in Table 2. Overall, $PM_{2.5}$ averages are relatively similar across the islands of Oahu, Maui, and Kauai. Hawaii island sees slightly higher levels of $PM_{2.5}$ in certain areas due to Kilauea's volcanic activity. The Pahala monitoring station is located less than 20 miles south of the Kilauea volcano. Because of the Pahala monitoring station's close proximity to the volcano, average levels of SO_2 in Pahala are about four times the state average. For the full sample, the average $PM_{2.5}$ is 6.14 μg/m^3 (with a standard deviation of 1.84) and the average SO_2 is 8.69 ppb (with a standard deviation of 6.17).

Table 2: Summary Statistics (Pollutant Measures)

	$PM_{2.5}$		SO_2	
Station	Mean	Std. Dev.	Mean	Std. Dev.
Hawaii Island				
Hilo	7.86	5.19	3.65	7.41
Kona	11.25	5.52	3.67	3.32
Mountain View	3.68	3.23	1.80	2.59
Ocean View	12.64	4.77	17.38	16.24
Pahala	4.77	2.59	27.90	21.98
Hawaii Island Average	8.27	3.35	11.98	8.37
Oahu				
Honolulu	3.71	2.44	0.38	0.48
Kapolei	4.95	2.03	0.14	0.32
Pearl City	4.09	2.05		
Sand Island	4.95	1.93		
Oahu Average	4.39	1.76	0.26	0.34
Maui				
Kahului	3.23	1.78		
Kihei	4.37	3.60		
Maui Average	4.07	2.53		
Kauai				
Niumalu	3.33	2.67	0.73	0.60
State Average	6.14	1.84	8.69	6.17

Notes: Data on pollutant measures come from the State of Hawaii Department of Health. Measures of $PM_{2.5}$ and SO_2 are reported for each pollutant monitoring station. The particulate $PM_{2.5}$ is reported in $\mu g/m^3$ and SO_2 is reported in ppb.

3 Research Design

3.1 Measuring Pollution at Schools

We begin by discussing how we use pollution measurements from monitoring stations in conjunction with variation in Hawaii's trade wind patterns to predict pollution at each of Hawaii's schools. We do this because of the spatial misalignment of the pollution monitoring stations and the schools. To do this, we use a Kriging procedure which delivers the best linear unbiased predictor (BLUP) of unobserved pollution at each school (Cressie, 1990). Predicted exposure at a given school and on a given day is a weighted average of the available measurements on that day within some vicinity of the school (Cressie, 1990). Normally, Kriging weights depend solely on the spatial correlations of pollution measurements across monitoring stations. However, we extend the procedure so that we can incorporate external variables (wind direction in our case) to generate more accurate predictions.

The weights that we employ depend on the distance between school s and monitoring station m (d_{sm}), the relative location of the monitoring station vis-a-vis the school (l_{sm}), and the wind direction on that day (NE_t). We denote the weight given to monitoring station m to predict pollution at school s as $\lambda(d_{sm}, l_{sm}, NE_t) \equiv \lambda_{sm}(NE_t)$. In the spirit of Halliday et al. (2019), we employ the variable NE_t, a binary variable indicating that the winds on that day were north-easterly. As previously discussed, such winds are called "trade winds" and tend to improve air quality throughout the state. If we denote the measured pollution at monitoring station m on day t with Π_{mt}, our prediction of exposure at school s on day t is

$$\widehat{P}_{st} = \sum_{m \in N(s)} \lambda_{sm}(NE_t)\Pi_{mt} \tag{1}$$

where $N(s)$ denotes a neighborhood of school s.[13] The weights are constrained to sum to one

[13] In fact, kriging is a "universal" predictor in that it uses all available measurements to make predictions. However, in non-stationary environments such as this, it is advisable to use more local measures. With this in mind, we employ three neighborhoods: Oahu, south Hawaii (includes Kilauea), and north Hawaii (the rest of the island of

which ensures that the predictor is unbiased. As many predictors can be written in the form of equation (1) including nearest neighbor and inverse distance weighting predictors, Kriging predictions will have lower prediction errors than many commonly used predictors as it is the BLUP.

While Kriging weights must sum to one, they can be negative or greater than unity. This allows the predictions to take on a value outside of the simplex generated by the pollution measurements. This is, in fact, a positive feature of Kriging - not a deficiency. To see this, we note that the monitoring stations on the island of Oahu are all in urban Honolulu on the southern shore of the island. However, many schools on this island are in rural parts of the island and/or on the northern facing shores placing them outside of the simplex generated by the monitoring stations. Because the weights are not constrained to be between zero and one, the predictions at these rural schools can be smaller than *all* of the Π_{mt} used to construct $\widehat{P}_{st}$. Other common predictors used in this literature such as nearest neighbor or inverse distance weighting do not share this property.

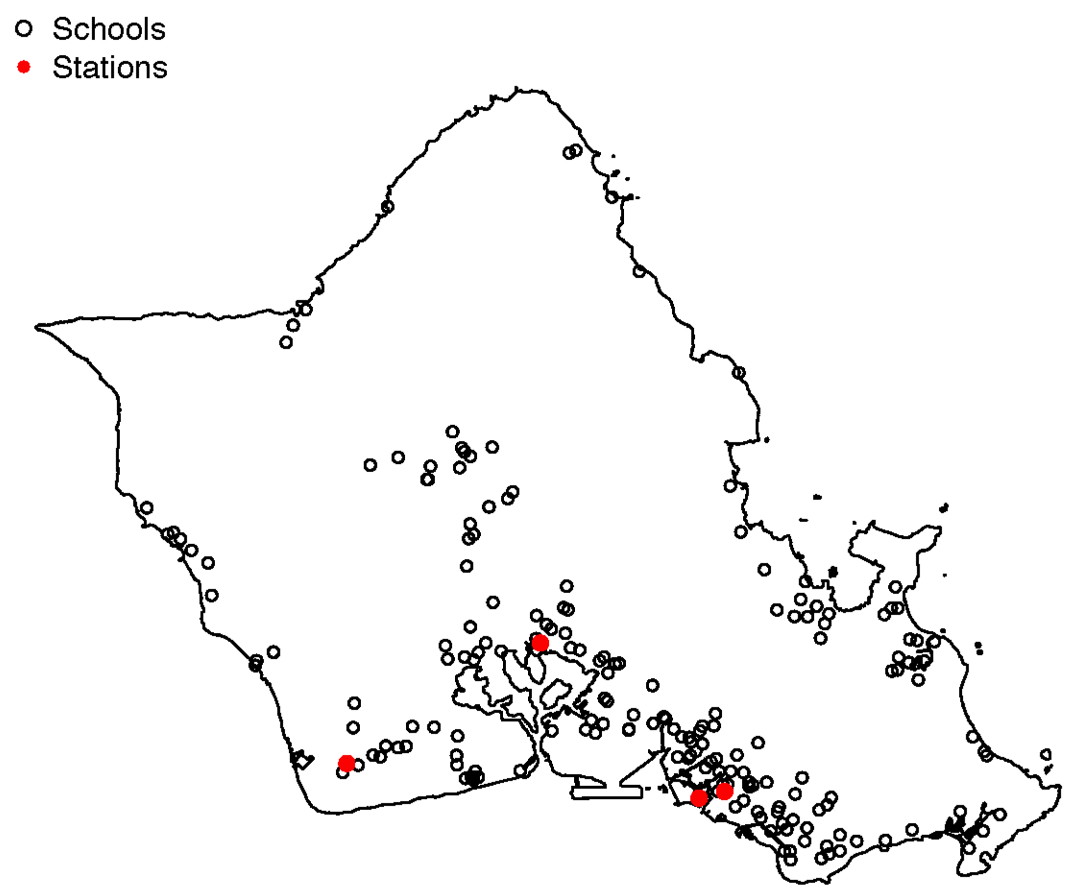
To illustrate the relative locations of the schools and the monitoring stations, we present Figures 1 and 2. These figures plot each school's location and each monitoring station's location (denoted by the solid red circles) on the islands of Oahu and Hawaii, respectively. Schools on Oahu are empty circles. Schools on south Hawaii are empty triangle whereas those on north Hawaii are crosses. These correspond to the three neighborhoods for which we compute the Kriging weights. The figure illustrates that the schools in our sample are often located well outside of the simplex generated by the monitoring stations.

