Discussion Paper Series

IZA DP No. 18904 August 2026

Air Pollution and Learning

A. Patrick Behrer
World Bank

Joshua Goodman
Boston University, NBER
and IZA@LISER

J. Parker Goyer
College Board

R. Jisung Park
University of Pennsylvania
and IZA@LISER

IZA@LISER NETWORK
ISSN: 2365-9793

IZA DP No. 18904 August 2026

Air Pollution and Learning*

Abstract

Nearly the entire world's population breathes air exceeding WHO pollution guidelines, but the extent to which that exposure impairs the accumulation of human capital is not well understood. We study this using longitudinal PSAT data on nearly 10 million U.S. high school students, comparing the same student's scores across attempts preceded by differing air quality and instrumenting for local PM2.5 with smoke from distant wildfires. A year of observed pollution exposure reduces learning by 0.04–0.06 standard deviations, or 14–19% of typical annual score growth. The damage comes almost entirely from moderate pollution days (8–12 g/m3), below the EPA's historical standard, and from exposure during the school year rather than summer, pointing to instructional disruption as a mechanism. Effects are three times larger in disadvantaged schools and among Black and Hispanic students, who are harmed more by the same exposure. Exposure to air pollution widens achievement gaps.

JEL classification

I2, I24, Q5, Q53

Keywords

air pollution, PM2.5, learning, student achievement, wildfire smoke

Corresponding author

Joshua Goodman
edecjosh@bu.edu

* Funding from the Russell Sage Foundation is gratefully acknowledged. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

1 Introduction

How does air pollution affect human capital accumulation? Answering this question is of first-order importance for several reasons. First, the scale of exposure is enormous: 99 percent of the world's population live in places that exceed the WHO's recommended levels of ambient particulate matter 2.5 microns or less in diameter ($PM_{2.5}$), particles so small that they enter deep in the lungs and pass to the bloodstream, giving them a wide physiological reach. Second, exposure is highly uneven, with poorer individuals far more likely to live in polluted areas globally (Rentschler and Leonova, 2023), and more vulnerable communities systematically more exposed within wealthy countries (Colmer et al., 2020; Jbaily et al., 2022). Third, because local air pollution is a classic negative externality, understanding its full social cost is critical to designing efficient policy, whether aimed at reducing emissions or mitigating pollution's effects. Finally, human capital accumulation is a pivotal input to economic growth (Gethin, 2025) and individual economic mobility (Heckman and Mosso, 2014), so any effect of pollution on learning carries consequences that compound over time.

This paper explores how sustained exposure to air pollution over the course of a school year affects the rate at which students learn. Our focus is on the cumulative rate of learning: how ongoing exposure to poor air quality affects the total human capital a student accumulates by the time they sit for important exams, pursue higher education, or enter the labor force. This is conceptually distinct from asking whether pollution impairs performance on a single exam day, a question the existing literature has addressed. We instead ask whether sustained exposure to polluted air slows the rate at which students actually learn, a question that has proven difficult to answer given data limitations.

Our outcome data come from the College Board and comprise the universe of high school students who took the PSAT exam at least twice between 2007 and 2016. This student-level longitudinal data contains nearly 19 million exam scores from nearly 10 million students attending nearly 25,000 U.S. high schools. The focus on repeat test-takers is central to our identification strategy. Because a large share of U.S. high school students take the PSAT more than once in grades 9 through 11, we can compare each student's performance across attempts that differ in the air quality experienced in the preceding year, holding fixed student-level factors like ability, household background, and school quality.

This within-student design has three advantages over much of the recent literature on pollution and human capital, where outcomes are typically measured at the school or district level, or students are observed only once. First, it avoids cross-sectional comparisons across schools or districts, which are vulnerable to bias from geographic differences in unobserved determinants of pollution and achievement. Second, it avoids the compositional bias from pollution itself changing who attends a given school or district. Third, our sample is substantially larger than any prior study of this question — more than sixteen times larger than the next largest examination of cumulative pollution effects on learning — giving us the statistical power to detect effects that may be modest on a per-day basis but consequential in aggregate.

We measure air pollution using daily $PM_{2.5}$ concentrations predicted by the machine learning model of Di et al. (2019), which combines ground-station monitoring, satellite imagery, chemical transport models, and weather reanalysis to estimate $PM_{2.5}$ across the contiguous United States at a 1km-by-1km resolution each day from 2000 to 2017. For each school, we average pollution across all grid cells within 10 kilometers of its location, using the school's ZIP code when precise coordinates are unavailable. We characterize each student's exposure using the number of days in the 365 days prior to their exam that fall into three ranges: non-polluted days ($PM_{2.5}$ below 8 $\mu g/m^3$), moderately polluted days (8–12 $\mu g/m^3$), and heavily polluted days (at or above 12 $\mu g/m^3$), and we further divide these into school-year and summer days using state-specific academic calendars.

Even conditional on student and school, local air quality in a given year is not randomly assigned. Our within-student design addresses selection on time-invariant student and school characteristics, but not the possibility that year-to-year changes in local pollution are driven by changes in local economic conditions, such as a new factory or a local labor demand shock, that independently affect learning. To isolate plausibly exogenous variation in pollution, we instrument for local $PM_{2.5}$ using data on wildfire smoke from distant fires from Childs et al. (2022). Smoke plumes can travel hundreds or thousands of miles, raising local $PM_{2.5}$ in ways that are largely orthogonal to local economic conditions, school operations, and household behavior (Zhang et al., 2025). Restricting to smoke originating at least 10 kilometers from a student's school isolates variation in local pollution driven by distant weather and fire activity rather than by anything happening in the local community in response to a fire. The instrument serves a second purpose as well: because our pollution measures are themselves estimates from a machine learning model, they are subject to measurement error that would otherwise attenuate our OLS estimates toward zero. Smoke-derived $PM_{2.5}$ is subject to independent measurement error uncorrelated with the error in total $PM_{2.5}$, so instrumenting purges this attenuation bias and recovers estimates closer to the true causal effect.

Pollution exposure meaningfully slows the rate at which students learn. IV estimates are substantially larger than OLS estimates, consistent with economic confounds or attenuation bias from measurement error. Each additional day of pollution exposure in the year prior to an exam reduces PSAT performance by roughly 0.21 thousandths of a standard deviation. Scaling our estimates by students' observed exposure, pollution reduces annual learning by roughly 0.055 standard deviations relative to a pollution-free year, equivalent to 19 percent of typical yearly score growth. To put this magnitude in perspective, it is comparable to the learning loss associated with approximately 5 to 10 additional days of student absence over a school year (Goodman, 2014; Aucejo and Romano, 2016), or reducing teacher value added by roughly a third of a standard deviation (Chetty et al., 2014). It is roughly one-fourth the observed test score impact that Biasi et al. (2025) find when schools install new HVAC systems, which provide a combination of temperature control and air quality control through improved ventilation. Our estimates suggest that a meaningful fraction of such infrastructure effects work through reducing exposure to air pollution.

The impact of pollution on learning is driven almost entirely by moderate pollution rather than by more extreme exposure. Harmful effects emerge at concentrations below the EPA's historical 12 $\mu g/m^3$ standard, indicating meaningful social costs accrue within levels of pollution regulators have historically treated as safe. Consistent with mechanisms that include reduced or impaired instructional time (Currie et al., 2009; Park et al., 2020), effects are also driven more by pollution experienced during the school year rather than summer, pointing toward cumulative short-run cognitive interference rather than longer-run health channels as the likely mechanism.

This average effect masks substantial heterogeneity across schools and student groups. Pollution's harm to learning is roughly three times larger in socioeconomically disadvantaged schools than in less disadvantaged ones. Pollution's harm is similarly concentrated among Black and Hispanic students, for whom the estimated effect is large and highly significant, while the corresponding effect for White and Asian students is statistically indistinguishable from zero. Both gaps are driven almost entirely by differential sensitivity to pollution rather than differential exposure. Exposure itself varies only modestly across school-challenge and racial groups, but the same exposure does far more damage to the more socially disadvantaged group in each case. One plausible contributor to this pattern is differential adaptation. Evidence from wildfire-smoke avoidance behavior shows that higher socioeconomic status individuals respond to smoke far more aggressively than others (Berestycki and Chen, 2026), and schools in lower income neighborhoods are less likely to have well-functioning HVAC systems (Park et al., 2020) which can help improve indoor air quality, and lower quality capital stock overall (Biasi et al., 2025), suggesting that unequal capacity to avoid exposure may help explain why the same polluted day does more harm in more disadvantaged schools. Our estimates suggest that the total learning cost of pollution accounts for over 10 percent of the achievement gaps between high- and low-challenge schools and between Black/Hispanic and White/Asian students.

Our work contributes to three related literatures. First, we build on a growing literature estimating the effects of environmental factors, such as air pollution, on student outcomes. High-stakes cognitive performance can be affected by exam day factors including temperature (Graff-Zivin and Shrader, 2016; Park, 2022) and air pollution (Ebenstein et al., 2016; Carneiro et al., 2021; Graff Zivin et al., 2020; Andersen et al., 2025).[1] Our focus on accumulated pollution exposure over the year prior to exams distinguishes genuine learning losses from test-day performance effects. Other papers use school district-level performance data to show that school year exposure to air pollution from wildfire smoke or coal plants reduces district level test scores (Wen and Burke, 2022; Gilraine and Zheng, 2022). Our student-level data and focus on retakers account for the endogenous compositional shifts pollution can induce. Student-level evidence from Florida links highway and industrial pollution to learning (Heissel et al., 2022; Persico and Venator, 2021). Both papers rely on binary exposure measures from a single state, which we complement with national estimates from a continuous measure.[2] Our sample is more than sixteen times larger than the largest of these studies

[1] Contemporaneous pollution can also impair cognitive performance on demanding tasks in non-academic settings (Archsmith et al., 2018; Huang et al., 2020; Dong et al., 2021). For a comprehensive review of pollution's cognitive and broader "nonhealth" effects, see Aguilar-Gomez et al. (2022).

[2] International evidence from Colombia and Vietnam also suggests learning impacts from air pollution exposure in those contexts (Martin et al., 2025; Dang et al., 2025).

and our instrument is based on a transparent, well-validated source of exogenous variation that accounts for potential measurement error in ways prior work often does not.[3] The learning losses we document may be partly explained by air pollution's impact on student absenteeism (Komisarow and Pakhtigian, 2022; Chen et al., 2018; Aziz and Elbakidze, 2025), particularly for students with asthma (Fitzpatrick et al., 2026), and may partly explain why school infrastructure projects focused on HVAC improvements have large impacts on test scores (Biasi et al., 2025).

Second, our results also speak to a growing literature assessing the causes and consequences of educational inequality: specifically, a longstanding literature assessing the factors that give rise to differences in educational achievement across socioeconomic factors such as race, gender, and income (Coleman et al., 1966; Reardon et al., 2019). Academic achievement gaps have remained persistently large both across and within countries and have been associated with substantial differences in a range of adult outcomes (Heckman et al., 2013; Chetty et al., 2020). While many potential determinants of such achievement gaps have been studied, including but not limited to teacher quality (Chetty et al., 2011), parental investments (Cunha et al., 2010; Todd and Wolpin, 2007; Guryan et al., 2008), school funding (Jackson et al., 2015; Lafortune et al., 2018), crime (Ang, 2021) and extreme temperature (Park et al., 2020), none to our knowledge have rigorously assessed the role of local air pollution.[4] Our estimate of pollution's contribution to US racial achievement gaps, on the order of 10 percent of the gap between Black/Hispanic and White/Asian students, is comparable in magnitude to the contribution of extreme temperature (3 to 7 percent) estimated using similar data (Park et al., 2020).

Finally, our paper contributes to the literature assessing the causal influence of environmental externalities on non-market outcomes, including but not limited to the effect of air pollution on health (Greenstone, 2002), the climate externality on mortality, morbidity, and learning (Barreca et al., 2016; Carleton et al., 2022; Dillender, 2021; Park et al., 2021), and the effects of water pollution on health and recreational value (Shapiro et al., 2019). This literature has increasingly emphasized that such externalities are not evenly borne. Air pollution exposure in the United States is highly

[3]See Table A1 for a list of comparison studies.

[4]A separate but related literature shows that exposure to pollution in-utero (Bharadwaj et al., 2017; Carneiro et al., 2024) or early in childhood (Voorheis, 2017) has adverse effects on learning later in life, with effects that may persist across generations (Colmer and Voorheis, 2020). These papers are unable to assess the effects of pollution on the schooling process.

unequally distributed, with lower-income communities experiencing higher levels of pollution of many kinds (Currie, 2011; Colmer et al., 2020). These disparities persist: the most polluted census tracts in 1981 remained the most polluted in 2016, even as average $PM_{2.5}$ concentrations fell by roughly 70 percent (Colmer et al., 2020). This inequality in exposure is compounded by inequality in beliefs about it: recent survey evidence finds that Americans systematically misperceive their own air pollution exposure and its health consequences, that these misperceptions differ by race, and that correcting beliefs alone would do little to close racial gaps in exposure or in willingness to pay for clean air (Tarduno and Walker, 2025). Our results extend this pattern from health to human capital: the same unequal exposure that generates disparate health burdens (Deryugina et al., 2019; Bishop et al., 2023) also generates disparate learning losses. We show that disadvantaged students are not merely more exposed to pollution but more harmed by any given level of exposure, implying that disparities in exposure understate disparities in harm.

2 Mechanisms and Policy Context

2.1 Why Air Pollution Might Affect Learning

$PM_{2.5}$ refers to particulate matter smaller than 2.5 microns in diameter, roughly a thirtieth the width of a human hair. Its size makes $PM_{2.5}$ of more concern than coarser particles (PM_{10}) or gaseous pollutants like ozone or NO_2. Particles this small bypass the nose and upper airway's natural filtering, penetrate deep into the lungs' alveoli, and in some cases cross into the bloodstream, giving $PM_{2.5}$ a physiological reach well beyond the respiratory system alone. Common sources include vehicle exhaust, industrial emissions, power generation, and, increasingly, as wildfire activity has grown more frequent, smoke.

There are several distinct biological and behavioral channels through which $PM_{2.5}$ exposure could plausibly affect learning:

Acute cognitive interference. A body of evidence documents that pollution impairs performance on demanding cognitive tasks — including exams — on the day of exposure itself, plausibly through short-run effects on attention, working memory, and processing speed (Aguilar-Gomez et al., 2022). The proposed physiological pathway involves oxidative stress and low-grade inflammation triggered by inhaled particulates, which can affect neural function in the short run even absent any lasting damage. This can affect both short-run performance on exams and the rate of daily learning during the period leading up to exams.

Respiratory and physical health. $PM_{2.5}$ is a well-documented trigger of asthma exacerbation and other respiratory symptoms, particularly in children, whose lungs are still developing and who breathe more air relative to body size than adults. Beyond asthma, elevated $PM_{2.5}$ has been linked to a range of general respiratory and cardiovascular strain that can leave students fatigued or unwell. This channel can affect learning both directly, by reducing a student's capacity to engage with instruction while present in class, and indirectly, by driving absences.

Disruption to school operations. Independent of any individual student's health, schools sometimes respond to elevated pollution by canceling outdoor recess and athletics, restricting time outdoors more generally, or in more severe cases (particularly during heavy wildfire smoke events) closing entirely. These operational responses reduce instructional time for an entire student body at once rather than affecting students individually.

2.2 US and International Policy Context

Air pollution has been federally regulated in the United States since the Clean Air Act, which directs the Environmental Protection Agency to set National Ambient Air Quality Standards (NAAQS) for $PM_{2.5}$ and to review them at least every five years. The primary (health-based) annual $PM_{2.5}$ standard stood at 12 $\mu g/m^3$ from 2012 until February 2024, when the EPA tightened it to 9 $\mu g/m^3$, citing evidence of health harm at lower concentrations than the prior standard reflected. That revision was challenged in court by industry groups and several states, and in March 2025 the EPA itself, under new leadership, announced it would reconsider the rule and asked the D.C. Circuit to vacate it. In June 2026, the D.C. Circuit upheld the 9 $\mu g/m^3$ standard, rejecting the challenges brought by EPA, states, and industry groups. Though further appeal remains possible, the legally binding annual standard as of 2026 is 9 $\mu g/m^3$.

U.S. standards are, by global comparison, relatively permissive. The World Health Organization's Global Air Quality Guidelines, last updated in September 2021, recommend an annual mean $PM_{2.5}$ concentration no higher than 5 $\mu g/m^3$ — half the 2005 guideline of 10 $\mu g/m^3$ and roughly half the current binding U.S. standard. These guidelines are explicitly non-binding and intended as a health-based benchmark rather than an enforceable standard, but they are widely used by governments and researchers as a reference point. The great majority of the world's population lives in places that exceed even this stricter benchmark: contemporaneous modeling work estimates that more than 90 percent of the global population is exposed to annual average $PM_{2.5}$ above the 5 $\mu g/m^3$ guideline, with exceedance most severe in South and East Asia, sub-Saharan Africa, and parts of the Middle East. Many middle- and low-income countries either lack enforceable national $PM_{2.5}$ standards altogether or set them well above the WHO guideline; the European Union's binding annual limit, for comparison, is 25 $\mu g/m^3$, itself far looser than the WHO recommendation. Ambient $PM_{2.5}$ concentrations, and the stringency of regulation governing them, thus vary enormously across countries, with the world's most populous regions typically facing both higher exposure and looser regulatory constraints than the United States.

3 Data and Descriptive Statistics

Our analysis combines two nationwide data sets, one containing student test scores and one measuring pollution exposure. Test scores come from student-level longitudinal data from the College Board, which administers a variety of standardized exams including the PSAT. We combine this student-level data with high-resolution measures of local ambient $PM_{2.5}$ concentrations, local weather data that includes daily temperature, precipitation and wind, and measures of wildfire smoke. We focus throughout on students who took the PSAT exam more than once, which allows us to compare each student's performance across attempts that may differ in the air quality experienced in the preceding year. Below we describe each data source in turn and then present summary statistics for the analysis samples.

3.1 Test Score Data

Our College Board test score data cover the universe of the millions of U.S. high school students who each year take the PSAT, an exam containing a reading and a math section.[5] The PSAT is

[5]Prior to 2015, the PSAT also contained a Writing section. In the 2015-16 academic year, College Board introduced two additional PSAT assessments (PSAT 8/9 and PSAT 10) to provide a vertically scaled, longitudinal way to measure student progress from grades 8-12. The youngest cohort in our data graduated in 2016 and did not have these new assessments available to them. In our sample, all students took the PSAT/NMSQT.

administered annually, with nearly all test administrations occurring during the second and third weeks of October. Most students first take it in tenth or eleventh grade, though some begin as early as ninth grade. Students take and retake the PSAT for a variety of reasons, including preparation for the SAT college entrance exam, qualification for National Merit Scholarships, and obtaining information about their college readiness.

The PSAT has several features that make it well-suited for studying the cumulative effects of pollution on learning. The PSAT is taken by a large share of U.S. secondary school students, including many who never take the SAT. Across the high school classes of 2007–2016, our data include more than 20 million PSAT takers, equivalent to roughly 61–69% of students in the corresponding U.S. graduating cohorts, providing unusually broad national coverage of student achievement.[6] Because the exam is offered during a well-defined testing window with advanced registration required, there is limited scope for endogenous taking or timing decisions. Proctoring is nationally harmonized and the test is centrally graded, limiting the potential for score manipulation. Finally, the test is designed to assess cumulative progress on skills learned during secondary school rather than generalized intelligence, making it well-suited for detecting the effects of environmental conditions on the rate of human capital accumulation rather than on test-taking ability.

Our data contain test scores and test dates from the universe of PSAT-takers from high school classes between 2001 and 2022, though we restrict our primary analysis to exam years between 2007 and 2016 given the availability of pollution and wildfire smoke measures. To allow for student fixed effects analyses, we further focus on the 39% of PSAT takers who retake the exam at least once, a rate computed over the full 2001–2022 population of takers; within our 2007–2016 analysis window, 45.0% of takers retake the exam. After excluding a small number of students who take the PSAT on atypical dates, we end up with a working dataset of 9.8 million PSAT retakers whose 18.8 million scores form the core of our analysis sample. Our primary outcome is a student's combined math and reading score, standardized by test administration and scaled by a factor of 1,000 so that coefficients can be interpreted in thousandths of a standard deviation.[7]

[6]U.S. graduating cohorts during the study period contained approximately 3.0–3.4 million students annually according to National Center for Education Statistics estimates.

[7]To put scores from both two-section and three-section PSAT regimes on a comparable scale, we construct a pseudo two-section score for the three-section period by averaging Reading and Writing and adding Math. We then standardize both pseudo and redesigned PSAT total scores separately by regime and within each fall administration (for example, Fall 2010).

In addition to PSAT scores, the College Board data also contain students' demographic and geographic information. Demographic information includes gender, race, and parental education. Geographic information includes students' home ZIP codes. High school identifiers, representing the final school on record for each student, allow us to assign students to specific $PM_{2.5}$ exposure levels at the school's location and to merge school-level characteristics. We classify schools by their level of socioeconomic challenge using the College Board's Challenge index, a nationally normed percentile summarizing the educational and socioeconomic conditions in the communities served by each high school. We define high challenge high schools as those falling into the highest tercile of the national distribution.

3.2 Air Pollution Data

We obtain daily local $PM_{2.5}$ measures predicted using the machine learning approach of Di et al. (2019). This approach combines observed $PM_{2.5}$ from ground-station monitoring data with satellite imagery, chemical transport models, and weather reanalysis products to produce daily predicted $PM_{2.5}$ levels over the contiguous United States from 2000 to 2017 at a 1km x 1km grid resolution. For each high school in our data, we construct daily pollution measures by averaging all grid cells within 10 kilometers of the school's latitude and longitude, or if coordinates are unavailable, all grid cells within the school's ZIP code. This produces a dataset of school-specific daily $PM_{2.5}$ levels for each day from 2000 to 2017.

We characterize students' pollution exposure as the total number of days falling into various concentration bins during the 365 days prior to the date they took the PSAT, focusing on the two most common exam dates in each year. Our primary measure distinguishes three ranges: "non-polluted" days with $PM_{2.5}$ below 8 $\mu g/m^3$, moderately polluted days with $PM_{2.5}$ between 8 and 12 $\mu g/m^3$, and heavily polluted days with $PM_{2.5}$ at or above 12 $\mu g/m^3$. Days below 8 $\mu g/m^3$ constitute the reference category in all specifications. Using state-specific school calendars, we further distinguish pollution exposure on school-year days from pollution exposure during summer months, which allows us to test the extent to which potential adverse effects are concentrated during instructional time.

3.3 Wildfire Smoke Data

We construct an instrument for local $PM_{2.5}$ using data from Childs et al. (2022), which provides the daily contribution of wildfire smoke to ambient $PM_{2.5}$ concentrations across the contiguous United States. These data are constructed by training a machine learning model on information from ground-based monitoring stations, satellite imagery, aerosol parcel transport models, and weather reanalysis products, and then applying the trained model to predict the level of $PM_{2.5}$ attributable to wildfire smoke within 10km x 10km grid cells for each day from 2006 to 2023. The resulting dataset has been used extensively in both the economics and public health literatures to document the adverse impacts of wildfire smoke on mortality (Gould et al., 2024; Connolly et al., 2024), morbidity (Zhang et al., 2025), and the contribution of wildfires to $PM_{2.5}$ trends across the United States (Burke et al., 2023).

We use a version of these data aggregated by Childs et al. (2022) to U.S. ZIP Code Tabulation Areas (ZCTAs), which provides predicted smoke-attributable $PM_{2.5}$ levels for each ZCTA on each day of the year. We assign schools to smoke exposure levels based on the ZCTA corresponding to each school's ZIP code. The smoke data distinguish between days with no detected smoke plume, days with a detected plume but no predicted smoke-attributable $PM_{2.5}$, and days with positive smoke-attributable $PM_{2.5}$. We group the first two categories together as days with zero smoke contribution, and characterize remaining days by the level of smoke-attributable $PM_{2.5}$. Positive smoke exposure is partitioned using the same $PM_{2.5}$ thresholds used to define total pollution exposure: moderate smoke days between 8 and 12 $\mu g/m^3$ and heavy smoke days at or above 12 $\mu g/m^3$. Because smoke-attributable $PM_{2.5}$ can be zero, we use one additional category for the instrument, light smoke days with smoke $PM_{2.5}$ between 1 and 8 $\mu g/m^3$. The overlap between our $PM_{2.5}$ data and the wildfire smoke data covers PSAT exam years from 2007 to 2016, which defines the sample period for our primary estimates. We additionally exclude schools within 10 kilometers of wildfire perimeters to avoid conflating pollution effects with direct effects of wildfires.

3.4 Summary Statistics

Table 1 presents summary statistics for two samples of students taking exams from 2007 to 2016, the period during which we have both pollution and wildfire smoke data. The first column contains all takers, students who took the PSAT at least once during this time. The second column shows all retakers, students who took the PSAT at least twice and are the focus of our main analysis. Comparing the two columns reveals that 45.0% of PSAT takers retake the exam during this time period and that those retakers are demographically similar to the overall PSAT-taking population.

The retaker sample is large and demographically diverse. We have nearly 19 million scores from nearly 10 million retakers attending nearly 25,000 schools. Given that our data span ten years, this implies that each year of data contains about one million students. As Panel A shows, 53% are female, 51% identify as White, 9% as Asian, 14% as Black, and 23% as Hispanic. Approximately 59% of retakers report that at least one parent holds a bachelor's degree or higher. The average student in our analysis sample attends a high school with a Challenge score of 41.2, meaning that the school is more advantaged than roughly 59 percent of U.S. domestic high schools (i.e. is in the 41st percentile of the challenge distribution). Panel B shows that retakers take the PSAT an average of 2.18 times, with first-attempt scores near zero by construction due to our exam-year standardization.

Students are exposed to a substantial number of pollution days in each year prior to the PSAT. As panel C shows, the average retaker in our sample experiences 189.1 days each year with $PM_{2.5}$ levels at or above 8 $\mu g/m^3$, meaning that slightly more than half of days qualify as polluted. Of those 189.1 pollution days, 102.7 involve moderate pollution levels ($PM_{2.5}$ between 8 and 12 $\mu g/m^3$) and 86.5 involve heavy pollution levels ($PM_{2.5}$ at or above 12 $\mu g/m^3$). Nearly three-fourths of pollution days fall during the school year, with the remaining fourth falling during the summer. As seen in panel D, the typical retaker experiences 23 days in which wildfire smoke contributes to ambient air pollution, an amount nearly identical to the full population of PSAT takers.

Test scores vary substantially across racial groups and school challenge levels. As panel A of Table 2 shows, White and Asian students have higher than average first-attempt PSAT scores (0.27 and 0.49 standard deviations above the national mean respectively), while Black and Hispanic students have lower than average scores (0.62 and 0.47 standard deviations below the mean respectively). PSAT scores are also closely related to school Challenge. Students attending schools in the bottom two-thirds of the national Challenge distribution average 0.17 standard deviations above the national mean, whereas students attending top-third schools average 0.55 standard deviations below the mean. These differences underscore the importance of a within-student empirical design that does not rely on cross-sectional comparisons across groups.

Pollution exposure also varies meaningfully across demographic groups and school types in ways that further motivate our within-student design. As panel B of Table 2 shows, White students experience fewer polluted days on average (184 per year) than Black (205), Hispanic (192), or Asian (191) students. This is true both of moderate pollution days, which range from 99 for White students to 109 for Hispanic students and for heavily polluted days, which range from 85 for White students to 98 for Black students. Students attending schools in the top third of the national Challenge distribution experience, on average, 203 polluted days per year, compared with 184 among students attending schools in the bottom two-thirds of the distribution. These patterns are consistent with the well-documented relationship between socioeconomic disadvantage and pollution exposure (Colmer et al., 2020; Jbaily et al., 2022). As panel C shows, exposure to pollution generated by wildfire smoke varies across race and high schools somewhat less and in less systematic ways than does air pollution generally.

Pollution exposure also varies substantially across places, students, and years. Figures 1 through 3 map the county-level geography of this exposure: polluted days are far more common in the eastern half of the country and in Southern and Central California than in the Mountain West, with moderately polluted days widespread across the East and South (Figure 2) and heavily polluted days concentrated in the Midwest, the Southeast, and Southern California (Figure 3). This pronounced cross-sectional variation is precisely the variation our within-student design avoids relying on; our estimates instead exploit year-to-year changes in exposure at a given school. Figure 4 illustrates the considerable variation in pollution exposure across students in our sample. Per our chosen definition of polluted days, some students experience fewer than 20 polluted days per year, whereas others experience moderate or heavy pollution throughout nearly all of the 365 days in a year.

Our IV strategy leverages variation in local $PM_{2.5}$ driven by plausibly exogenous changes in wildfire smoke. Figure 5 shows the spatial distribution of average wildfire smoke across the US. While fairly concentrated in the West, Pacific Northwest, and Northeast, wildfire smoke appears to be measurably present throughout the country, consistent with its propensity to be carried long distances (Burke et al., 2021). As shown in Figure 5, many parts of the country experience at least 3 micrograms per cubic meter of smoke pollution per year, with some experiencing more than twice that amount. While the majority of days have little to no detected smoke, there is considerable year-to-year variation in the number of days with moderate or heavy smoke exposure, as depicted in Figure 6. This variation reflects differences in fire activity and prevailing wind patterns across our sample period rather than systematic trends, providing the kind of quasi-random variation across years that our identification strategy requires.

4 Empirical Strategy

Our empirical strategy exploits within-student variation in prior-year air quality across repeated PSAT attempts to estimate the effect of cumulative pollution exposure on learning. Because many students take the PSAT more than once, we can compare the same student's performance across attempts that differ in the pollution they experienced in the preceding year, controlling for everything fixed about that student over time. This within-student design addresses the concern that students in more polluted places differ systematically from those in less polluted places in ways that independently affect learning. It does not, however, fully address the possibility that year-to-year changes in local pollution are driven by changes in local economic conditions that also affect learning, such as a new manufacturing facility that simultaneously raises local $PM_{2.5}$ and stimulates the local economy. To isolate variation in pollution that is plausibly free of this confound, we combine the within-student design with an instrumental variables strategy that uses wildfire smoke from distant fires as a source of exogenous variation in local $PM_{2.5}$. We first describe the OLS specification and then the IV strategy.

4.1 Within-Student OLS Design

The intuition behind our OLS design is that the same student, attending the same school, may experience meaningfully different levels of air pollution in the year preceding one PSAT attempt compared to another. By comparing a student's performance across these attempts, we can ask whether the attempt preceded by more polluted days yields lower scores, while holding constant the student's underlying ability, motivation, household background, and the long-run quality of their school.

Formally, let $Score_{isctk}$ denote the standardized PSAT score for student i, attending high school s, from graduating cohort c, taking the PSAT on exam date t for the kth time. We construct $Score_{isctk}$ by standardizing the composite PSAT score within each fall administration and scaling by a factor of 1,000, so that coefficients can be interpreted in thousandths of a standard deviation. We characterize pollution exposure using bins that count the number of days in the 365 days prior to the exam on which daily average $PM_{2.5}$ fell within each concentration range, with days below 8 $\mu g/m^3$ serving as the reference category in all specifications.