We now briefly discuss the GMM procedure that we use to estimate the Kriging weights. A detailed treatment of this can be found in Appendix A.1. Denoting $M(s) \equiv \#N(s)$, we define

$$\lambda_s(b) \equiv (\lambda_{s1}(b), ..., \lambda_{sM(s)}(b), \alpha_b)'$$

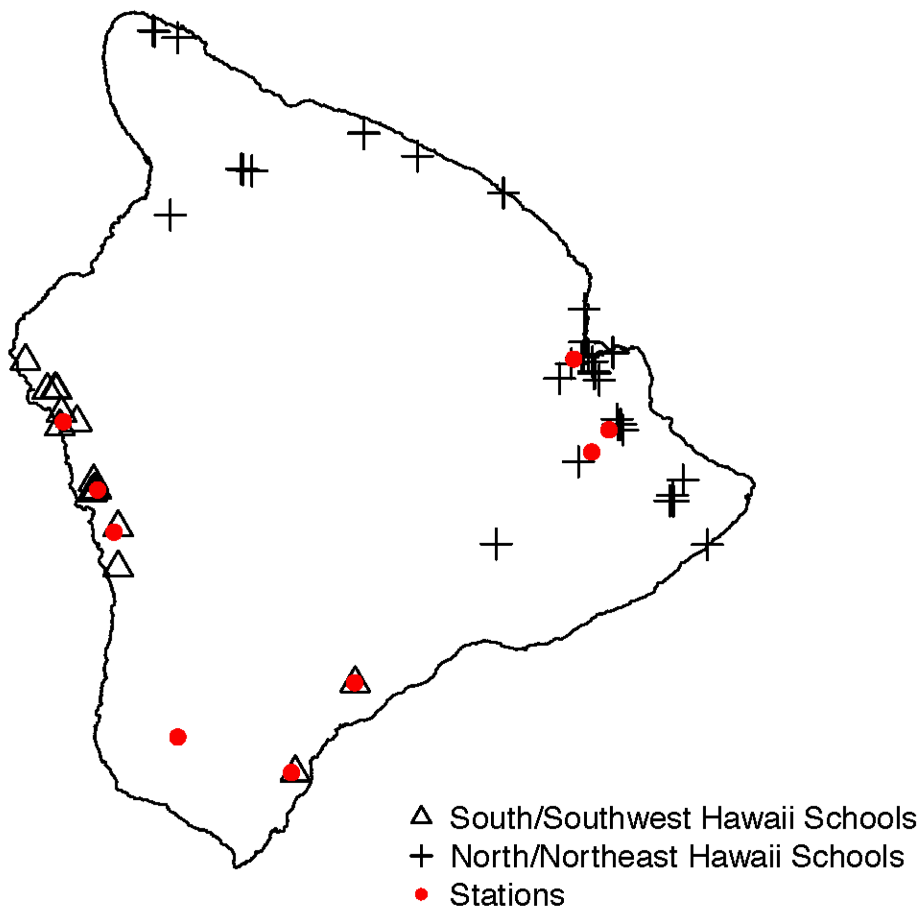
Hawaii).

Figure 1: Schools and Monitoring Stations on Oahu



Notes: This figure displays all pollutant monitoring stations and schools located on the island of Oahu.

Figure 2: Schools and Monitoring Stations on Hawaii Island



Notes: This figure displays all pollutant monitoring stations and schools located on the island of Hawaii.

which is a vector that includes the Kriging weights, $\lambda_{sm}(b)$ for $b \in \{0, 1\}$ (b is an indicator for northeasterly winds), and the Lagrangian multiplier on the constraint that the weights must sum to one denoted by α_b for $b \in \{0, 1\}$. The weights depend on the semivariogram between stations m and n on trade wind and non-trade wind days. The semivariogram is one minus the spatial correlation between the two locations and equals zero at a given location *i.e.* We denote this measure by $\gamma_{mn}(b)$. when $m = n$. The Kriging weights can then be derived as $\lambda_s(b) = \Gamma(b)^{-1}\Gamma_s(b)$ where

$$\Gamma(b) \equiv \begin{bmatrix} \gamma_{11}(b) & \dots & \gamma_{1M(s)}(b) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{M(s)1}(b) & \dots & \gamma_{M(s)M(s)}(b) & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix}$$

and

$$\Gamma_s(b) \equiv \begin{bmatrix} \gamma_{1s}(b) \\ \vdots \\ \gamma_{M(s)s}(b) \\ 1 \end{bmatrix}.$$

Hence, the task of computing the Kriging weights hinges on computing the semivariogram for $NE_t = 1$ and $NE_t = 0$. We refer the reader to the appendix for details on the derivation and, specifically, what optimization problem delivers these weights.

To compute the semi-variograms in $\Gamma(b)$ and, especially, $\Gamma_s(b)$ (which requires out-of-sample prediction), we postulate a parametric model indexed by a vector. To estimate this parameter, first, we compute the empirical semivariagrams for trade wind ($NE_t = 1$) and non-trade wind days ($NE_t = 0$). Next, we model the spatial correlation between stations m and n as

$$1 - \gamma_{mn}(NE_t) = \exp\left(-d_{mn} \times \left(\phi NE_t + \sum_{j \in \mathcal{L}} (\delta_j 1_{mn}(j) + \beta_j \times 1_{mn}(j) \times NE_t)\right)\right).$$