Our baseline specification uses a single measure of total polluted days, $PollutedDays_{st}$, which counts all days in the year prior to exam date t during which school s experienced $PM_{2.5}$ at or above 8 $\mu g/m^3$:

$$Score_{isctk} = \beta PollutedDays_{st} + \mu_i + \lambda_{ctk} + X'_{st}\gamma + \varepsilon_{isctk} \tag{1}$$

Student fixed effects μ_i absorb all time-invariant student characteristics, including ability, motivation, and household background. Our coefficient of interest β is thus identified from the relationship between within-student changes in prior year pollution exposure and within-student changes in test scores across test takes. Cohort-test date-attempt number fixed effects λ_{ctk} flexibly control for a variety of potential confounds, including differential selection into test-taking across high school classes, differential difficulty of the test across testing dates, and differential test performance based on past number of test takes.

Though we begin without additional controls beyond the above fixed effects, we successively add a vector X_{st} that includes a rich set of controls that vary by school and test date. These include prior-year average maximum temperature and precipitation at the student's school, exam-day $PM_{2.5}$, exam-day temperature and precipitation, and Census-division-specific linear time trends in exam year to account for region-specific changes in air quality, schooling, or test-taking behavior over time. We also control for county-level unemployment in the exam year to address the possibility that local economic shocks simultaneously influence local air quality and educational achievement. Standard errors are clustered by high school, the level of the identifying variation.

To investigate whether effects differ across pollution intensity levels, we also decompose pollution exposure into moderately polluted days ($ModerateDays_{st}$, with $PM_{2.5}$ between 8 and 12 $\mu g/m^3$) and heavily polluted days ($HeavyDays_{st}$, with $PM_{2.5}$ at or above 12 $\mu g/m^3$), estimating:

$$Score_{isctk} = \beta_1 ModerateDays_{st} + \beta_2 HeavyDays_{st} + \mu_i + \lambda_{ctk} + X'_{st}\gamma + \varepsilon_{isctk} \tag{2}$$

This decomposition allows us to test whether the learning impacts of pollution are driven by the frequency of more common moderate pollution days or by heavier pollution events. The same fixed effects and control vector apply in both specifications, and both are estimated on the same sample of students and attempts.

Identification within this design comes from comparing the same student's PSAT performance across attempts preceded by different pollution histories, after removing the influence of fixed student characteristics, common shocks to students taking the exam at the same time, and the time-varying controls described above.[8] There is substantial within-student variation in prior-year pollution, with a standard deviation of roughly 13 days for total polluted days and 11 days for heavily polluted days.

4.2 Instrumental Variables Strategy

While the within-student design addresses selection on time-invariant characteristics, OLS estimates may still be biased if year-to-year changes in local pollution are correlated with year-to-year changes in local economic conditions that independently affect learning. A new factory that raises local $PM_{2.5}$ while also creating jobs and raising household incomes would cause OLS to understate the true harm of pollution; a local recession that reduces both pollution and school resources would cause OLS to overstate it. The direction of this bias is theoretically ambiguous, which makes it difficult to draw confident conclusions from OLS estimates alone.

A further concern is attenuation bias from measurement error. Our school-level $PM_{2.5}$ measures are constructed from a machine learning model that predicts pollution concentrations at 1km $\times$ 1km resolution, which introduces noise into the pollution exposure assigned to any given student. Classical

[8] See Figure A1 for illustration of the residual variation in pollution exposure that drives our estimates after netting out these fixed effects.

measurement error of this kind biases OLS estimates toward zero, causing them to understate the true effect of pollution on learning. By instrumenting for total $PM_{2.5}$ with smoke-derived $PM_{2.5}$, which is subject to independent measurement error uncorrelated with the error in total $PM_{2.5}$, the IV strategy purges this attenuation bias and recovers estimates closer to the true causal effect.

To address both concerns, we instrument for local $PM_{2.5}$ using variation in wildfire smoke from distant fires. Wildfire smoke plumes can travel thousands of miles from their source, raising local $PM_{2.5}$ concentrations in communities far removed from any active fire. For schools located well away from wildfire perimeters, year-to-year variation in smoke exposure is driven primarily by distant weather patterns and fire activity rather than by local economic conditions or other community-level factors that might independently affect student learning. This makes smoke-derived $PM_{2.5}$ a plausible instrument for total $PM_{2.5}$. It is strongly correlated with local pollution levels, but its variation is largely orthogonal to the local economic conditions that confound OLS estimates.

We implement this strategy using the smoke data described above, which characterizes each school-year observation by the number of days with light ($0 < PM_{2.5} < 8\ \mu g/m^3$), moderate ($8 \leq PM_{2.5} < 12\ \mu g/m^3$), and heavy ($PM_{2.5} \geq 12\ \mu g/m^3$) smoke-attributable $PM_{2.5}$ in the prior year. Our baseline first stage instruments for total polluted days using all three smoke bins:

$$\begin{aligned} PollutedDays_{st} = \pi_1 LightSmoke_{st} + \pi_2 ModerateSmoke_{st} + \pi_3 HeavySmoke_{st} \\ + \mu_i + \lambda_{ctk} + X'_{st}\gamma + \varepsilon_{isctk} \end{aligned} \tag{3}$$

where $LightSmoke_{st}$, $ModerateSmoke_{st}$, and $HeavySmoke_{st}$ denote the number of days with light, moderate, and heavy smoke-attributable $PM_{2.5}$ respectively. Student and cohort-test date-attempt number fixed effects, as well as additional controls, match those in the OLS specification.

The corresponding second stage is:

$$Score_{isctk} = \beta^{IV} \widehat{PollutedDays}_{st} + \mu_i + \lambda_{ctk} + X'_{st}\gamma + \varepsilon_{isctk} \tag{4}$$

where $Pollut\hat{e}dDays_{st}$ is the fitted value from the first stage equation (3). We also run specifications decomposing pollution days into moderate and heavy days and estimate a system in which both

$ModerateDays_{st}$ and $HeavyDays_{st}$ are jointly instrumented by the three smoke bins, mirroring the intensity decomposition in equation (2).

The key identifying assumption is that wildfire smoke affects student performance only through its effect on local $PM_{2.5}$ concentrations. That is, conditional on total $PM_{2.5}$, smoke exposure has no independent effect on learning. The most obvious threat to this assumption is that schools located near an active wildfire may experience both elevated smoke and a range of other disruptions that affect learning through channels other than air quality. To address this concern, we exclude all school-year observations for schools located within 10 kilometers of a wildfire perimeter in the relevant exam year. We later test the robustness of our results to changing that choice of exclusion distance. A secondary concern is that smoke exposure may be correlated with other weather conditions that independently affect learning. We address this by controlling flexibly for prior-year and exam-day temperature and precipitation in all specifications.

5 Results

5.1 OLS with Student Fixed Effects

We begin by presenting within-student OLS estimates of the effect of prior-year pollution on PSAT performance, and then show how these estimates change when we instrument for pollution using wildfire smoke. The comparison between OLS and IV estimates is itself informative: as we discuss below, the IV estimates are substantially larger, consistent with economic confounds and attenuation bias from measurement error in our school-level pollution measures.

Our OLS estimates suggest consistently negative impacts of pollution, particularly from days with moderate levels of $PM_{2.5}$. Table 3 reports estimated impacts of pollution from Equations 1 and 2, starting with the sparsest specification in column 1 (which includes only student and cohort-test date-attempt number fixed effects) and adding further controls in subsequent columns. Estimates from that sparsest specification suggest that pollution days generally have negative but small and statistically insignificant effects on learning.

Across specifications that add a variety of controls, we see consistently negative impacts of moderate pollution days on learning, on the order of 0.05-0.07 thousandths of a standard deviation per additional day of exposure. Column 2 adds controls for prior year temperature and precipitation, to account for correlations between pollution and weather exposure, and exam day pollution and weather, to account for correlations between test-taking conditions and cumulative exposure to pollution. Column 3 adds Census division-specific trends to account for geographic differences in both pollution and test score time trends. Column 4 further accounts for county-level unemployment rates, given concerns that pollution can be driven in part by economic conditions. As such controls are added, the damage from overall pollution days stabilizes around 0.02 thousandths of a standard deviation per additional day of exposure, while the impact of each moderate pollution day is generally around 0.07 thousandths of a standard deviation. Heavy pollution days have positive but statistically insignificant coefficients, suggesting little clear evidence of impact on learning.

The difference in estimated impacts between moderate and heavy pollution days is potentially surprising. One explanation is that students and schools engage in greater adaptive behavior on the most severely polluted days by staying indoors, canceling outdoor activities, or otherwise reducing exposure, thus attenuating the measured impact.[9] A second possibility, which partly motivates our IV strategy, is that high pollution days are disproportionately driven by local economic activity that has offsetting positive effects on student performance, causing OLS to understate the true harm of heavy pollution.

5.2 IV with Student Fixed Effects

We begin our IV analysis by showing that wildfire smoke has a clear, strong relationship to $PM_{2.5}$ exposure. Table 4 presents first-stage results from equation 3. Panel A shows how smoke-attributable pollution predicts the number of polluted days, while Panel B shows how it predicts the number of moderately and heavily polluted days separately. Panel A shows that all three smoke pollution categories increase the number of polluted days, with the largest effects coming from moderate smoke pollution days. Panel B further shows how this relationship varies across pollution intensity. Light and moderate smoke pollution days strongly and clearly raise the frequency of moderate $PM_{2.5}$ pollution days, turning low pollution days into moderate ones. In contrast, heavy smoke pollution days reduce that frequency because such days turn low and moderate pollution days into heavy

[9]Related to this, there is also evidence that differences between measured and actual exposure are much greater on heavily polluted days than on low pollution days (Dangel, 2025).

ones. Similarly, moderate and heavy smoke pollution days substantially increase the frequency of heavy $PM_{2.5}$ pollution days, as expected. The magnitude and significance of these results is robust to all of the additional controls we include across columns 2 through 4. The instrument is strong across all specifications: F-statistics exceed 300 in all cases, well above conventional thresholds for instrument strength (Lee et al., 2022).

Air pollution driven by wildfire smoke has a substantial negative impact on learning. Table 5 shows the IV estimates from equation 4 in panel A. Panel B shows estimates from treating moderate and heavy pollution days as separate endogenous regressors. The simplest specification in panel A column 1 suggests that each additional pollution day triggered by wildfire smoke reduces student learning by 0.28 thousandths of a standard deviation, a highly statistically significant result. In column 2, controlling for prior year weather and exam day conditions barely changes this effect to 0.25 thousandths of a standard deviation. Further controlling for division-specific time trends slightly decreases the estimated effect to a still highly statistically significant 0.21 thousandths of a standard deviation, while controlling for county-level unemployment has little additional effect.

The negative impacts of wildfire-driven air pollution on learning are driven by the effects of moderate pollution days (panel B). Across all specifications, the negative effect of each additional moderate pollution day is remarkably stable, at about 0.50 thousandths of a standard deviation. In our simplest specification with no controls beyond fixed effects, wildfire-driven heavy pollution days have a negative impact of 0.17 thousandths of a standard deviation. That estimated effect shrinks, however, with inclusion of prior year and exam day conditions, and declines to near zero with inclusion of division-specific time trends. We thus see consistent evidence that moderate pollution days interfere with learning and less consistent evidence about the impact of heavy pollution days. Changing the fire exclusion distance leaves our results qualitatively unchanged.[10]

The substantially larger IV estimates are consistent with two sources of downward bias in OLS. First, the positive correlation between local economic activity and pollution may cause OLS to understate pollution's true harm, particularly for heavily polluted days that are more likely to be driven by industrial activity. The IV estimates, which isolate variation in pollution driven by distant wildfire smoke, are free of this confound. Second, classical measurement error in our school-level $PM_{2.5}$

[10]See Table A2 for details.

measures attenuates OLS coefficients toward zero. The IV purges this bias by using smoke-derived $PM_{2.5}$ as an instrument, which is subject to independent measurement error uncorrelated with the error in total $PM_{2.5}$. Together these results provide strong causal evidence that particulate matter meaningfully reduces academic performance, with effects concentrated in the moderate pollution range that students experience on a routine basis.

Table 6 extends the preferred IV specification (column 3 in Table 5) to examine whether the estimated effect of pollution reflects only the most recent exposure window or a broader history of exposure. Column (1) reproduces the preferred IV estimate on the subset of observations with available lagged exposure data. Consistent with the main results, moderate $PM_{2.5}$ exposure during the year preceding the PSAT is associated with a statistically significant decline in PSAT scores.

Column (2) augments this specification by including one-year-lagged pollution exposure. The contemporaneous coefficient remains negative and statistically significant, whereas the lagged coefficient is imprecisely estimated. Because pollution exposure is highly persistent across years, current and lagged exposure are strongly correlated, limiting the ability of this specification to separately identify their effects. Consequently, the imprecision of the lagged coefficient should not be interpreted as evidence that earlier exposure has no effect.

5.3 School-Year versus Summer Pollution Exposure

The results above establish that prior-year pollution reduces learning but do not tell us through what mechanism. A key distinction is between mechanisms that operate specifically during instructional time, such as cognitive interference on polluted school days, disruption to classroom instruction, or increased absenteeism, and longer-run health channels that would be expected to accumulate regardless of whether exposure occurs during the school year or summer. To distinguish between these possibilities, we decompose prior-year pollution into school-year and summer components, using state-specific school calendars to classify each day. If effects operate through instructional disruption or short-run cognitive impairment, we would expect the school-year component to drive our results. If effects instead reflect cumulative health impacts, summer pollution should matter as well.

Formally, we replace the full-year pollution measures in our main specifications with separate counts of polluted days during the school year and during summer months. School-year and summer pollution are then instrumented using corresponding measures of smoke-derived $PM_{2.5}$ from each season separately, with light, moderate, and heavy smoke bins linked separately to school-year and summer periods.

Learning is harmed by air pollution during the school year rather than during the summer. Table 7 presents IV estimates of the school-year versus summer decomposition, with column 1 and column 2 separately including school-year and summer measures and column 3 including both simultaneously. The results are striking. In that final column, the estimated effect of an additional school-year polluted day on PSAT performance is -0.19 thousandths of a standard deviation, while the effect of summer pollution days is positive and statistically insignificant. Panel B sharpens this result. Each additional school-year moderately polluted day reduces scores by 0.70 thousandths of a standard deviation, while the corresponding summer exposure coefficients are again positive and statistically insignificant.

These timing estimates strongly suggest that learning disruptions arising from polluted school days, rather than cumulative long-run health exposure, drive the relationship between particulate matter and academic performance. The absence of any comparable effect during summer months, when students are not engaged in formal instruction, is difficult to reconcile with a purely health-based mechanism and instead points toward instructional disruption and short-run cognitive interference as the likely primary channels through which pollution reduces learning.

5.4 Heterogeneity by School and Student Characteristics

A central question in understanding the educational impacts of air pollution is whether the same environmental exposure generates different learning consequences across populations facing different social and economic situations. If pollution affects learning primarily through biological or mechanical channels, its marginal impact should be similar across students and schools facing comparable exposure. In contrast, if learning losses are shaped by institutional, household, or individual behavioral responses, such as schools' and families' ability to mitigate exposure or adjust instructional and living environments, then the educational effects of pollution may vary systematically across both school contexts and student groups.

To assess whether pollution effects differ across socioeconomic conditions, we examine heterogeneity by both school-level and student-level characteristics. At the school level, we stratify high schools based on their College Board Challenge metric, a nationally normed percentile measure summarizing how difficult it is for students from a given high school to access higher educational opportunities, constructed using all U.S. domestic high schools.[11] We define high-challenge schools as those at or above the 67th percentile, corresponding to the top (most disadvantaged) third of schools nationally. At the student level, we stratify students by race, comparing Black and Hispanic students who are historically educationally disadvantaged relative to White and Asian students. We then re-estimate our preferred IV specifications separately within each group, focusing on school-year pollution to concentrate the analysis on the period when institutional and behavioral responses to pollution are most likely to shape learning outcomes.

Air pollution is substantially more harmful to the learning of students in more socioeconomically challenged schools than to students in less challenged schools. As seen in the first two columns of Table 8, the estimated effect of an additional school year pollution day on learning is a highly statistically significant -0.302 thousandths of a standard deviation in the most challenged third of high schools, compared to a statistically insignificant -0.101 thousandths of a standard deviation for schools in the bottom two-thirds of the challenge distribution. Panel B shows that in high-challenge schools, the estimated effect of a moderately polluted school day is -0.838 thousandths of a standard deviation, nearly four times larger than the effect in lower-challenge schools. Heavy pollution days appear to be damaging only in high-challenge schools.

Air pollution is substantially more harmful to the learning of Black and Hispanic students than to White and Asian students. As seen in the last two columns of Table 8, the estimated effect of an additional school year pollution day on learning is a highly statistically significant -0.311 thousandths of a standard deviation for Black and Hispanic students, compared to a positive and statistically insignificant effect for White and Asian students. Moderately and heavily polluted school days both have large, statistically significant negative impacts on Black and Hispanic students, while neither level of pollution has clear negative effects on White and Asian students.

[11] Challenge percentiles are first constructed at the Census tract level using tract-level characteristics from national surveys and all SAT, PSAT, and AP assessment takers across five consecutive graduating cohorts. High school Challenge percentiles are then computed as weighted averages of the Challenge percentiles of the Census tracts from which a school's assessment takers originate.

These results imply that particulate pollution disproportionately harms students with fewer community- or family-based resources. Importantly, similar patterns emerge across two conceptually distinct measures: a school-level measure of educational and socioeconomic challenge that is constructed independently of race, and a student-level measure based only on race. Although these measures are correlated in practice, their independent construction suggests that the observed heterogeneity may not be driven by any single dimension of disadvantage.

The concentration of impacts in the moderate pollution range suggests that the mechanism operates through the accumulation of many routine exposure days rather than through more extreme events. One interpretation is that disadvantaged schools and families have fewer resources to mitigate exposure on moderately polluted days — through improved HVAC systems, scheduling adjustments, or other protective measures — while extreme pollution events may trigger non-discretionary responses such as school closures that partially equalize exposure-related risks across groups. This pattern of greater vulnerability in disadvantaged settings has important implications for understanding how pollution contributes to educational inequality, a question we return to below.

6 Magnitudes

To translate our IV estimates into economically interpretable magnitudes, we scale the estimated per-day effects by the observed exposure profile of students in the IV-eligible sample and express the resulting achievement costs in standard deviation units and time-based learning equivalents. Throughout, we interpret these quantities as local average treatment effects for the subpopulation of students whose pollution exposure is shifted by distant wildfire smoke.

6.1 Total achievement cost of observed exposure

First, we compute the total achievement cost of pollution by comparing our estimates to a world in which students experience no polluted days, either moderate or heavy. The total annual achievement cost of pollution for the average student is the sum of the estimated per-day effects, weighted by average exposure in each concentration bin:

$$\text{Total cost} = \hat{\beta}_{8\text{–}12} \cdot \mathbb{E}[D_{8\text{–}12}] + \hat{\beta}_{\geq 12} \cdot \mathbb{E}[D_{\geq 12}] \tag{5}$$

Plugging in the IV coefficients from our preferred specification (Column 3, Table 5, Panel B) and student-weighted mean exposure levels from the IV-eligible sample yields an implied annual reduction in PSAT performance of −0.055 standard deviations.[12] If we repeat this exercise collapsing moderate and heavy days into our single measure of polluted day, our calculation implies a total reduction of −0.040 standard deviations.[13]

To express these effects in time-based units, we estimate an empirical learning rate from within-student score growth between consecutive PSAT attempts. The mean gain between the first and second attempt is 0.294 standard deviations.[14] This suggests that the total achievement costs of −0.040 to −0.055 standard deviations computed above are equivalent to 14 to 19 percent of typical yearly score growth. This cost comes almost entirely from moderate pollution exposure rather than from extreme events. The IV coefficient on heavily polluted days is small and imprecisely estimated (Table 5, Panel B), so the scaling exercise is driven almost entirely by the 104 moderate-pollution days the average student experiences.

6.2 Contribution to achievement gaps

Beyond the average effect, our estimates speak to whether air pollution widens existing achievement gaps. Two distinct channels could contribute. Groups may differ in how many polluted days they experience (differential exposure), and they may differ in how much each polluted day harms learning (differential sensitivity). Our heterogeneity estimates in Table 8 indicate that both channels operate, and that the second is quantitatively dominant. We therefore construct a measure of pollution's contribution to each gap that incorporates both, by evaluating the total achievement cost of observed exposure separately within each group using that group's own estimated effects.

For any group g, we define the pollution-attributable achievement cost as in equation (5), but using group-specific coefficients and group-specific exposure:

[12] See Appendix A1.1 for details on exposure levels.
[13] See Appendix A1.2 for details on construction of these average-student costs.
[14] See Appendix A1.3.2 for details.

$$\text{Cost}_g = \hat{\beta}_{8\text{–}12}\mathbb{E}[D^g_{8\text{–}12}] + \hat{\beta}_{\geq 12}\mathbb{E}[D^g_{\geq 12}] \tag{6}$$

where the coefficients are the school-year IV estimates from Table 8 and the exposure means are drawn from Table 2. Appendices A1.4 and A1.5 give the full group-specific decompositions for the racial and school-challenge gaps, respectively. The contribution of pollution to the gap between a more disadvantaged group A and a comparison group B is then the difference in these costs, $\text{Cost}_A - \text{Cost}_B$, which we express as a share of the observed first-attempt score gap. Unlike a conventional exposure-only decomposition, this quantity does not hold the pollution–score relationship fixed across groups; it should be read as an accounting summary of how much of the observed gap coincides with the combined pattern of differential exposure and differential sensitivity, not as a causal decomposition of the gap's underlying sources.

Pollution exposure accounts for a meaningful portion of the achievement gap between Black and Hispanic students and White and Asian students. The two groups differ only modestly in exposure, with Black and Hispanic students experiencing about 8 more moderate-pollution days and 3 more heavy-pollution days per year. The groups differ sharply, however, in sensitivity to pollution. Evaluating equation (6) for each group yields a pollution cost of roughly 0.118 standard deviations for Black and Hispanic students, against a cost near zero for White and Asian students. The resulting difference of approximately 0.125 standard deviations is equivalent to about 15 percent of the observed 0.83 standard deviation first-attempt gap between the two groups. The analogous comparison across school challenge terciles tells a similar story: the pollution cost is roughly 0.115 standard deviations in top-third highest-challenge schools versus 0.024 in lower two-thirds-challenge schools, a difference of about 0.092 standard deviations, or roughly 13 percent of the observed 0.73 standard deviation gap between lower- and high-challenge high schools.

Decomposing these differences into their two channels confirms that sensitivity, not exposure, drives them. For the challenge comparison, valuing the 9.6-day difference in moderate exposure at the low-challenge coefficient contributes only about 0.002 standard deviations, leaving essentially all of the 0.092 standard deviation total attributable to the fact that students in high-challenge schools lose substantially more learning per polluted day. The comparison across racial groups is more lopsided still, since the White and Asian coefficients are near zero: almost none of the implied contribution reflects differential exposure.

We regard these shares as illustrative upper bounds rather than point estimates. They rest on group-specific coefficients estimated in separate subsamples, and the comparison-group coefficients that do much of the work — the near-zero, occasionally wrong-signed effects for non-URM and low-challenge students — are the least precisely estimated in Table 8. A more conservative reading of this evidence is that the learning harm from pollution is concentrated almost entirely among Black and Hispanic students and students in high-challenge schools, with no statistically detectable effect for their more advantaged counterparts. Under either framing, the conclusion for educational inequality is the same: pollution does not merely add a roughly equal burden that happens to fall on differently-exposed groups, but does disproportionate damage precisely where baseline achievement is already lowest. The students most exposed to pollution are also those whose learning is most sensitive to it, so environmental disparities and educational disparities compound.[15]

6.3 Implications of declining pollution over time

A complementary way to assess the aggregate importance of our estimates is to ask how much of the change in student achievement over our sample period can be accounted for by the secular decline in pollution exposure. Mean exposure to days with $PM_{2.5}$ at or above 8 μg/m^3 fell by approximately 102 days over the estimation sample, reflecting broad improvements in air quality driven by Clean Air Act regulation and the ongoing shift away from coal-fired power generation.[16] Applying the IV per-day coefficient from the 8+ specification to this reduction implies an improvement of approximately 0.021 standard deviations, attributable to declining pollution exposure over this period, equivalent to roughly 7 percent of a typical year's worth of learning.[17]

Two caveats are worth stating. First, this calculation is descriptive: it scales a causal per-day estimate by an observed time-series change in exposure, but does not itself identify the causal effect of the pollution trend on achievement, since other factors affecting learning evolved concurrently.

[15]This is consistent with speculation in Biasi et al. (2025) that the reason bond elections for HVAC have higher impacts in more disadvantaged schools is that systems in these schools are worse at baseline. If that were the case one would expect the impact of the same dose of ambient pollution to have larger adverse impacts in these classrooms.

[16]See Figure A2 for details.

[17]See Appendix A1.6 for details of this accounting exercise.

Second, the estimate holds the pollution-score relationship fixed over time, which may not be appropriate if the composition of pollution sources or the adaptive capacity of schools changed over the sample period. Subject to these qualifications, the calculation suggests that the long-run decline in ambient particulate matter has meaningfully supported student achievement, and conversely that policy reversals which raise $PM_{2.5}$ concentrations back toward historical levels could impose achievement costs of similar magnitude.

7 Conclusion

This paper provides new evidence on how ambient air pollution affects the rate at which students learn. Using longitudinal PSAT data on nearly ten million U.S. high school students and a within-student design that exploits year-to-year variation in local $PM_{2.5}$, combined with an instrumental variables strategy based on distant wildfire smoke, we find that cumulative pollution exposure over the school year causally reduces academic performance. IV estimates are substantially larger than OLS estimates, consistent with both local economic conditions masking the true harm of pollution in naive comparisons and attenuation bias from measurement error in our pollution data.

The effects we document are economically meaningful and arise primarily from routine, moderate exposure rather than extreme pollution events. Scaling our preferred IV estimates by students' observed exposure implies an annual reduction in test-score growth of approximately 0.055 standard deviations (roughly 19 percent of typical yearly gains), driven almost entirely by days with $PM_{2.5}$ between 8 and 12 $\mu g/m^3$, levels common across most of the United States and below current regulatory thresholds. The concentration of these effects during the school year, together with the absence of a comparable effect during summer months, points toward instructional disruption and short-run cognitive interference rather than longer-run health accumulation as the operative channel, though we cannot rule out some role for the latter.

These harms are also unevenly distributed. Students in high-challenge schools and Black and Hispanic students experience pollution-driven learning losses several times larger than their more advantaged peers. Differential sensitivity, not differential exposure, accounts for most of this gap. As a result, pollution appears to widen, rather than simply add proportionally to, existing achievement disparities. Because pollution exposure is itself unequally distributed along the same demographic and socioeconomic lines, our results suggest that environmental inequality and educational inequality are mutually reinforcing.

Beyond their implications for the specific relationship we study, these results speak to broader questions about how environmental conditions shape human capital formation. They add to a growing body of evidence, alongside work on heat, school infrastructure, and indoor air quality, that the physical conditions in which children learn are a meaningful and policy-relevant input into achievement, distinct from more traditional determinants such as school spending or teacher quality. Our finding that measurable harm occurs well below the EPA's historical 12 μg/m^3 threshold is directly relevant to ongoing debates over particulate matter regulation, where the reconsideration of these standards is presently underway. The scale of our estimates also helps explain why returns to investments such as school HVAC upgrades have been found to be so large.

Several questions remain open for future work. First, our evidence is most consistent with short-run instructional disruption, but distinguishing between specific mechanisms — reduced attention and cognition during class time, teacher and staff absenteeism, disrupted school operations, or reduced time outdoors — would have different implications for mitigation strategies and remains largely unresolved. Second, our outcome measure, while unusually large in scale, captures only one dimension of human capital; extending this work to other outcomes, including postsecondary enrollment, persistence, and eventual labor market outcomes, would clarify whether these learning losses translate into longer-run economic costs. Finally, as wildfire activity itself grows more frequent and severe with a changing climate, understanding how school- and district-level adaptation — through air filtration, schedule flexibility, or remote instruction on high-pollution days — can mitigate these harms is an increasingly important area for both research and policy.

References

Aguilar-Gomez, S., H. Dwyer, J. S. Graff Zivin, and M. J. Neidell (2022, March). This is Air: The "Non-Health" Effects of Air Pollution. Working Paper 29848, National Bureau of Economic Research.

Andersen, C. M., J. Brandt, J. H. Christensen, L. M. Frohn, C. Geels, T. Hener, M. Simonsen, and L. Skipper (2025). Air pollution and cognition in children: Evidence from national tests in Denmark. *Journal of Human Resources*.

Ang, D. (2021). The effects of police violence on inner-city students. *Quarterly Journal of Economics 136*(1), 115–168.

Archsmith, J., A. G. 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.

Aucejo, E. M. and T. F. Romano (2016). Assessing the effect of school days and absences on test score performance. *Economics of Education Review 55*, 70–87.

Aziz, M. and L. Elbakidze (2025). Invisible barrier: The impact of air quality on chronic school absenteeism in the US. *Environmental and Resource Economics 88*(3), 655–680.

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

Berestycki, C. and M. K. Chen (2026). Behavioral responses to wildfire smoke: Insights from smartphone location data. *Proceedings of the National Academy of Sciences 123*(18), e2527320123.

Bharadwaj, P., M. Gibson, J. G. Zivin, and C. Neilson (2017). Gray Matters: Fetal Pollution Exposure and Human Capital Formation. *Journal of the Association of Environmental and Resource Economists 4*(2), 505–542.

Biasi, B., J. Lafortune, and D. Schönholzer (2025). What works and for whom? Effectiveness and efficiency of school capital investments across the US. *The Quarterly Journal of Economics 140*(3), 2329–2379.

Bishop, K. C., J. D. Ketcham, and N. V. Kuminoff (2023). Hazed and confused: The effect of air pollution on dementia. *Review of Economic Studies 90*(5), 2188–2214.

Burke, M., M. L. Childs, B. de la Cuesta, M. Qiu, J. Li, C. F. Gould, S. Heft-Neal, and M. Wara (2023). The contribution of wildfire to PM2.5 trends in the USA. *Nature 622*(7984), 761–766.

Burke, M., A. Driscoll, S. Heft-Neal, J. Xue, J. Burney, and M. Wara (2021). The changing risk and burden of wildfire in the United States. *Proceedings of the National Academy of Sciences 118*(2), e2011048118.

Carleton, T., A. Jina, M. Delgado, M. Greenstone, T. Houser, S. Hsiang, A. Hultgren, R. Kopp, K. McCusker, I. Nath, J. Rising, A. Rode, H.-J. Seo, J. H. Stock, and A. Zhang (2022). Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits. *Quarterly Journal of Economics 137*(4), 2037–2105.

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*(6), 1051–1077.

Carneiro, J., M. A. Cole, and E. Strobl (2024). Foetal exposure to air pollution and students' cognitive performance: Evidence from agricultural fires in Brazil. *Oxford Bulletin of Economics and Statistics 86*(1), 156–186.

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.

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How does your kindergarten classroom affect your earnings? Evidence from Project STAR. *Quarterly Journal of Economics 126*(4), 1593–1660.