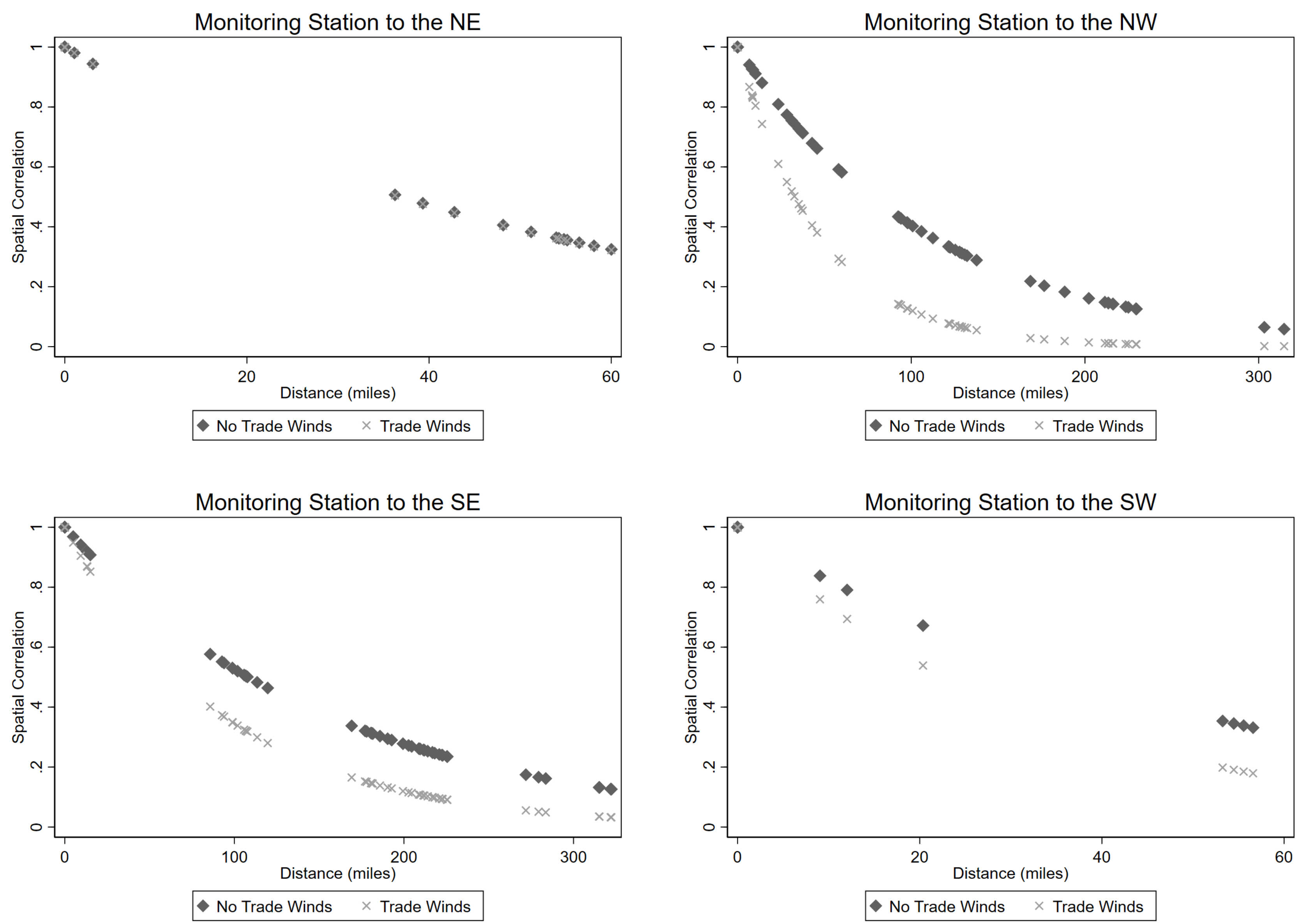
This functional form allows the spatial correlation: (1) to decline with the distance between locations; (2) to decline when trade winds are blowing; (3) to depend on relative locations according to wind direction. The semivariogram is zero when $d_{mn} = 0$. We then note that estimation of the parametric model can proceed using a simple Poisson regression or Generalized Linear Model (with a log link function) package in Stata or R and is relatively easy to implement.

3.2 Covariogram Estimates

We plot the covariograms for $PM_{2.5}$ and SO_2 in Figures 3 and 4. Each figure has four plots corresponding to the relative locations of the monitoring stations: NE, SE, SW, and NW. For any pair of stations, (m, n), the location corresponds to m relative to n. We also plot the covariograms for trade and non-wind days in each plot. Each dot in these plots is a pair of stations.[14]

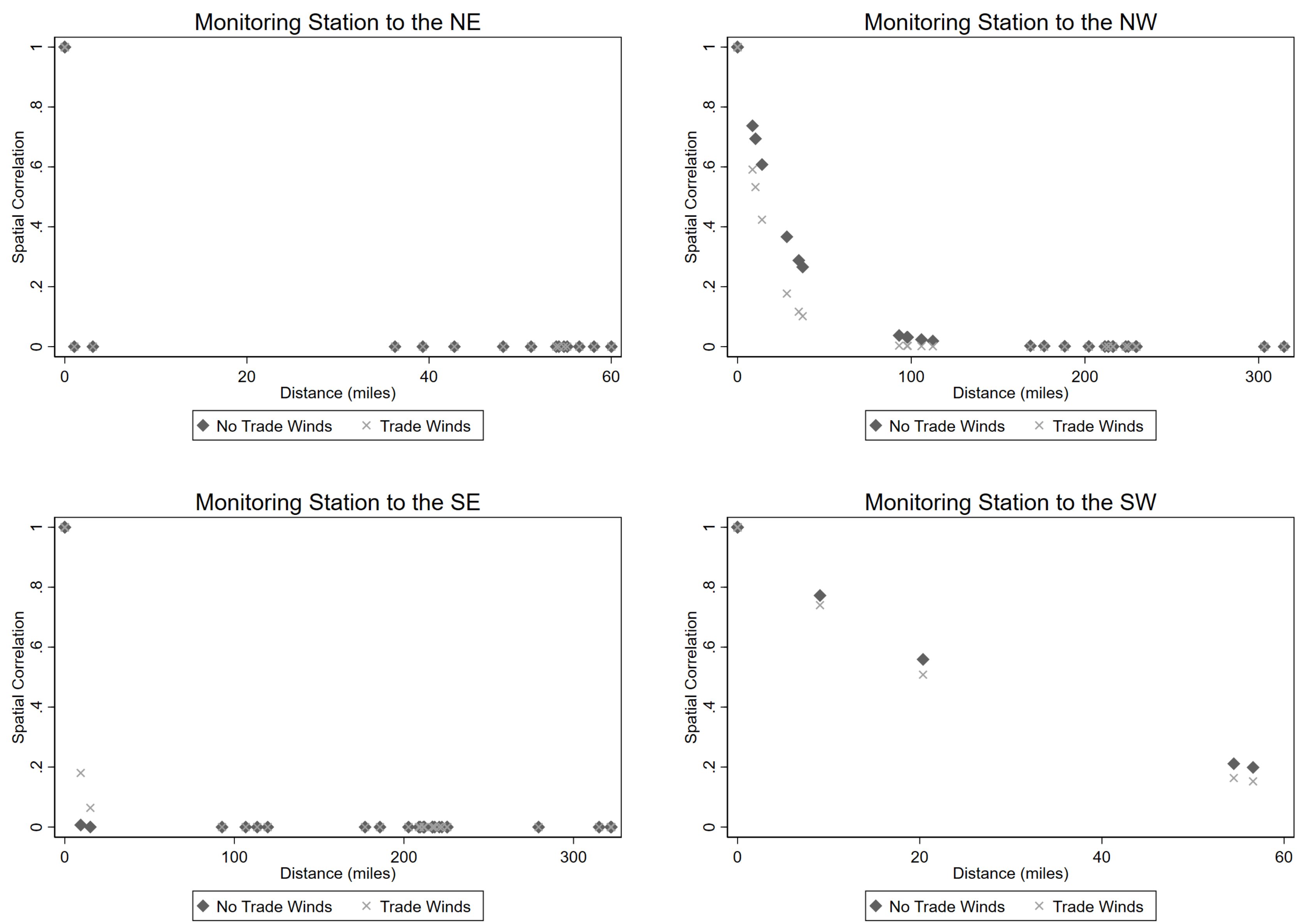
[14]The GMM procedure that we use only employs the bottom triangle of the covariance matrix since the information in the upper triangle is redundant. This implies, however, that the samples used in the off-diagonal plots in the top right and bottom left plots of Figures 3 and 4 are different and so the estimations are also different. For example, for a pair of stations (m, n), if m is northeast of n then $1_{mn}(NE) = 1$ and $1_{mn}(SW) = 0$. On the other hand, for (n, m), we would have $1_{mn}(NE) = 0$ and $1_{mn}(SW) = 1$. The same is true for the plots on the diagonal of the figure (not the covariance matrix). This is highly technical point but some readers may have wondered why these plots are different. This is why.

Figure 3: Covariograms for $PM_{2.5}$



Notes: Displays estimated covariograms by the relative locations of the monitoring stations and by the direction of the trade winds.

Figure 4: Covariograms for SO_2



Notes: Per Figure 3

In Figure 3, we see a large degree of spatial correlation in $PM_{2.5}$. In the diagonal plots corresponding to relative locations NE and SW, we see that the spatial correlation remains above 0.2 for up to a distance of 60 miles. In the off-diagonal plots corresponding to NW and SE, the spatial correlations remain well above zero when there are no trade winds, but they are substantially smaller when the trade winds are blowing. Overall, this figure strongly indicates that trade winds result in lower spatial correlations in $PM_{2.5}$.

Figure 4 shows a much more muted degree of spatial correlation for SO_2. When the monitoring stations are either to the northeast or southeast, the figure shows that there is essentially no spatial correlation in SO_2. When the stations are to the northwest or the southwest, there appears to be moderate spatial correlation through about 100 miles and trade winds modestly dampen it. The more modest degree of spatial correlation in SO_2 makes it more difficult to predict SO_2 exposure at schools. This might underscore the noisier estimate of the effects of SO_2 on outcomes that we will present.

An important feature of Kriging is that it delivers the arithmetic mean within the neighborhood of the school when there is no spatial correlation. This implies that the Kriging predictions for SO_2 will be closer to the neighborhood means than they are for $PM_{2.5}$. All told, the spatial correlations in $PM_{2.5}$ are more informative than they are in SO_2.

In Table 3, we provide the means and standard deviations of the predictions of pollution at each school on days with and without trade winds.[15] We provide descriptive statistics for the three neighborhoods for which we computed the weights: Oahu, south Hawaii, and north Hawaii. The table shows that on trade wind days levels of $\text{PM}_{2.5}$ are lower on Oahu (2.94 vs. 4.20 $\mu g/m^3$) as well as on north Hawaii (5.65 vs. 7.07 $\mu g/m^3$). However, we do not see lower particulate levels on trade wind days on south Hawaii. Presumably, the reason for this is that many of the schools on south Hawaii are located to the west of Kilauea and, so the trade winds

[15] We replaced the prediction with a zero when the Kriging prediction was negative. We also replaced all predictions above the 99th percentile with a missing value. In the raw pollution data, 22.10% of $PM_{2.5}$ predictions were negative and 1.08% of SO_2 predictions were negative. These numbers might seem large but we note per Halliday et al. (2019) many days in those data essentially had very low pollution levels.

blow $PM_{2.5}$ towards those schools. Finally, we see a similar pattern for SO_2 with lower levels on trade wind days on Oahu and north Hawaii.

Table 3: Summary Statistics (Pollutant Measures) by Tradewind Status

	Oahu		South/Southwest Hawaii		North/Northeast Hawaii	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
No Tradewinds						
$PM_{2.5}$	4.20	3.11	7.71	6.79	7.07	5.21
SO_2	0.80	0.78	4.95	4.52	3.55	4.03
Tradewinds						
$PM_{2.5}$	2.94	6.12	8.87	6.44	5.65	4.01
SO_2	0.60	0.63	6.79	4.94	1.36	0.98

Notes: Data on pollutant measures come from the State of Hawaii Department of Health. Measures of $PM_{2.5}$ and SO_2 are reported for Oahu, West/Northwest Hawaii island, and East/Southeast Hawaii for days with and without tradewinds (northeasterly winds). The particulate $PM_{2.5}$ is reported in $\mu g/m^3$ and SO_2 is reported in ppb.

3.3 Estimation Equation

To identify the impact of pollution on student cognitive functioning, we employ Kriging-based measures of pollution exposure at each school and estimate a linear regression via OLS of student standardized test scores onto pollution exposure while adjusting for school fixed effects, seasonality, and student demographic characteristics. Specifically, we estimate the model:

$$Y_{igset} = \alpha + \beta \widehat{P}_{st} + \gamma X_{ige} + \sigma_s + \mu_m + \theta_y + \upsilon_{igset} \tag{2}$$

where Y_{igset} is the standardized test score (i.e. the raw score minus its means divided by its standard deviation) of student i enrolled in grade g at school s taking exam e (math or English) on day t. Our main variable of interest, $\widehat{P}_{st}$, is the prediction of exposure to $PM_{2.5}$ or SO_2 at school s on day t discussed above. We scale all estimates of β up by 100 in order to make the estimates more readable by providing more significant figures. This and the fact that the dependent variable is a z-score implies that the interpretation of β is that a one unit increase in the pollutant increases test scores by β % of a standard deviation. The vector X_{ige} contains time-varying student characteristics such as indicators for economic disadvantage status and recipiency of English language services, time-invariant student characteristics such as indicators for gender and ethnicity, and an indicator for the student's grade and the type of exam.[16] We include school fixed effects, denoted by σ_s, in order to control for variation at the school level. The terms, μ_m and θ_y, are month and academic year fixed effects respectively and υ_{igset} is the error term. Standard errors are clustered by school. Identification of β in equation (2) comes from plausibly exogenous variation in $\widehat{P}_{st}$ *within* schools, across time.[17]

We conclude with a few remarks about the calculation of standard errors in the presence of a generated regressor ($\widehat{P}_{st}$ in our case). With a generated regressor, standard errors should account for sampling uncertainty in $\widehat{P}_{st}$. In a standard situation in which the regressor is generated from the same sample that is used in the second stage estimation, bootstrapping the generated regressor and then bootstrapping the second stage coefficient estimates provides a common solution.

However, two points make this solution less viable in our scenario. First, the Kriging procedure that is used to generate the regressor takes several hours on a fast machine. This implies that a single standard error with 100 replications could take a week to compute using standard bootstrapping procedures.[18] Second, as is common in the literature on the impacts of pollution, the generated regressor comes from a seperate sample with a separate sampling scheme than the primary estimation sample. Accordingly, the asymptotic distribution computed in Appendix

[16]Controlling for the economic status of each students' family is particularly important in this specification because the DOE grants geographic exceptions (GE) to students who wish to enroll in a school outside of their district of residence under several qualifying circumstances (e.g. parents are faculty at the receiving school, a program is offered at the receiving school but not at the student's district school, etc). Because of the GE, students who reside in areas with relatively lower average household incomes may attend schools in districts that have higher average household incomes and thus better education programs.

[17]As an additional robustness check, we replace all time-invariant student controls with student fixed effects, which account for all unobserved time-invariant differences across students (e.g. innate ability). Our preferred model avoids student fixed effects since the time frame of our study is limited to several years (2015 to 2018) which, therefore, creates some considerable power issues.

[18]We do note, however, that there are faster alternatives for extremum estimations that could be considered (Andrews, 2002).

6A of Wooldridge (2010) for linear models with generated regressors does not apply as these calculations presume a single sample.

4 Results

4.1 Balance Test

Though the Hawai‘i context likely provides exogenous variation for the identification of the effects of pollutants on learning outcomes, we can still test whether there are observable differences in student characteristics correlated with pollutant exposures. In Appendix Table A1, we regress both of our pollution measures on the full vector of student and exam characteristics. Overall, we find little to no evidence of correlations between observable characteristics and pollutant levels. Excluding the estimates of the grade fixed effects in the second part of the table, the only significant coefficient in either column is for receipt of English language services which appears to predict $PM_{2.5}$ levels. However, this is the only significant variable in either column. We do see that the grade indicators predict SO_2 levels but not $PM_{2.5}$ levels, but this is easily dealt with by the inclusion of grade fixed effects in the estimations. Finally, we note that an F-test that all of the covariates in the estimates are zero resoundingly fails to reject the null. All told, we suspect that this significant estimate in the table is the consequence of Type I error. Thus, we conclude that the variation in pollutants has no systematic relationships with observable confounders, and the detected statistical significance likely arises from Type I error.