Chetty, R., J. N. Friedman, and J. E. Rockoff (2014). Measuring the impacts of teachers I: Evaluating bias in teacher value-added estimates. *American Economic Review 104*(9), 2593–2632.

Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter (2020). The opportunity atlas: Mapping the childhood roots of social mobility. *Nature 586*, 398–403.

Childs, M. L., J. Li, J. Wen, S. Heft-Neal, A. Driscoll, S. Wang, C. F. Gould, M. Qiu, J. Burney, and M. Burke (2022). Daily local-level estimates of ambient wildfire smoke PM2.5 for the contiguous US. *Environmental Science & Technology 56*(19), 13607–13621.

Coleman, J. S., E. Q. Campbell, C. F. Hobson, J. McPartland, A. M. Mood, F. D. Weinfeld, and R. L. York (1966). Equality of educational opportunity. Technical report, U.S. Department of Health, Education, and Welfare, Office of Education, Washington, D.C.

Colmer, J., I. Hardman, J. Shimshack, and J. Voorheis (2020). Disparities in PM2.5 air pollution in the United States. *Science 369*(6503), 575–578.

Colmer, J. and J. Voorheis (2020, November). The grandkids aren't alright: The intergenerational effects of prenatal pollution exposure. Working paper, Center for Economic Studies, U.S. Census Bureau.

Connolly, R., M. E. Marlier, D. A. Garcia-Gonzales, J. Wilkins, J. Su, C. Bekker, J. Jung, E. Bonilla, R. T. Burnett, Y. Zhu, et al. (2024). Mortality attributable to PM2.5 from wildland fires in California from 2008 to 2018. *Science advances 10*(23), eadl1252.

Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica 78*(3), 883–931.

Currie, J. (2011). Inequality at birth: Some causes and consequences. *American Economic Review 101*(3), 1–22.

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*(4), 682–694.

Dang, H.-A. H., M. N. Do, and C. V. Nguyen (2025). The impacts of climate change and air pollution on children's education outcomes: Evidence from Vietnam. *Economics of Education Review 106*, 102656.

Dangel, A. (2025). Mind the PM2.5 gap! Comparing pollution exposure estimates from wearable sensors and ambient monitors. Presentation at the Annual Conference of the European Association of Environmental and Resource Economists (EAERE), Bergen, Norway. Conference presentation.

Deryugina, T., G. Heutel, N. H. Miller, D. Molitor, and J. Reif (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review 109*(12), 4178–4219.

Di, Q., H. Amini, L. Shi, I. Kloog, R. Silvern, J. Kelly, M. B. Sabath, C. Choirat, P. Koutrakis, A. Lyapustin, et al. (2019). An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. *Environment international 130*, 104909.

Dillender, M. (2021). Climate Change and Occupational Health: Are There Limits to Our Ability to Adapt? *Journal of Human Resources 56*(1), 184–224.

Dong, R., R. Fisman, Y. Wang, and N. Xu (2021). Air pollution, affect, and forecasting bias: Evidence from Chinese financial analysts. *Journal of Financial Economics 139*(3), 971–984.

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*(4), 36–65.

Fitzpatrick, M. D., K. Sadowski, C. Chen, S. Day, K. Konty, and K. Argenio (2026). Asthma management in schools: A tool for reducing absenteeism? *Economics of Education Review 113*, 102813.

Gethin, A. (2025). Distributional growth accounting: Education and the reduction of global poverty, 1980–2019. *The Quarterly Journal of Economics 140*(4), 2571–2618.

Gilraine, M. and A. Zheng (2022, May). Air Pollution and Student Performance in the U.S. Working Paper 30061, National Bureau of Economic Research.

Goodman, J. (2014). Flaking out: Student absences and snow days as disruptions of instructional time. NBER Working Paper 20221, National Bureau of Economic Research, Cambridge, MA. No journal version as of this writing; cite as a working paper.

Gould, C. F., S. Heft-Neal, M. Johnson, J. Aguilera, M. Burke, and K. Nadeau (2024). Health effects of wildfire smoke exposure. *Annual Review of Medicine 75*(1), 277–292.

Graff-Zivin, J. and J. Shrader (2016). Temperature extremes, health, and human capital. *The Future of Children 26*(1), 31–50.

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

Greenstone, M. (2002). The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 Clean Air Act amendments and the Census of Manufactures. *Journal of political economy 110*(6), 1175–1219.

Guryan, J., E. Hurst, and M. Kearney (2008). Parental education and parental time with children. *Journal of Economic Perspectives 22*(3), 23–46.

Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annu. Rev. Econ. 6*(1), 689–733.

Heckman, J. J., R. Pinto, and P. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103*(6), 2052–2086.

Heissel, J. A., C. Persico, and D. Simon (2022). Does Pollution Drive Achievement? The Effect of Traffic Pollution on Academic Performance. *Journal of Human Resources 57*(3), 747–776.

Huang, J., N. Xu, and H. Yu (2020, October). Pollution and Performance: Do Investors Make Worse Trades on Hazy Days? *Management Science 66*(10), 4455–4476.

Jackson, C. K., R. C. Johnson, and C. Persico (2015). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics 131*(1), 157–218.

Jbaily, A., X. Zhou, J. Liu, T.-H. Lee, L. Kamareddine, S. Verguet, and F. Dominici (2022). Air pollution exposure disparities across US population and income groups. *Nature 601*(7892), 228–233.

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

Lafortune, J., J. Rothstein, and D. W. Schanzenbach (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics 10*(2), 1–26.

Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter (2022). Valid t-ratio inference for IV. *American Economic Review 112*(10), 3260–3290.

Martin, D., D. A. Romero, and D. Salcedo (2025, February). The hidden cost of mobility: Public transit pollution and academic achievement. Working paper, SSRN. SSRN Working Paper No. 5125511.

Park, R. J. (2022). Hot temperature and high-stakes performance. *Journal of Human Resources 57*(2), 400–434.

Park, R. J., A. P. Behrer, and J. Goodman (2021). Learning is inhibited by heat exposure, both internationally and within the United States. *Nature Human Behaviour 5*(1), 19–27.

Park, R. J., J. Goodman, M. Hurwitz, and J. Smith (2020). Heat and learning. *American Economic Journal: Economic Policy 12*(2), 306–339.

Persico, C. L. and J. Venator (2021). The effects of local industrial pollution on students and schools. *The Journal of Human Resources 56*, 406 – 445.

Reardon, S. F., D. Kalogrides, and K. Shores (2019). The geography of racial/ethnic test score gaps. *American Journal of Sociology 124*(4), 1164–1221.

Rentschler, J. and N. Leonova (2023). Global air pollution exposure and poverty. *Nature communications 14*(1), 4432.

Shapiro, J. S., D. A. Keiser, et al. (2019). Consequences of the Clean Water Act and the demand for water quality. *The Quarterly Journal of Economics 134*, 349–396.

Tarduno, M. A. and R. Walker (2025). Misperceptions about air pollution: Implications for willingness to pay and environmental inequality. Technical report, National Bureau of Economic Research.

Todd, P. E. and K. I. Wolpin (2007). The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital 1*(1), 91–136.

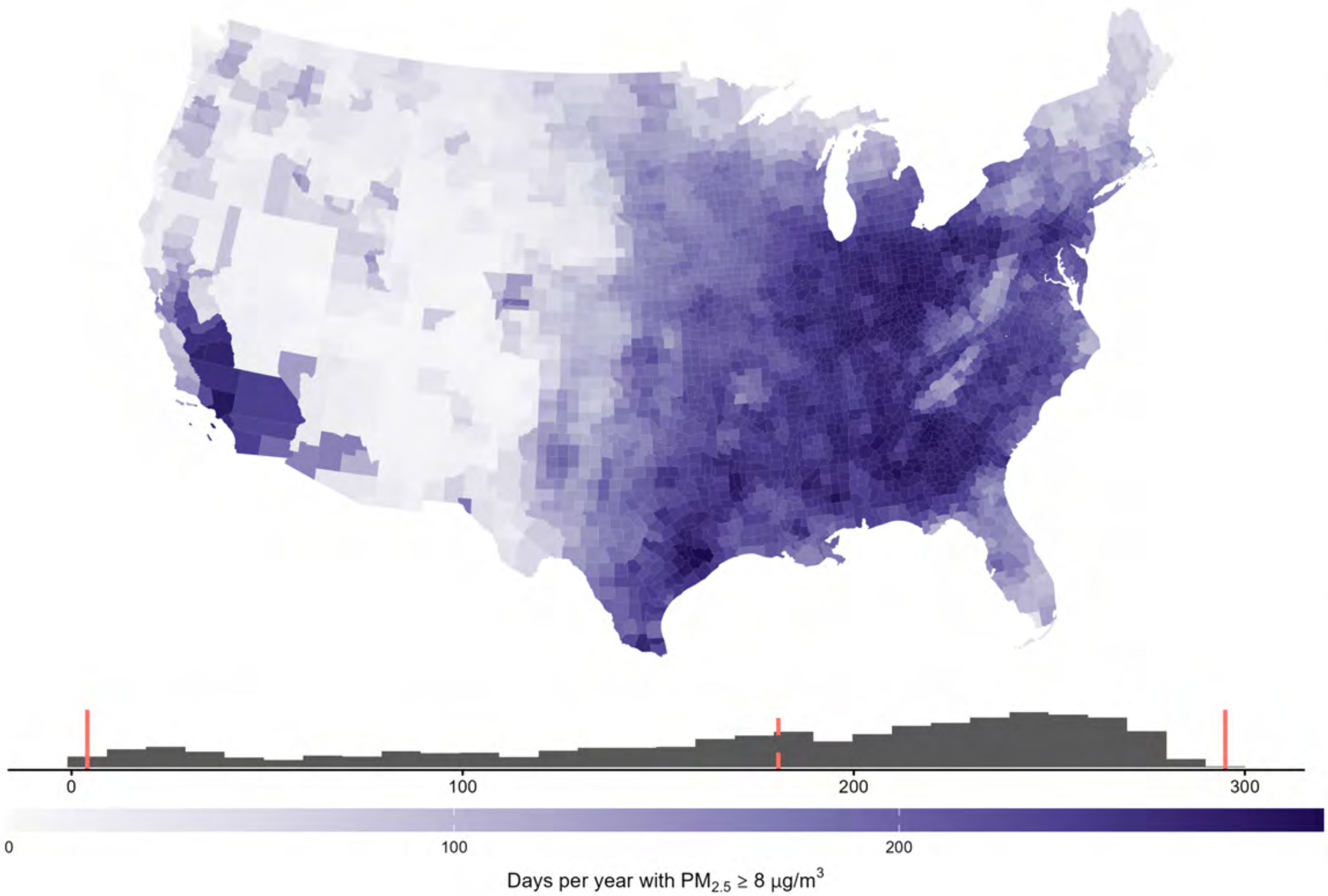
Voorheis, J. (2017, May). Air quality, human capital formation and the long-term effects of environmental inequality at birth. Working paper, Center for Administrative Records Research and Applications, U.S. Census Bureau.

Wen, J. and M. Burke (2022). Lower test scores from wildfire smoke exposure. *Nature Sustainability 5*(11), 947–955.

Zhang, Q., Y. Wang, Q. Xiao, G. Geng, S. J. Davis, X. Liu, J. Yang, J. Liu, W. Huang, C. He, et al. (2025). Long-range PM2.5 pollution and health impacts from the 2023 Canadian wildfires. *Nature 645*(8081), 672–678.

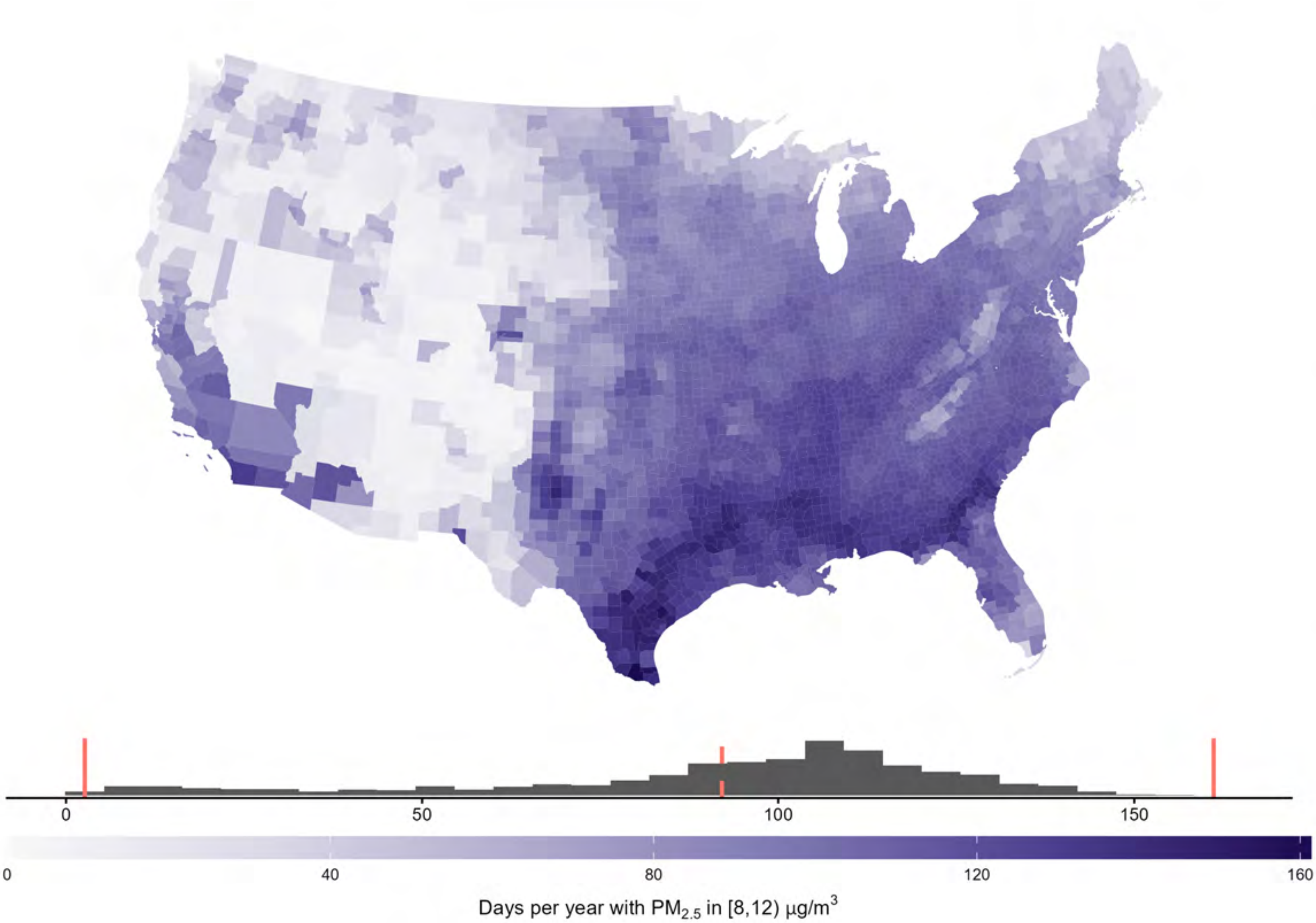
Zhang, X., H. Liu, X. Wu, L. Jia, K. Gadhave, L. Wang, K. Zhang, H. Li, R. Chen, R. Kumbhar, et al. (2025). Lewy body dementia promotion by air pollutants. *Science 389*(6764), eadu4132.