4.2 Baseline Results

We report our first set of OLS estimations in Table 4. Column (1) presents our results for the effect of standardized $PM_{2.5}$ levels on student test scores, as estimated in equation (2). With moderate precision, we estimate (at the 10% level) a drop in student standardized test scores on days with higher $PM_{2.5}$ levels. This average effect is small; a one unit increase in $PM_{2.5}$

Table 4: Effect of Pollution on Math and Reading Scores, OLS Estimates

	(1)	(2)	(3)	(4)
$PM_{2.5}$	–0.132*	–0.076		
	(0.073)	(0.071)		
SO_2			–0.211	–0.224
			(0.174)	(0.168)
R^2	0.267	0.851	0.267	0.838
School FE	X	X	X	X
Month FE	X	X	X	X
Year FE	X	X	X	X
Individual FE		X		X

Notes: Standard errors are clustered by school. Control variables include gender, economically disadvantaged students, English language service recipients, exam subject, grade level and ethnicity. All estimations control for school, month and academic year fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

leads to a 0.13 percent of a standard deviation drop in student test scores (recall that our pollutant effects are scaled up by a factor of 100). With a full sample standard deviation of 1.84 (see Table 2), a one standard deviation increase in $PM_{2.5}$ corresponds with a 0.24 percent drop in student test scores, on average. In column (2), we replace student-invariant controls with student fixed effects and see that the estimate is no longer significant, but this is likely the result of low power. In the final two columns, we do not see any impact of SO_2 on test scores. However, we will show that these weaker effects mask important underlying heterogeneity by geography, the level of baseline pollution, and SES within schools.

4.3 Impacts by Geographical Region

In Table 5, we estimate our model stratified by three geographical regions: Oahu, south Hawaii, and north Hawaii. The important finding in this table is that we now see much larger impacts of both $PM_{2.5}$ and SO_2 on south Hawaii while we see no impacts elsewhere. The point estimates of the effects of this pollutants are -0.652 and -0.347 and are significant at the 1% and 5%, respectively. These estimates indicate that, on south Hawaii, a one standard deviation increase in $PM_{2.5}$ and SO_2 results in test score declines of 1.20 and 2.14 percent of a standard

Table 5: Effects of Pollutants on Exam Scores for Students by Region

	(1)	(2)	(3)	(4)	(5)	(6)
	Oahu		South/Southwest Hawaii		North/Northeast Hawaii	
$PM_{2.5}$	–0.070		–0.652***		0.167	
	(0.0927)		(0.219)		(0.180)	
SO_2		–0.623		–0.347**		0.206
		(1.94)		(0.145)		(0.347)
R^2	0.273	0.273	0.284	0.269	0.201	0.200

Notes: Standard errors are clustered by school. Control variables include gender, economically disadvantaged students, English language service recipients, exam subject, grade level and ethnicity. All estimations control for school, month and academic year fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

deviation.

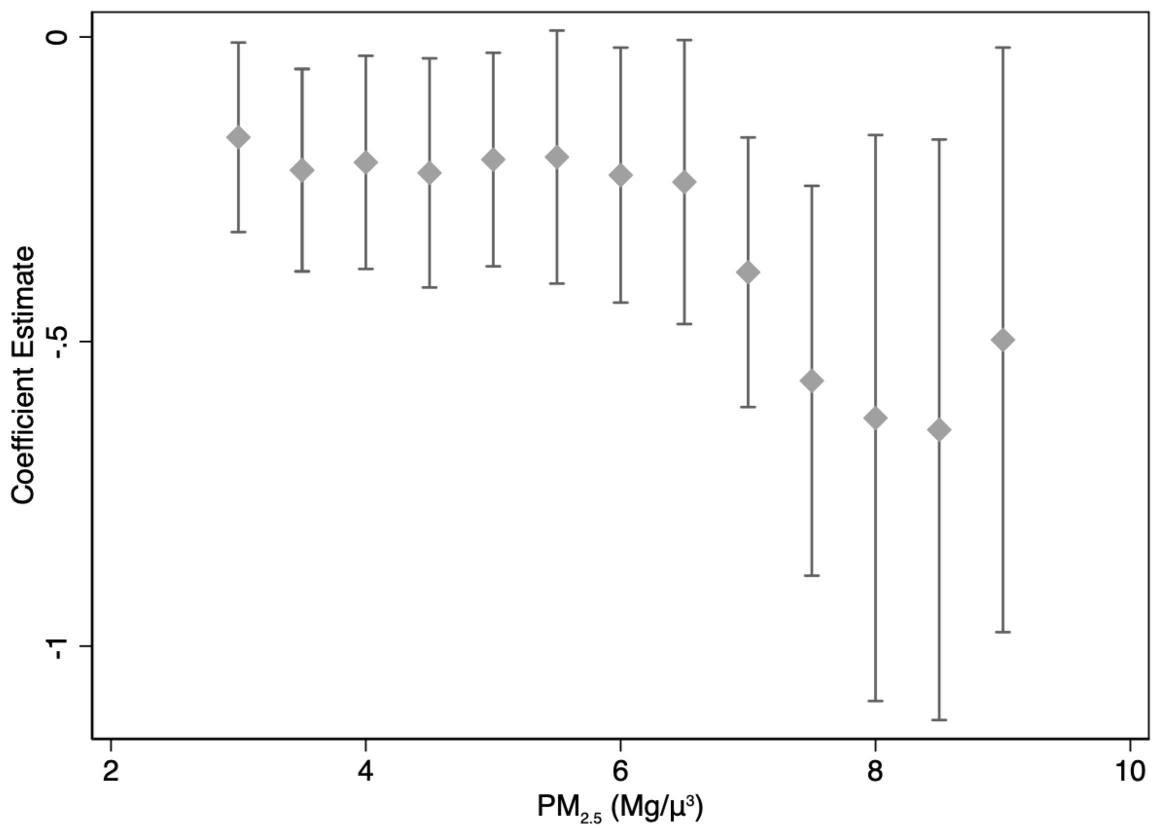
We note that pollution levels are substantially higher on south Hawaii than on Oahu or north Hawaii as shown in Table 3. For example, south Hawaii has substantially worse pollution than Oahu regardless of whether or not the trade winds are blowing. In addition, air quality in south Hawaii is notably worse than in north Hawaii on trade wind days but not on days in which they are not.

Thus, these effects might indicate that the effects of these pollutants are non-linear in their levels. Small exposure to either particulates or SO_2 appears to have no effects on Oahu and, to a lesser extent, north Hawaii. However, the effects on south Hawaii, where air quality is notably worse, are very large.

4.4 Impacts by baseline pollutant levels

Are the effects of $PM_{2.5}$ on student cognitive performance in fact larger when baseline pollution levels are higher? A key advantage to the Hawai‘i context (as indicated by Table 3) comes from its rich variation in baseline pollutant levels across schools. This is due to where schools are located on each island and their proximity to the Kilauea volcano. This allows us to test for possible nonlinear effects in how pollutants affect student performance, particularly across schools with lower baseline levels. In Figure 5 we report coefficients across different samples by school, where each subsample progressively focuses on schools with higher baseline $PM_{2.5}$ levels.[19]

Figure 5: Differential Effects by Baseline $PM_{2.5}$ Levels



Notes: The y-axis represents the coefficient for the effect of the pollutant, $PM2.5$ on student z-scores by the mean level of exposure to $PM_{2.5}$ within each school. The x-axis represents the threshold at which each school's mean exposure is greater than a given level of $PM_{2.5}$. Standard errors are clustered by school. Control variables include female, economically disadvantaged families, non-native English speaking, math exam, grade level and ethnicity. All estimations control for school, month and academic year fixed effects.

We find that as we focus on schools with higher baseline levels of pollution, the negative effects of pollutants sharply increase. When we restrict our regression sample to pollution levels between 3 to 6 $\mu g/m^3$ we see reductions in test scores in the neighborhood of 0.37-0.74 percent of a standard deviation with respect to a one standard deviation increase in $PM_{2.5}$. However,

[19] Baseline $PM_{2.5}$ levels are calculated by taking the average $PM_{2.5}$ level across the full sample of days for each school.

when baseline exposure is greater than 6 $\mu g/m^3$, there is a precipitous drop to reductions of in the neighborhood of 1.1 percent of a standard deviation (also with respect to a one standard deviation increase in $PM_{2.5}$). Accordingly, the pernicious effects of $PM_{2.5}$ are largest when exposure is the greatest. This result is very much consistent with the results in Table 5 which show that the effects of pollutants are largest near Kilauea.

4.5 Heterogeneity by economic disadvantage status

Does poor air quality have larger effects on the most disadvantaged pupils within a school? Case et al. (2002) show that children from poorer backgrounds are at higher risk of developing a host of health problems than better off children. This suggests that more well off children will be in better health which could confer more resiliency when combating the pernicious effects of air pollution. In this sense, air pollution could exacerbate pre-existing inequities within schools.

To investigate this, in Table 6, we estimate equation (2) while including an interaction term between the pollutant level and an indicator for whether the pupil was eligible for the free school lunch program - a proxy for student economic disadvantage. In the first column of the table, we display the estimates of the effects of particulates and we observe drastically different effects by economic status. The interaction between $PM_{2.5}$ and the disadvantaged indicator is -0.572 and significant at the 1% level whereas the direct effect is -0.044 but not significant. This implies that the the harmful effects from $PM_{2.5}$ for disadvantaged students are over ten times the magnitude of their effects for their more advantaged counterparts.[20] In column (2), we also find that SO_2 harms disadvantaged students substantially more than students who are better off. A similar calculation indicates that the effects of SO_2 on poorer pupils are larger by sixfold.

The magnitudes of these effects on cognitive performance of disadvantaged students pupils are not trivial. Once again, using the descriptive statistics from Table 2, we calculate that a one standard deviation increase in $PM_{2.5}$ decreases test scores for disadvantaged pupils by 1.13%

[20] We compared the sum of the interaction and the direct effect of particulates to the direct effect by itself.

Table 6: Effect of Pollution on Exam Scores for Economically Disadvantaged Students

	(1)	(2)
Economically Disadvantaged * $PM_{2.5}$	–0.572***	
	(0.182)	
$PM_{2.5}$	–0.044	
	(0.083)	
Economically Disadvantaged * SO_2		–0.426***
		(0.153)
SO_2		–0.084
		(0.182)
Economically Disadvantaged	–27.519***	–29.210***
	(1.255)	(1.113)
R^2	0.267	0.264

Notes: Standard errors are clustered by school. Control variables include gender, English language service recipients, exam subject, grade level and ethnicity. All estimations control for school, month and academic year fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

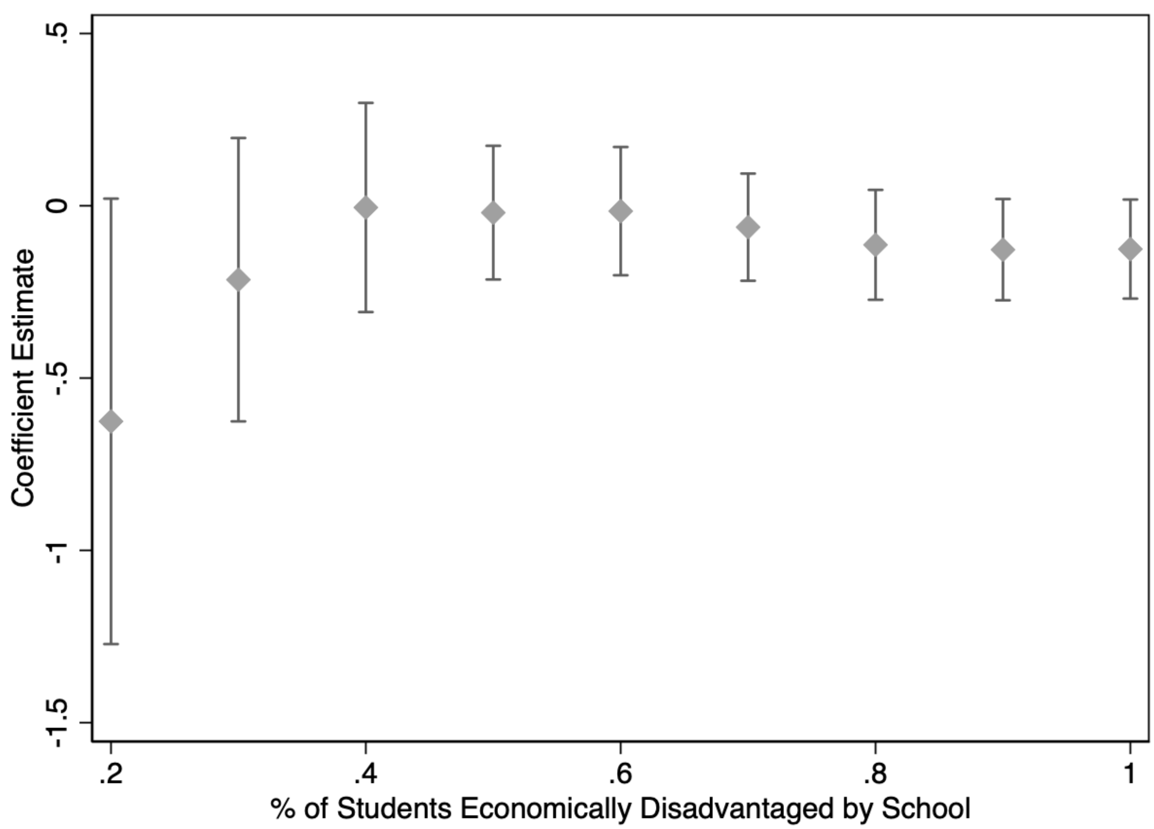
of a standard deviation. A similar calculation indicates that a one standard deviation increase in SO_2 decreases test scores for poorer pupils by 3.15% of a standard deviation.

Could these larger impacts of pollution for disadvantaged students be driven by selection? For example, there could be a potential correlation between where disadvantaged students enroll and school characteristics including the school's location, its baseline pollution level, or the school's potential resources to combat the harmful effects of pollutants (e.g. air conditioning). We do not believe that this is the case for the simple reason that that estimates in Table 6 all include school fixed effects.

However, we can also offer an alternative test of this possibility to eliminate any lingering doubts. In Figure 6 we estimate equation (2) while focusing on subsamples of schools by the fraction of the school's students who were economically disadvantaged. The estimates in the far left of the figure correspond to schools with 20% or fewer pupils on the school lunch program whereas the far right of the figure includes all schools. As you move from the left to the right of the figure, the sample of schools becomes more disadvantaged.

Interestingly, we find little difference in how pollutants harm student learning by the school's fraction of economically disadvantaged students. This suggests that the observed disparity by student disadvantage status arises from differences across students within each school, and not due to differences across schools. In other words, disadvantaged students do not appear to be especially harmed by pollutants due to their school's location or school resources. Rather, it appears as if poorer pupils are more adversely impacted by pollution than richer pupils.