Figure 1: Spatial Distribution of Pollution



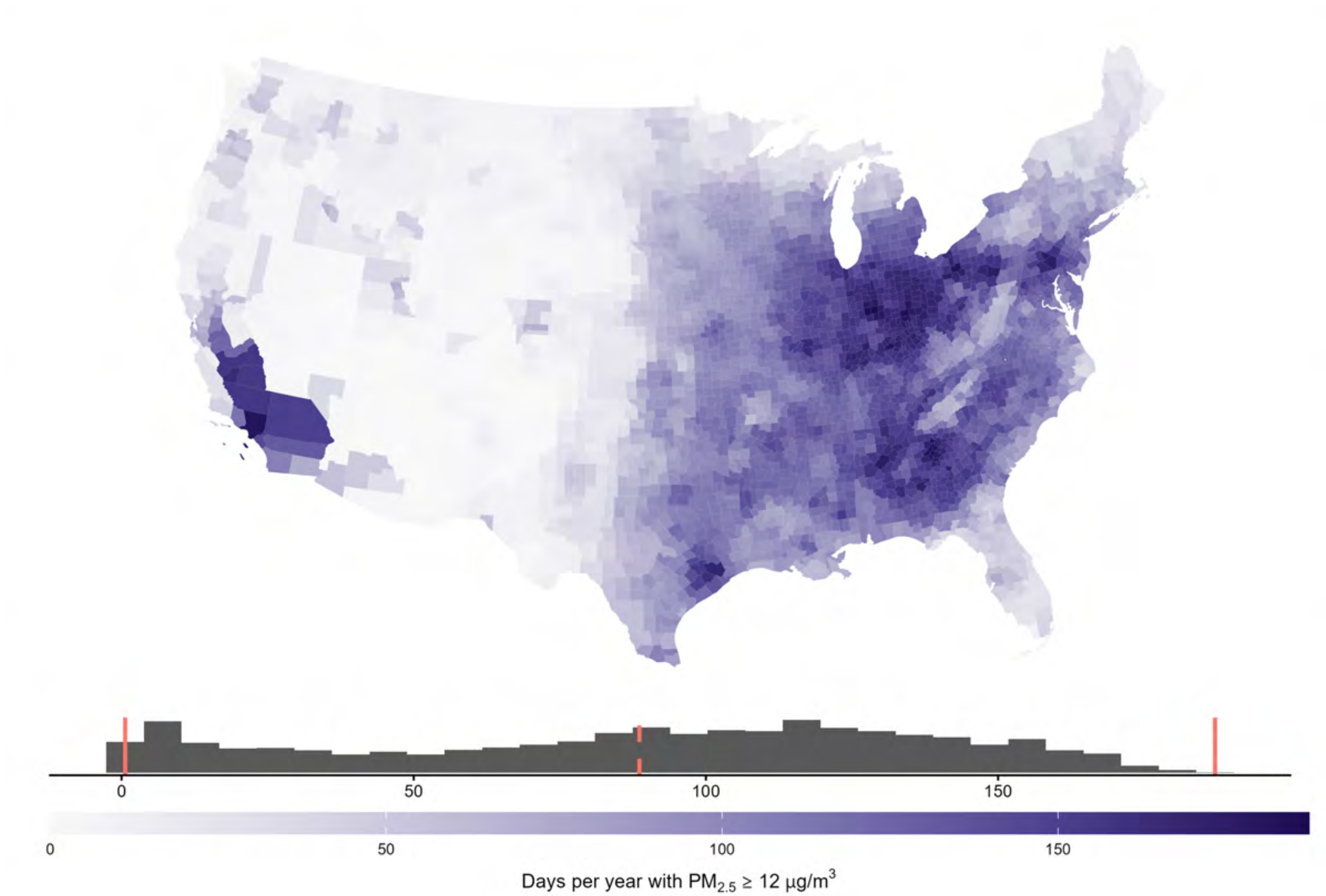
Notes: Shown above is the county-level average annual number of polluted days for PSAT takers in our sample. Polluted days are those with $PM_{2.5}$ of at least 8 μg/m^3. The bottom panel shows the distribution of county-level values, with red lines marking the minimum, mean, and maximum.

Figure 2: Spatial Distribution of Moderately Polluted Days



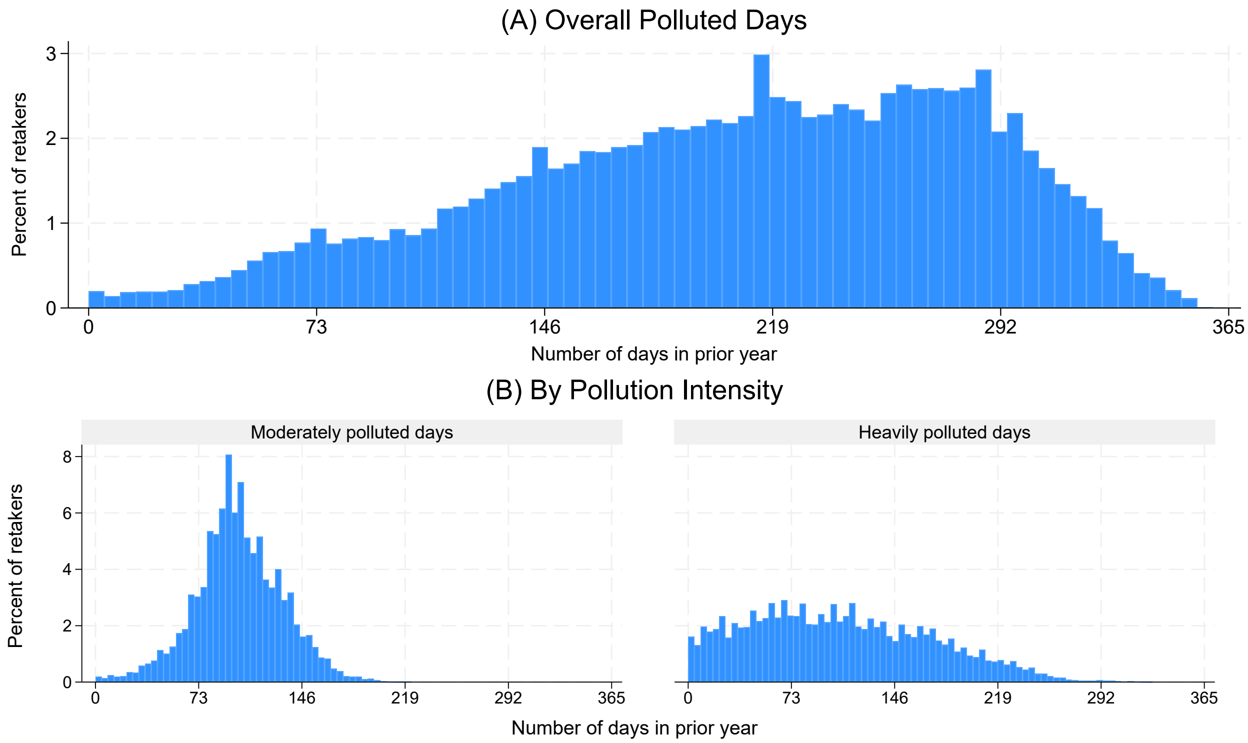
Notes: The above figure shows county-level average number of polluted days in PSAT exam years for all students from the high school classes of 2001-22 who took the PSAT at least once in grades 9-11. We defined a moderately polluted day as one where $PM_{2.5}$ was ≥ 8 and <12. County refers to the physical location of the high school the student last attended. The bottom panel shows the distribution of county-level values, with red lines marking the minimum, mean, and maximum. Pollution in PSAT years 1997-2000 and 2017-2019 was not available. COVID-impacted PSAT years 2020 and 2021 were also excluded.

Figure 3: Spatial Distribution of Heavily Polluted Days



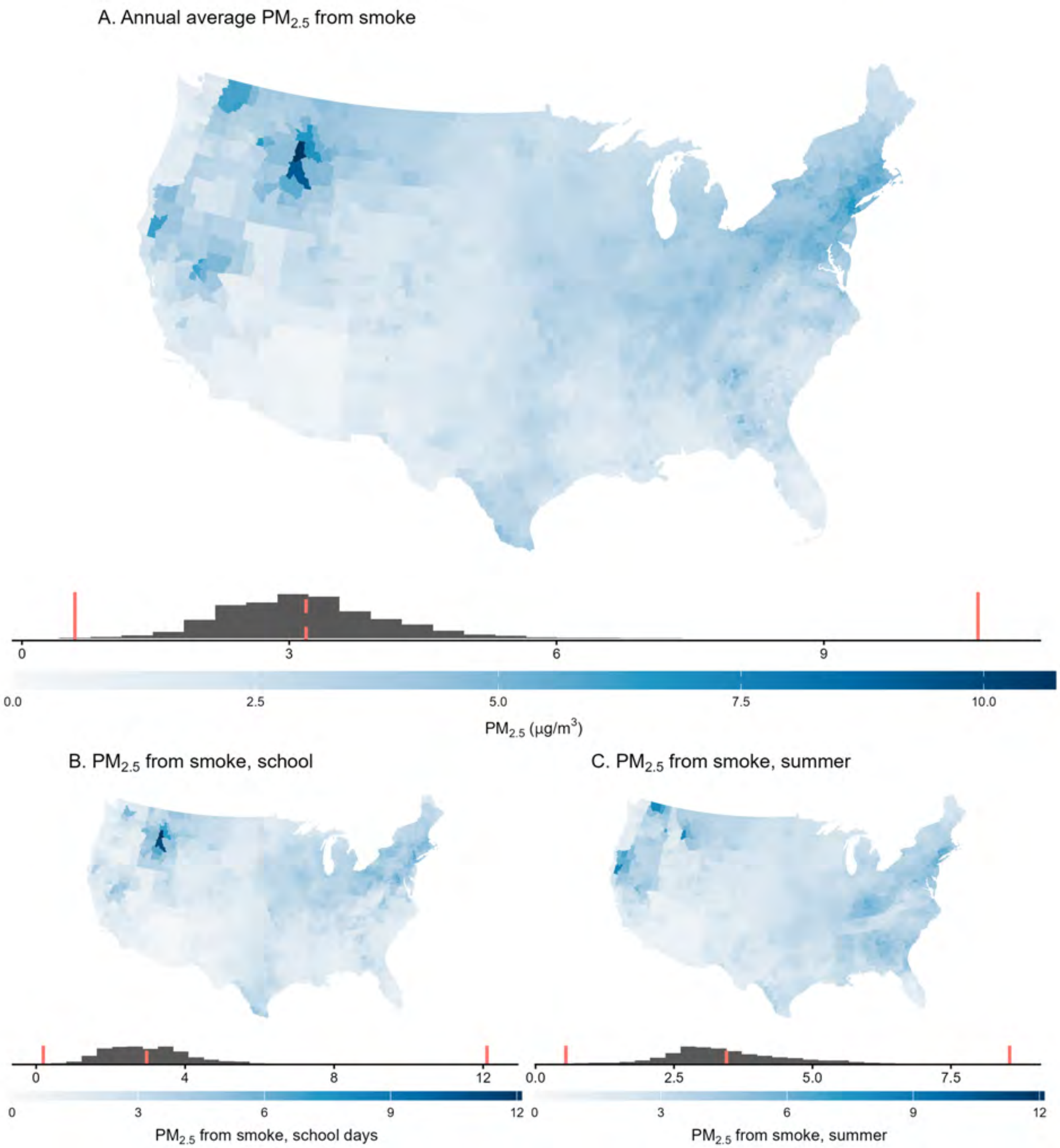
Notes: The above figure shows county-level average number of polluted days in PSAT exam years for all students from the high school classes of 2001-22 who took the PSAT at least once in grades 9-11. We defined a heavily polluted day as one where $PM_{2.5}$ was $\geq$12. County refers to the physical location of the high school the student last attended. The bottom panel shows the distribution of county-level values, with red lines marking the minimum, mean, and maximum. Pollution in PSAT years 1997-2000 and 2017-2019 was not available. COVID-impacted PSAT years 2020 and 2021 were also excluded.

Figure 4: Student-Level Variation in Pollution Exposure



Notes: Shown above is the distribution of students' average number of polluted days experienced at their school location in the year prior to their first PSAT take. School location is defined using a 2 km zone around the school's latitude and longitude, if available, or else using the school's ZIP code. The graphs include all students from the high school classes of 2001-22 who took the PSAT more than once in grades 9-11 and had available pollution data. In panel A, polluted days are those with $PM_{2.5}$ of at least 8 $\mu g/m^3$. Panel B distinguishes between moderately polluted days ($PM_{2.5}$ of 8-12 $\mu g/m^3$) and heavily polluted days ($PM_{2.5} \geq 12$ $\mu g/m^3$).

Figure 5: Spatial Distribution of Average Smoke $PM_{2.5}$



Notes: Panel A shows annual average smoke-attributable $PM_{2.5}$ concentrations. Panels B and C decompose smoke $PM_{2.5}$ into average concentrations during the school year and summer, respectively.

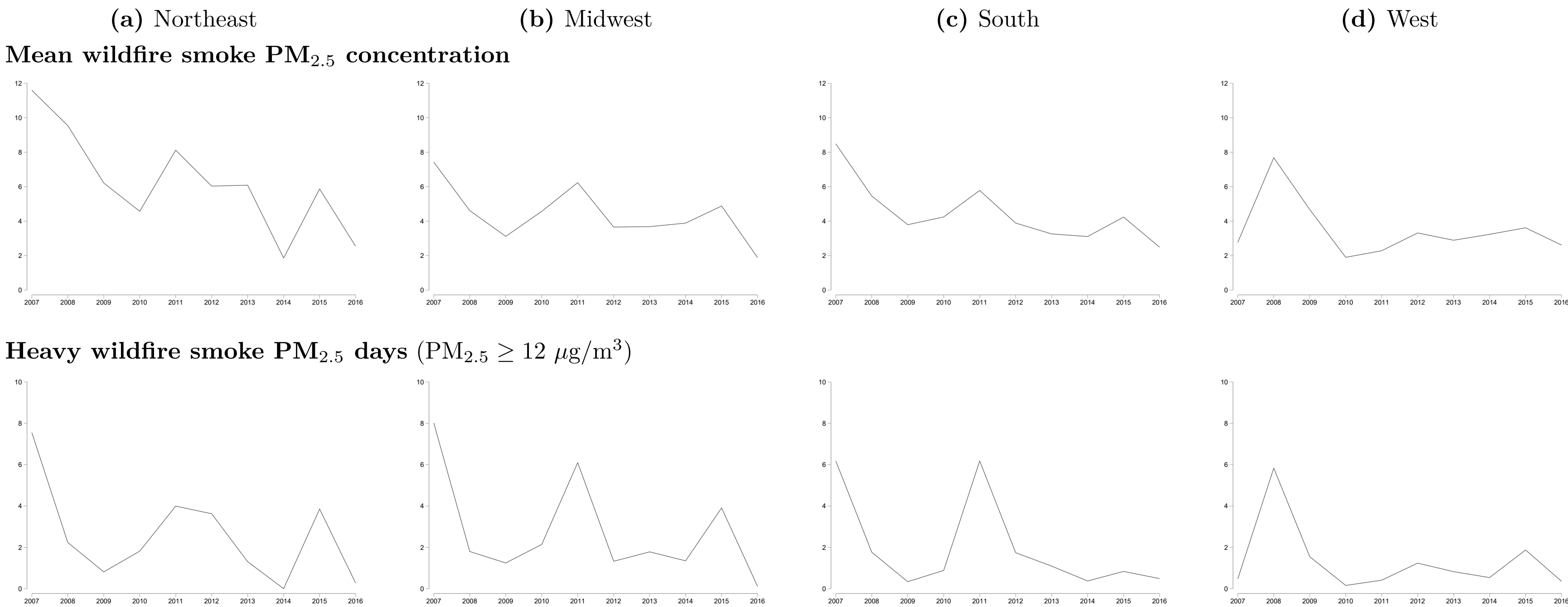


Figure 6: Regional trends in wildfire smoke $PM_{2.5}$ exposure

Notes: The top row plots the student-weighted mean concentration of $PM_{2.5}$ attributable to wildfire smoke in the year prior to the PSAT exam. The bottom row plots the student-weighted mean number of heavy wildfire smoke $PM_{2.5}$ days (days with $PM_{2.5} \geq 12\ \mu g/m^3$) in the prior year. Exposure measures are derived from the wildfire smoke dataset of Childs et al. (2022), aggregated to Census region by exam year, and weighted by the number of PSAT takers in each school–year. The sample includes exam years 2007–2016, the period for which wildfire smoke $PM_{2.5}$ data are available.