Figure 6: Differential Effects by % of Economically Disadvantaged Students Within Each School



Notes: The y-axis represents the coefficient for the effect of the pollutant, $PM2.5$ on student z-scores by the percentage of students who are economically disadvantaged within each school. The x-axis represents the threshold at which each school has less than a given percentage of economically disadvantaged students. Standard errors are clustered by school. Control variables include female, economically disadvantaged families, non-native English speaking, math exam, grade level and ethnicity. All estimations control for school, month and academic year fixed effects.

5 Conclusion

Using variation in air quality in the Hawaiian islands due to volcanic activity, we estimate the impacts of $PM_{2.5}$ and SO_2 on student performance. Because of the state's normally pristine air quality conditions, variation in pollutants are primarily determined by wind direction and volcanic emissions from Kilauea volcano on the island of Hawai'i. Exploiting this variation in pollution across the state, we find that worsening air quality decreases student exam scores. Specifically, we find that a standard deviation increase in $PM_{2.5}$ decreases test scores by about 0.24% of a standard deviation on average.

We also find that schools with higher baseline pollution levels tend to see even worse effects of air quality on student performance. Specifically, when we focus on schools with pollution levels between 3 to 6 $\mu g/m^3$, we see reductions in test scores of about 0.37-0.74 percent of a standard deviation with respect to a one standard deviation increase in $PM_{2.5}$. However, schools with average $PM_{2.5}$ exposure above 6 $\mu g/m^3$ see reductions in test scores of 1.1 percent of a standard deviation at the same margin. We also observe these nonlinear effects when estimating our main specification by geographic region. The negative, statistically significant effects of pollution on test scores are concentrated within south Hawaii, which has notably higher levels of pollution than in other areas across the state.

Finally, the negative effects of pollution on student performance are much larger for poorer students. Pupils who are economically disadvantaged experience ten times the effect of $PM_{2.5}$ and six times the effect of SO_2 on exam scores when compared to their more advantaged peers. These effects are not driven by school level characteristics but are instead a result of student level differences within schools. This is in line with previous literature which shows that poorer children are subject to worse health outcomes (Case et al., 2002) which may imply greater susceptibility to environmental insults.

All told, the findings from our study have implications for environmental justice. We show that poor students face additional obstacles accumulating human capital when air quality is poor relative to those who are more financially stable. This suggests that air pollution contributes to the strong persistence in socioeconomic status across generations that we observe in the United States.

References

Aguilar-Gomez, S., H. Dwyer, J. S. G. Zivin, and M. J. Neidell (2022): "This is Air: The" Non-Health" Effects of Air Pollution," Tech. rep., National Bureau of Economic Research.

Andrews, D. W. (2002): "Higher-order improvements of a computationally attractive k-step bootstrap for extremum estimators," *Econometrica*, 70, 119–162.

Archsmith, J., A. Heyes, and S. Saberian (2018): "Air quality and error quantity: Pollution and performance in a high-skilled, quality-focused occupation," *Journal of the Association of Environmental and Resource Economists*, 5, 827–863.

Bedi, A. S., M. Y. Nakaguma, B. J. Restrepo, and M. Rieger (2021): "Particle pollution and cognition: Evidence from sensitive cognitive tests in Brazil," *Journal of the Association of Environmental and Resource Economists*, 8, 443–474.

Burnett, R. T., M. Smith-Doiron, D. Stieb, S. Cakmak, and J. R. Brook (1999): "Effects of particulate and gaseous air pollution on cardiorespiratory hospitalizations," *Archives of Environmental Health: An International Journal*, 54, 130–139.

Carneiro, J., M. A. Cole, and E. Strobl (2021): "The effects of air pollution on students' cognitive performance: Evidence from brazilian university entrance tests," *Journal of the Association of Environmental and Resource Economists*, 8, 1051–1077.

Carré, J., N. Gatimel, J. Moreau, J. Parinaud, and R. Léandri (2017): "Does air pollution play a role in infertility?: a systematic review," *Environmental Health*, 16, 1–16.

Case, A., D. Lubotsky, and C. Paxson (2002): "Economic status and health in childhood: The origins of the gradient," *American Economic Review*, 92, 1308–1334.

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

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

Chen, S., C. Guo, and X. Huang (2018): "Air pollution, student health, and school absences: Evidence from China," *Journal of Environmental Economics and Management*, 92, 465–497.

Choumert-Nkolo, J., A. Lamour, and P. Phélinas (2021): "The economics of volcanoes," *Economics of Disasters and Climate Change*, 1–23.

Cressie, N. (1990): "The origins of kriging," *Mathematical geology*, 22, 239–252.

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

Currie, J., J. Voorheis, and R. Walker (2020): "What caused racial disparities in particulate exposure to fall? New evidence from the Clean Air Act and satellite-based measures of air quality," Tech. rep., National Bureau of Economic Research.

Di, Q., Y. Wang, A. Zanobetti, Y. Wang, P. Koutrakis, C. Choirat, F. Dominici, and J. D. Schwartz (2017): "Air pollution and mortality in the Medicare population," *New England Journal of Medicine*, 376, 2513–2522.

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

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

Halliday, T. J., J. Lynham, and Á. de Paula (2019): "Vog: using volcanic eruptions to estimate the health costs of particulates," *The Economic Journal*, 129, 1782–1816.

He, J., H. Liu, and A. Salvo (2019): "Severe air pollution and labor productivity: Evidence from industrial towns in China," *American Economic Journal: Applied Economics*, 11, 173–201.

Heissel, J. A., C. Persico, and D. Simon (2020): "Does Pollution Drive Achievement? The Effect of Traffic Pollution on Academic Performance," *Journal of Human Resources*, 1218–9903R2.

Koken, P. J., W. T. Piver, F. Ye, A. Elixhauser, L. M. Olsen, and C. J. Portier (2003): "Temperature, air pollution, and hospitalization for cardiovascular diseases among elderly people in Denver." *Environmental health perspectives*, 111, 1312–1317.

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

Linares, B., J. M. Guizar, N. Amador, A. Garcia, V. Miranda, J. R. Perez, and R. Chapela (2010): "Impact of air pollution on pulmonary function and respiratory symptoms in children. Longitudinal repeated-measures study," *BMC Pulmonary Medicine*, 10, 1–9.

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

Marcotte, D. E. (2017): "Something in the air? Air quality and children's educational outcomes," *Economics of Education Review*, 56, 141–151.

Montero, J.-M., G. Fernández-Avilés, and J. Mateu (2015): *Spatial and spatio-temporal geostatistical modeling and kriging*, vol. 998, John Wiley & Sons.

Nieuwenhuijsen, M. J., X. Basagaña, P. Dadvand, D. Martinez, M. Cirach, R. Beelen, and B. Jacquemin (2014): "Air pollution and human fertility rates," *Environment international*, 70, 9–14.

Perin, P. M., M. Maluf, C. E. Czeresnia, D. A. N. F. Januário, and P. H. N. Saldiva (2010): "Impact of short-term preconceptional exposure to particulate air pollution on treatment outcome in couples undergoing in vitro fertilization and embryo transfer (IVF/ET)," *Journal of assisted reproduction and genetics*, 27, 371–382.

Slama, R., S. Bottagisi, I. Solansky, J. Lepeule, L. Giorgis-Allemand, and R. Sram (2013): "Short-term impact of atmospheric pollution on fecundability," *Epidemiology*, 871–879.

Wooldridge, J. M. (2010): *Econometric analysis of cross section and panel data*, MIT press.