Table 1: Summary Statistics

	Eligible for Analysis	
	Takers	Retakers
Panel A. Demographics		
Female	0.52	0.53
White	0.52	0.51
Asian	0.07	0.09
Black	0.15	0.14
Hispanic	0.22	0.23
Parent BA	0.55	0.59
High School Challenge	43.77	41.15
Panel B. PSAT scores		
PSAT Attempts	1.54	2.18
First PSAT z-score	-0.11	-0.02
Panel C. Pollution		
Prior-year 8+ pollution days	185.7	189.1
Moderate [8,12) pollution days	100.7	102.7
Heavy [12,inf) pollution days	85.1	86.5
Schoolyear 8+ pollution days	138.6	140.7
Summer 8+ pollution days	47.2	48.5
Panel D. Smoke-attributable pollution		
Light [1,8) smoke pollution days	19.3	18.3
Moderate [8,12) smoke pollution days	2.4	2.3
Heavy [12,inf) smoke pollution days	2.0	2.1
N(attempts)	30,777,374	18,793,469
N(students)	21,805,403	9,821,498
N(schools)	29,347	24,838

Notes: Means of key variables are shown. Column 1 contains all students in our analysis sample who took the PSAT at least once. Column 2 contains students who took the PSAT at least twice. Summary statistics reflect the restricted analysis sample: PSAT takers with available pollution and smoke data, excluding COVID-affected years (2020–2021) and school-years within 10 km of a fire. Regression estimation samples are smaller because fixed-effects estimation additionally excludes observations with missing regression variables and iteratively drops singleton fixed-effect groups. Parent BA means at least one parent or guardian has a Bachelor's degree. High school challenge is a percentile relative to all U.S. high schools, with higher numbers representing more disadvantaged communities. Polluted days are those with $PM_{2.5}$ of at least 8 $\mu g/m^3$. These are also segmented into pollution days during the previous school year versus summer. Pollution days in the previous year are also segmented into bins: moderately polluted ($8 \leq PM_{2.5} < 12\ \mu g/m^3$) and heavily polluted ($PM_{2.5} \geq 12\ \mu g/m^3$) days. Smoke pollution days are defined using smoke-attributable $PM_{2.5}$ and classified into light, moderate, and heavy categories using the same $PM_{2.5}$ thresholds as the total pollution measure.

Table 2: Summary Statistics by Subgroups of Retakers

	Student Race				Challenge of High School	
	White	Asian	Black	Hispanic	Bottom Two-Thirds (1–66)	Top Third (67–100)
Panel A. PSAT Scores.						
PSAT Attempts	2.17	2.19	2.20	2.18	2.18	2.18
First PSAT z-score	0.27	0.49	-0.62	-0.47	0.17	-0.55
Panel B. Pollution.						
Prior-year 8+ pollution days	184.3	191.4	204.9	191.9	183.9	203.1
Moderate [8,12) pollution days	99.1	103.3	106.6	108.7	100.2	109.8
Heavy [12,inf) pollution days	85.2	88.1	98.3	83.2	83.8	93.3
Schoolyear 8+ pollution days	136.5	144.7	152.2	143.0	136.4	152.2
Summer 8+ pollution days	47.8	46.7	52.6	48.9	47.6	50.9
Panel C. Smoke-attributable pollution.						
Light [1,8) smoke pollution days	18.7	17.5	16.6	18.8	17.9	19.3
Moderate [8,12) smoke pollution days	2.5	2.1	2.4	2.0	2.3	2.2
Heavy [12,inf) smoke pollution days	2.2	1.9	2.5	1.8	2.1	2.1
N(attempts)	9,415,522	1,625,764	2,600,039	4,107,552	13,820,336	4,873,853
N(students)	4,901,415	856,583	1,331,871	2,180,268	7,219,416	2,546,399
N(schools)	14,350	1,473	3,272	4,176	14,704	7,222

Notes: Means of key variables are shown. The table contains subgroups of students who took the PSAT at least twice in years with available pollution and smoke data, excluding school-years with distance to fires $\leq$ 10 km (the sample analyzed with instrumental variables regression). High school challenge is a percentile relative to all U.S. high schools, with higher numbers representing more disadvantaged communities. Schools are segmented by a fixed national Challenge cutoff: high-challenge schools have Challenge ≥ 67 (national percentile), and lower-challenge schools have Challenge < 67. Polluted days are those with $PM_{2.5}$ of at least 8 $\mu g/m^3$. These are also segmented into pollution days during the previous school year versus summer. Pollution days in the previous year are also segmented into bins: moderately polluted ($8 \leq PM_{2.5} < 12$ $\mu g/m^3$) and heavily polluted ($PM_{2.5} \geq 12$ $\mu g/m^3$) days. Smoke pollution days are defined using smoke-attributable $PM_{2.5}$ and classified into light, moderate, and heavy categories using the same $PM_{2.5}$ thresholds as the total pollution measure.

Table 3: OLS - Impact of Prior Year Pollution on Test Scores

	(1)	(2)	(3)	(4)
(A) Overall pollution				
Pollution days	-0.006	-0.022	-0.017	-0.017
	(0.016)	(0.016)	(0.016)	(0.016)
(B) By pollution intensity				
Moderate pollution days	-0.053**	-0.057**	-0.067***	-0.065***
	(0.022)	(0.022)	(0.022)	(0.022)
Heavy pollution days	0.029	0.006	0.023	0.021
	(0.022)	(0.022)	(0.022)	(0.022)
N	16,742,701	16,742,701	16,742,701	16,742,701
Prior year weather	No	Yes	Yes	Yes
Exam day pollution	No	Yes	Yes	Yes
Exam day weather	No	Yes	Yes	Yes
Division-specific trends	No	No	Yes	Yes
County unemployment rate	No	No	No	Yes

Notes: Heteroskedasticity robust standard errors clustered by student high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Coefficients in each column and panel come from a regression of thousandths of a standard deviation in PSAT (math plus reading) z-scores on the pollution measures shown. Pollution is measured as days in designated bin ranges in the 365 days preceding the PSAT take. In Panel A, polluted days are those with $PM_{2.5}$ of at least 8 $\mu g/m^3$. Panel B distinguishes between moderately polluted days ($8 \leq PM_{2.5} < 12$ $\mu g/m^3$) and heavily polluted days ($PM_{2.5} \geq 12$ $\mu g/m^3$). All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. Column 2 adds controls for prior year temperature and precipitation, test day pollution levels, and test day temperature and precipitation. Column 3 additionally adds 9 Census division-specific linear time trends. Column 4 additionally adds controls for the county-level unemployment rate in the PSAT take year. The sample comprises all students from the high school classes of 2007–16 who took the PSAT more than once and had available pollution and smoke data (the sample that could be analyzed using instrumental variables).

Table 4: Instrumental Variables - Impact of Prior Year Smoke Pollution on Prior Year Pollution (First Stage)

	(1)	(2)	(3)	(4)
(A) First stage: Overall pollution				
Light smoke pollution days	0.211***	0.310***	0.312***	0.312***
	(0.020)	(0.021)	(0.021)	(0.021)
Moderate smoke pollution days	1.862***	2.047***	2.027***	2.036***
	(0.065)	(0.065)	(0.065)	(0.066)
Heavy smoke pollution days	0.285***	0.347***	0.360***	0.350***
	(0.055)	(0.054)	(0.054)	(0.054)
F-statistic:	438.09	564.42	541.96	542.55
(B) First stage: By pollution intensity				
Moderate pollution days				
Light smoke pollution days	0.412***	0.456***	0.459***	0.460***
	(0.014)	(0.015)	(0.015)	(0.014)
Moderate smoke pollution days	0.353***	0.403***	0.408***	0.385***
	(0.054)	(0.056)	(0.056)	(0.056)
Heavy smoke pollution days	-1.180***	-1.138***	-1.111***	-1.089***
	(0.047)	(0.047)	(0.047)	(0.047)
Heavy pollution days				
Light smoke pollution days	-0.201***	-0.146***	-0.146***	-0.148***
	(0.016)	(0.017)	(0.017)	(0.016)
Moderate smoke pollution days	1.510***	1.644***	1.619***	1.651***
	(0.057)	(0.055)	(0.055)	(0.055)
Heavy smoke pollution days	1.465***	1.485***	1.471***	1.439***
	(0.050)	(0.050)	(0.050)	(0.049)
F-statistic:	331.62	384.52	384.24	384.73
Prior year temperature and precipitation	No	Yes	Yes	Yes
Exam day pollution	No	Yes	Yes	Yes
Exam day temperature and precipitation	No	Yes	Yes	Yes
Division-specific linear trends	No	No	Yes	Yes
County-level unemployment rate	No	No	No	Yes

Notes: Heteroskedasticity robust standard errors clustered by student high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Panel A reports first-stage estimates for overall pollution days. Panel B reports first-stage estimates for moderate and heavy pollution days. Coefficients in each column and panel come from the first-stage regression of the pollution measure shown on the excluded smoke pollution instruments, controlling for the covariates and fixed effects included in the corresponding second-stage specification. Pollution and smoke-attributable pollution are each measured as days in designated bin ranges in the 365 days preceding the PSAT take. The reference category for the excluded instruments is the number of days with no smoke or 0 $PM_{2.5}$ from smoke. Light smoke days are days with smoke $0 < PM_{2.5} < 8$. Moderate smoke days are days with smoke $8 \leq PM_{2.5} < 12$. Heavy smoke days are days with smoke $PM_{2.5} \geq 12\ \mu g/m^3$. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. Column 2 adds controls for prior year temperature and precipitation, test day pollution levels, and test day temperature and precipitation. Column 3 adds 9 Census division-specific linear time trends. Column 4 adds controls for the county-level unemployment rate in the PSAT take year. Attempt-student rows for schools $\leq$ 10 km from a fire in a given year were excluded. The sample comprises all students from the high school classes of 2007–16 who took the PSAT more than once and had available pollution and smoke data.

Table 5: Instrumental Variables - Impact of Prior Year Smoke Pollution on PSAT Scores (Second Stage)

	(1)	(2)	(3)	(4)
(A) Second Stage: Overall pollution				
Pollution days	-0.276***	-0.245***	-0.208***	-0.200***
	(0.053)	(0.052)	(0.051)	(0.051)
(B) Second stage: By pollution intensity				
Moderate pollution days	-0.482***	-0.498***	-0.508***	-0.506***
	(0.078)	(0.077)	(0.077)	(0.077)
Heavy pollution days	-0.174***	-0.106*	-0.022	-0.025
	(0.054)	(0.055)	(0.054)	(0.054)
N	16,742,701	16,742,701	16,742,701	16,742,701
Prior year temperature and precipitation	No	Yes	Yes	Yes
Exam day pollution	No	Yes	Yes	Yes
Exam day temperature and precipitation	No	Yes	Yes	Yes
Division-specific linear trends	No	No	Yes	Yes
County-level unemployment rate	No	No	No	Yes

Notes: Heteroskedasticity robust standard errors clustered by student high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Coefficients in each column and panel come from the second stage of the instrumental variables regression, where pollution is instrumented using smoke-attributable pollution. Pollution is measured as days in designated bin ranges in the 365 days preceding the PSAT take. In Panel A, smoke-attributable pollution instruments for overall polluted days. In Panel B, smoke-attributable pollution instruments separately for moderately polluted days ($8 \leq$ PM$_{2.5}$ <12 $\mu g/m^3$) and heavily polluted days (PM$_{2.5}$ ≥ 12 $\mu g/m^3$). The omitted first-stage category is days with no smoke contribution. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number. Column 2 adds controls for prior year temperature and precipitation, test day pollution levels, and test day temperature and precipitation. Column 3 additionally adds 9 Census division-specific linear time trends. Column 4 additionally adds controls for the county-level unemployment rate in the PSAT take year. Attempt-student rows for schools $\leq$10 km from a fire in a given year were excluded. The sample comprises all students from the high school classes of 2007–16 who took the PSAT more than once and had available pollution and smoke data.

Table 6: Instrumental Variables: Lagged Full-Year Pollution Exposure and PSAT Scores

	(1)	(2)
Moderate $PM_{2.5}$ days	-0.311***	-0.388***
	(0.075)	(0.072)
Heavy $PM_{2.5}$ days	-0.010	-0.022
	(0.058)	(0.055)
Moderate $PM_{2.5}$ days (1-year lag)		-0.117
		(0.082)
Heavy $PM_{2.5}$ days (1-year lag)		-0.034
		(0.056)
N	15,114,549	15,114,549

Notes: Heteroskedasticity robust standard errors clustered by student high school are in parentheses (* p<.10 ** p<.05 *** p<.01). All columns report instrumental variables estimates for full-year pollution intensity bins. Coefficients in each column and panel come from a regression of thousandths of a standard deviation in PSAT (math plus reading) z-scores on the pollution measures shown, where pollution is instrumented by pollution from smoke. Column 1 reproduces the baseline IV specification on the subset of observations eligible for the lag-1 analysis. Column 2 additionally includes one lag of full-year moderate and heavy pollution exposure in levels, instrumented using the corresponding lagged smoke pollution bins. All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number, as well as controls for current and lagged prior-year temperature and precipitation, test day pollution levels, test day temperature and precipitation, and Census division-specific linear time trends. Pollution is measured as days in designated bin ranges in the 365 days preceding the PSAT take. Moderate pollution days correspond to $8 \leq PM_{2.5} < 12$ $\mu g/m^3$, and heavy pollution days correspond to $PM_{2.5} \geq 12$ $\mu g/m^3$. Attempt-student rows for schools $\leq$ 10 km from a fire in a given year were excluded. The sample comprises retaking students with available pollution, smoke, and weather data, restricted to the lag-1 eligible sample.

Table 7: Instrumental Variables - Impact of School Year and Summer Pollution

	(1)	(2)	(3)
(A) Overall pollution			
School year pollution days	-0.097		-0.186**
	(0.068)		(0.078)
Summer pollution days		-0.039	0.329
		(0.193)	(0.234)
(B) By pollution intensity			
School year moderate pollution days	-0.347***		-0.702**
	(0.105)		(0.325)
School year heavy pollution days	-0.032		-0.152
	(0.067)		(0.137)
Summer moderate pollution days		-2.228***	0.587
		(0.411)	(1.009)
Summer heavy pollution days		-0.603***	0.551
		(0.218)	(0.486)
N	16,742,701	16,742,701	16,742,701

Notes: Heteroskedasticity robust standard errors clustered by student high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Coefficients in each column and panel come from a regression of thousandths of a standard deviation in PSAT (math plus reading) z-scores on the pollution measure shown, where pollution is instrumented by pollution from smoke. Pollution is measured as days in the year preceding the PSAT take, separating the calendar into days during the school year and days during the summer break. In panel A, polluted days are those with $PM_{2.5}$ of at least 8 $\mu g/m^3$. Panel B distinguishes between moderately polluted days ($PM_{2.5}$ of 8-12 $\mu g/m^3$) and heavily polluted days ($PM_{2.5} \geq 12$ $\mu g/m^3$). The reference category for the instrument in the first-stage regression is the bin representing number of days with no smoke or 0 $PM_{2.5}$ from smoke. Light smoke days are days with smoke $0 < PM_{2.5} < 8$. In Panel A, smoke days are classified as days with smoke $PM_{2.5} \geq 8$. In Panel B, moderate and heavy smoke days are also distinguished. All regressions instrument pollution using pollution from smoke from distant fires and include student fixed effects and fixed effects for each combination of cohort, test date and take number, as well as controls for prior year temperature and precipitation, test day pollution levels, test day temperature and precipitation, and Census division-specific linear time trends. Attempt-student rows for schools ≤ 10 km from a fire in a given year were excluded. The sample comprises all students from the high school classes of 2007-16 who took the PSAT more than once and had available pollution and smoke pollution data.