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

A.1 Technical Details of the Kriging Procedure

To fix ideas, we let $s \in \{1, ..., S\}$ denote the school, and $t \in \{1, ..., T\}$ denote the time periods, and $m \in N(s)$ denote the monitoring station where $N(s)$ is the neighborhood of school s. We consider three neighborhoods: the island of Oahu, the southwestern part of Hawaii (that is most exposed to Kilauea's emissions), and the remainder of Hawaii. These are depicted in Figures 1 and 2. We denote the pollution measurement at a given monitoring station on a particular day as Π_{mt}. The predicted exposure is then

$$\widehat{P}_{st} = \sum_{m \in N(s)} \lambda_{sm}(NE_t)\Pi_{mt}$$

where the kriging weights are $\lambda_{sm}(NE_t) \equiv \lambda(d_{sm}, l_{sm}, NE_t)$. Once again, bear in mind that $NE_t \in \{0, 1\}$.

The weights are chosen to guarantee that the predictions are unbiased and that the prediction error has minimum variance. Unbiasedness requires that the weights sum to unity. If we let Π_{st} represent the true pollution measurement at school s, both criteria can formally be written as

$$\min_{\{\lambda_{sm}(b)\}_{m \in N(s)}} V\left(\widehat{P}_{st} - \Pi_{st}\right)$$

$$\text{subject to} \sum_{m \in N(s)} \lambda_{sm}(b) = 1$$

This minimization problem is solved twice: once for trade wind days ($b = 1$) and once for non-trade wind days ($b = 0$). This delivers two sets of weights which depend on the prevailing winds for that day.

Following Montero et al. (2015) (see p. 86) and making some local stationarity assumptions, the first order conditions that guarantee these criteria are

$$\sum_{m\in N(s)} \lambda_{sm}(b)\gamma_{nm}(b) + \alpha_b = \gamma_{ns}(d_b) \text{ for } n \in N(s) \tag{A.1}$$

$$\sum_{m\in N(s)} \lambda_{sm}(b) = 1 \tag{A.2}$$

where both conditions hold for $b \in \{0, 1\}$ and α_b is the Lagrangian multiplier on the constraint in A.2 which guarantees the unbiasedness of the prediction. The object, $\gamma_{nm}(b)$, is the semi-variogram between locations m and n when $NE_t = b$. For each school in $N(s)$, equations A.1 and A.2 constitute a set of $\#N(s) + 1$ equations in as many unknowns. So, we then have $\#N(s) \times (\#N(s) + 1)$ equations in total. If we index the monitoring stations in $N(s)$ from one to $M(s) \equiv \#N(s)$ (with some abuse of notation) and define

$$\lambda_s(b) \equiv (\lambda_{s1}(b), ..., \lambda_{sM(s)}(b), \alpha_b)'$$

then the Kriging weights are $\lambda_s(b) = \Gamma(b)^{-1}\Gamma_s(b)$ where

$$\Gamma(b) \equiv \begin{bmatrix} \gamma_{11}(b) & \dots & \gamma_{1M(s)}(b) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ \gamma_{M(s)1}(b) & \dots & \gamma_{M(s)M(s)}(b) & 1 \\ 1 & \dots & 1 & 0 \end{bmatrix}$$

and

$$\Gamma_s(b) \equiv \begin{bmatrix} \gamma_{1s}(b) \\ \vdots \\ \gamma_{M(s)s}(b) \\ 1 \end{bmatrix}.$$

Hence, the task of computing the Kriging weights is reduced to computing the semivariogram for $NE_t = 1$ and $NE_t = 0$.

To compute the semivariogram, we postulate a functional form for the semivariogram. We assume that

$$\gamma_{mn}(NE_t) = 1 - \exp\left(-d_{mn} \times \left(\phi NE_t + \sum_{j\in\mathcal{L}}(\delta_j 1_{mn}(j) + \beta_j \times 1_{mn}(j) \times NE_t)\right)\right) \quad \text{(A.3)}$$

where $\mathcal{L} \equiv \{NE, SE, SW, NW\}$ (which collects the relative location variables) and $1_{mn}(j)$ is an indicator for the location of n relative to m where m is held fixed. Note that distance enters multiplicatively to ensure that the semivariogram is zero when $d_{mn} = 0$ and, by construction, $\gamma_{mn}(NE_t) \in [0, 1]$. This assumption reduces the spatial covariance structure to a smaller number of parameters which allows us to make extrapolations and interpolations needed to construct $\Gamma_s(b)$. We collect these parameters in the vector θ.

We estimate these parameters using GMM. We let $\widetilde{\gamma_{mn}}(b)$ denote the *empirical* semivariogram for the monitoring station pair (m, n) for $b \in \{0, 1\}$. Similarly, $\widetilde{\Gamma(b)}$ is the matrix that collects the empirical semivariograms. Then for each pair (m, n), we can compare two collections of moment conditions

$$q_b(\theta) = \text{vec}\left(\text{lower triangle}\left(\widetilde{\Gamma(b)} - \Gamma(b;\theta)\right)\right)$$

for $b \in \{0, 1\}$. We further collect these in the $(M(s)+1)\times M(s)\times 2$ vector $q(\theta) \equiv (q_0(\theta)', q_1(\theta)')'$. We estimate θ by minimizing $q(\theta)'q(\theta)$.

Importantly, θ is easy to estimate. First, compute the empirical semivariagrams for trade wind ($NE_t = 1$) and non-trade wind days ($NE_t = 0$). Second, we re-write A.3 as

$$1 - \gamma_{mn}(NE_t) = \exp\left(-d_{mn} \times \left(\phi NE_t + \sum_{j\in\mathcal{L}}(\delta_j 1_{mn}(j) + \beta_j \times 1_{mn}(j) \times NE_t)\right)\right)$$

and we then note that estimation can proceed by using a simple Poisson regression package in

Stata or R. Once θ is estimated, we can then estimate the Kriging weights $\lambda_s(b)$.

A.2 Additional Tables and Figures

Table A1: Balance Test

	(1) $PM_{2.5}$	(2) SO_2
Female	-0.016	-0.003
	(0.013)	(0.008)
Economically Disadvantaged	-0.035	0.007
	(0.031)	(0.015)
Received English Language Services	0.097**	0.005
	(0.048)	(0.021)
Math Exam	-0.184	0.082
	(0.224)	(0.096)
Filipino	0.001	-0.010
	(0.024)	(0.009)
Native Hawaiian	0.023	0.017
	(0.030)	(0.017)
Pacfic Islander	-0.024	-0.017
	(0.032)	(0.016)
White	0.007	-0.007
	(0.023)	(0.012)
Other Ethnicity	-0.024	-0.011
	(0.030)	(0.016)

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors clustered by school. Regressors include female, economically disadvantaged families, non-native English speaking, math exam, grade level and ethnicity. All estimations control for school, month and academic year fixed effects. The F-test tests for whether the covariates in each model are jointly equal to zero.

Table A1: Balance Test Continued

	(1)	(2)
	$PM_{2.5}$	SO_2
4th Grade	0.161	0.012
	(0.183)	(0.050)
5th Grade	-0.038	-0.062
	(0.134)	(0.068)
6th Grade	0.086	-0.198***
	(0.189)	(0.062)
7th Grade	-0.217	-0.245***
	(0.221)	(0.088)
8th Grade	0.016	-0.317***
	(0.219)	(0.095)
11th Grade	-0.160	-0.738
	(0.312)	(0.640)
Month FE	X	X
Academic Year FE	X	X
School FE	X	X
R^2	0.356	0.504
F-test	1.337	1.186
p-value	0.432	0.742

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: Standard errors clustered by school. Regressors include female, economically disadvantaged families, non-native English speaking, math exam, grade level and ethnicity. All estimations control for school, month and academic year fixed effects. The F-test tests for whether the covariates in each model are jointly equal to zero.