Table 8: Heterogeneity by Disadvantaged Status

	School Challenge		Student Race	
	Bottom Two-Thirds (1–66)	Top Third (67–100)	Non-URM (White/Asian)	URM (Black/Hispanic)
(A) Overall pollution				
School year pollution days	-0.101	-0.302***	0.043	-0.311***
	(0.067)	(0.109)	(0.030)	(0.037)
(B) By pollution intensity				
School year moderate pollution days	-0.221**	-0.838***	0.027	-0.936***
	(0.102)	(0.230)	(0.047)	(0.068)
School year heavy pollution days	-0.020	-0.250**	0.053	-0.189***
	(0.080)	(0.111)	(0.037)	(0.038)
N	12,325,159	4,333,762	9,885,801	5,934,440

Notes: Heteroskedasticity robust standard errors clustered by student high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Coefficients in each column and panel come from a regression of thousandths of a standard deviation in PSAT (math plus reading) z-scores on the pollution measures shown. Regression is estimated using two-stage least squares instrumental variables. Columns 1–2 present estimates by school national challenge percentile group (top third versus bottom two-thirds). Columns 3–4 present estimates by student race group (non-URM versus URM). Columns 2 and 4 represent more disadvantaged students. Pollution is measured as days in the year preceding the PSAT take, counting only days during the school year. In Panel A, polluted days are those with $PM_{2.5}$ of at least 8 $\mu g/m^3$. Panel B distinguishes between moderately polluted days ($PM_{2.5}$ of 8–12 $\mu g/m^3$) and heavily polluted days ($PM_{2.5} \geq 12$ $\mu g/m^3$). All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number, as well as controls for prior year temperature and precipitation, test day pollution levels, test day temperature and precipitation, and Census division-specific linear time trends. The sample comprises all students from the high school classes of 2007–16 who took the PSAT more than once and had available pollution data (took the PSAT at least twice in exam years 2007-2016), and whose school in a given attempt-year was more than 10 km from the nearest fire. Challenge is a nationally normed percentile metric developed by the College Board to characterize how challenging it is for a given high school's students to access educational and socioeconomic opportunities. It incorporates data on college attendance, household structure, family income, and other measures drawn from national data sources, including the American Community Survey. Estimates in the left panel are presented by a fixed national cutoff: top-third schools have challenge national percentile scores ≥ 67, relative to all U.S. domestic high schools, and bottom two-thirds schools have challenge national percentile scores < 67. Students were grouped into two categories based on self-reported race. The non-URM category includes Asian and White students. The URM category includes Black and Hispanic students. Students with unknown or other races were excluded from Non-URM/URM categorization.

Appendix – For Online Publication

Table of Contents

A1 Method for Interpreting Magnitudes

A1.1 Student-Level Pollution Exposure and Scaling

This section explains how we construct the student-level pollution exposure measures used throughout the magnitude calculations. We characterize the level of exposure across students in the analytic sample.

A1.1.1 Average Exposure Across Students

We begin by restricting the data to observations included in the main estimation sample. The unit of observation in the underlying dataset is a student-by-exam record, so students may appear multiple times if they take the PSAT more than once. To obtain a student-level exposure measure, we collapse the data to one observation per student.

For each student, we compute the mean number of days with pollution in each exposure category across all of their observed exam years. Specifically, we calculate the average number of days per year in which particulate matter concentrations fall within the moderate range ($PM_{2.5}$ between 8 and 12 $\mu g/m^3$), the heavy range ($PM_{2.5}$ above 12 $\mu g/m^3$), and the combined set of polluted days ($PM_{2.5}$ at or above 8 $\mu g/m^3$). This produces a student-level average exposure profile that reflects the pollution conditions experienced over the period in which the student is observed.

After constructing these student-level averages, we summarize the distribution of exposure across students. We compute the mean of these student-level averages, which gives the expected number of polluted days experienced by the average student, with each student weighted equally. This differs from aggregating at the observation level, where students with more test attempts would receive greater weight.

These summary statistics serve as inputs for subsequent magnitude calculations, where estimated pollution effects are scaled by the typical exposure experienced by students. By working at the student level and weighting each student equally, this approach ensures that magnitude interpretations correspond to the experience of the average student in the analytic sample.

A1.2 Average-Student Effects of Pollution Exposure

This section computes the effect of pollution exposure for a representative student by scaling the estimated per-day pollution effects using the exposure profile of the average student. These quantities serve as the baseline magnitudes used in subsequent interpretation and counterfactual analyses.

A1.2.1 Total Achievement Cost for the Average Student

We first translate the estimated pollution effects into changes in test scores for a representative student. The goal is to quantify the total achievement cost associated with the level of pollution exposure experienced by the average student in the analytic sample.

To do so, we combine the estimated per-day effect of pollution on PSAT scores with the average number of polluted days experienced by students, as described in Appendix A1.1. This yields a model-implied reduction in test scores relative to a benchmark in which no polluted days occur. Throughout, we interpret this quantity as the total effect of exposure accumulated over the course of a year, rather than the effect of a marginal change in exposure.

We report estimates of this quantity under alternative specifications that differ in how pollution exposure is defined.

Baseline Scaling Using Student-Level Exposure We begin by estimating the total achievement cost of pollution for the average student using the binned exposure specification. This calculation combines the estimated per-day effects of moderate- and heavy-pollution days with the average number of days in each category experienced by students, as constructed in Appendix A1.1.

Specifically, let $\beta_{8\text{–}12}$ and $\beta_{\geq 12}$ denote the estimated effects of one additional day of moderate and heavy pollution, respectively. Let $\mathbb{E}[D_{8\text{–}12}]$ and $\mathbb{E}[D_{\geq 12}]$ denote the average number of days in each category across students, where each student is weighted equally. The total achievement cost for the average student is then given by:

$$\Delta_{\text{avg}} = \beta_{8\text{–}12}\,\mathbb{E}[D_{8\text{–}12}] + \beta_{\geq 12}\,\mathbb{E}[D_{\geq 12}].$$

We estimate this quantity using a nonlinear combination of the regression coefficients and the average exposure levels, which allows us to recover both the point estimate and its associated standard error and confidence interval. This approach accounts for uncertainty in the estimated pollution effects when constructing the scaled magnitude.

The resulting estimate represents the model-implied annual change in PSAT scores for the average student, measured relative to a benchmark in which no moderate- or heavy-pollution days occur. This interpretation relies on the linearity of the estimated relationship between pollution exposure and test scores, so that the total effect of exposure is the sum of the per-day effects across pollution categories.

A1.3 Interpreting Magnitudes in Intuitive Units

This section translates the estimated pollution effects into more interpretable units to facilitate economic interpretation. Building on the average-student effects constructed in Appendix A1.2, we express these magnitudes in standardized test-score units and relative to observed within-student score changes.

These transformations provide complementary ways of understanding the size of the estimated effects relative to the distribution of scores and typical patterns of learning, without altering the underlying empirical estimates.

A1.3.1 Standardized Units

We first interpret the estimated effects in terms of standardized test-score units and their implied position in the distribution of scores. These measures provide a distributional benchmark for assessing the magnitude of the estimated effects.

Effects in Standard Deviation Units

We first express the estimated effects in standardized test-score units. Because the PSAT outcome is scaled such that 1000 points correspond to one standard deviation, we convert point-scale effects into standard deviation units by dividing the estimated change in scores by 1000.

Let Δ denote the average-student scaled effect in points, as defined in Appendix A1.2. The corresponding effect in standard deviation units is given by:

$$\Delta_{\mathrm{SD}} = \frac{\Delta}{1000}.$$

Confidence intervals are transformed analogously by applying the same scaling to the lower and upper bounds.

We apply this transformation both to the baseline estimate based on the binned specification and to the alternative estimate based on the 8+ specification. This provides a standardized measure of effect size that is directly comparable across specifications and facilitates interpretation relative to the distribution of test scores.

A1.3.2 Comparison to Observed Score Changes

As an additional benchmark, we compare the magnitude of the estimated pollution effects to observed within-student changes in PSAT scores across test attempts. This provides a reference point based on actual score variation in the data, rather than a constructed transformation.

To do so, we compute within-student changes in test scores between consecutive PSAT attempts. For each student, we calculate the difference in scores between the first and second attempts, as well as additional differences for students with more than two attempts. We then summarize the distribution of these changes across students, including the mean, median, standard deviation, and the mean absolute change in scores.

Using these quantities, we construct two comparison metrics. First, we compute the ratio of the estimated effect from Appendix A1.2 to the mean within-student score change between the first and second attempts. This yields a measure of the extent to which the estimated pollution effect corresponds to typical observed changes in scores. Second, to abstract from the direction of score changes, we compare the absolute value of the estimated effect to the mean absolute within-student change. This provides a magnitude-based benchmark that is not affected by the sign of average score changes.

These comparisons situate the estimated pollution effect relative to the distribution of actual score movements observed in the data, offering an empirical benchmark for interpreting the size of the effect.

A1.4 Decomposition of the Black/Hispanic–White/Asian Achievement Gap

We assess pollution's contribution to the Black/Hispanic–White/Asian achievement gap using a heterogeneous-slope decomposition that allows the estimated pollution effect to differ across groups, paralleling the treatment of school challenge in Section A1.5. Both exposure and the pollution coefficient may vary by race.

To do so, we estimate the baseline binned specification separately for Black and Hispanic students and for White and Asian students. Let $\beta^U_{8\text{–}12}$ and $\beta^U_{\geq 12}$ denote the estimated effects of moderate- and heavy-pollution days for Black and Hispanic (URM) students, and let $\beta^N_{8\text{–}12}$ and $\beta^N_{\geq 12}$ denote the corresponding effects for White and Asian (non-URM) students. Let $\overline{D}^U_{8\text{–}12}$ and $\overline{D}^U_{\geq 12}$ denote the mean number of moderate- and heavy-pollution days experienced by Black and Hispanic students, and define the analogous quantities for White and Asian students.

Using these group-specific coefficients and exposure means, we compute a group-specific "achievement cost" for each racial group:

$$\text{Cost}_U = \beta^U_{8\text{–}12}\,\overline{D}^U_{8\text{–}12} + \beta^U_{\geq 12}\,\overline{D}^U_{\geq 12},$$
$$\text{Cost}_N = \beta^N_{8\text{–}12}\,\overline{D}^N_{8\text{–}12} + \beta^N_{\geq 12}\,\overline{D}^N_{\geq 12}.$$

The heterogeneous-slope contribution of pollution to the Black/Hispanic–White/Asian gap is then defined as the difference between these two group-specific achievement costs:

$$\text{Contribution}^{\text{het}}_{UN} = \text{Cost}_U - \text{Cost}_N.$$

We express this quantity relative to the observed racial gap by dividing by the difference in mean first-attempt PSAT scores between the two groups and multiplying by 100. This exercise allows pollution exposure to have different effects across groups, so the resulting object reflects both differences in exposure and differences in the estimated pollution–score slope. It is the racial-gap analog of the heterogeneous-slope challenge decomposition and is the calculation behind the main-text estimate of pollution's contribution to the Black/Hispanic–White/Asian gap.

This exercise should be interpreted as an accounting decomposition under heterogeneous slopes, not as a causal estimate of the effects of race or of the socioeconomic conditions correlated with it. Because the contribution is a nonlinear function of coefficients estimated in separate samples, formal inference for this object generally requires a bootstrap procedure.

A1.5 Decomposition of the High–Low Challenge Achievement Gap

We assess pollution's contribution to the high–low challenge achievement gap using a heterogeneous-slope decomposition that allows the estimated pollution effect to differ across groups. Both exposure and the pollution coefficient may vary by school challenge level.

To do so, we estimate the baseline binned specification separately for students in low-challenge schools and for students in high-challenge schools. Let $\beta^L_{8\text{–}12}$ and $\beta^L_{\geq 12}$ denote the estimated effects of moderate- and heavy-pollution days in low-challenge schools, and let $\beta^H_{8\text{–}12}$ and $\beta^H_{\geq 12}$ denote the corresponding effects in high-challenge schools. Let $\overline{D}^L_{8\text{–}12}$ and $\overline{D}^L_{\geq 12}$ denote the mean number of moderate- and heavy-pollution days experienced by students in low-challenge schools, and define the analogous quantities for high-challenge schools.

Using these group-specific coefficients and exposure means, we compute a group-specific "achievement cost" for each challenge category:

$$\text{Cost}_L = \beta^L_{8\text{–}12}\,\overline{D}^L_{8\text{–}12} + \beta^L_{\geq 12}\,\overline{D}^L_{\geq 12},$$

$$\text{Cost}_H = \beta^H_{8\text{–}12}\,\overline{D}^H_{8\text{–}12} + \beta^H_{\geq 12}\,\overline{D}^H_{\geq 12}.$$

The heterogeneous-slope contribution of pollution to the high–low challenge gap is then defined as the difference between these two group-specific achievement costs:

$$\text{Contribution}^{\text{het}}_{HL} = \text{Cost}_H - \text{Cost}_L.$$

We express this quantity relative to the observed high–low challenge gap by dividing by the difference in mean first-attempt PSAT scores and multiplying by 100. This exercise allows pollution exposure to have different effects across groups, so the resulting object reflects both differences in exposure and differences in the estimated pollution-score slope.

This exercise should be interpreted as an accounting decomposition under heterogeneous slopes, not as a causal estimate of the effects of changing school assignment or socioeconomic conditions.

A1.6 Changes in Achievement Associated with Declining Pollution Exposure

A1.6.1 Baseline Accounting Using the 8+ Specification

We next assess how changes in pollution exposure over time translate into changes in student achievement. Specifically, we construct a descriptive accounting measure of how much higher test scores would be expected to be in recent years relative to earlier years, given observed declines in pollution exposure.

To do so, we use the 8+ pollution specification, which defines exposure as the number of days with $PM_{2.5}$ at or above 8 μg/m^3. Let β_{8+} denote the estimated per-day effect of polluted days, and let $\mathbb{E}[D^{\text{now}}_{8+}]$ and $\mathbb{E}[D^{\text{past}}_{8+}]$ denote the mean number of polluted days in a recent year and in a comparison year several years earlier, respectively.

We define the change in exposure as:

$$\Delta D_{8+} = \mathbb{E}[D^{\text{now}}_{8+}] - \mathbb{E}[D^{\text{past}}_{8+}].$$

The implied change in test scores associated with this change in exposure is then given by:

$$\Delta Y = -\beta_{8+} \cdot \Delta D_{8+}.$$

In practice, we implement this calculation by comparing the most recent year in the sample to the closest available year approximately nine years earlier, reflecting the time span over which IV-based estimates are available. We compute mean exposure in each year and scale the difference by the estimated per-day effect.

This quantity represents the model-implied change in PSAT scores attributable to the observed decline in polluted days over this period, holding the estimated pollution–score relationship fixed.

It should be interpreted as a descriptive accounting exercise rather than a causal estimate of how changes in pollution over time affect achievement, as it does not account for other factors that may have evolved concurrently.

A2 Appendix Exhibits

A2.1 Appendix Tables

Table A1: Other papers that examine learning impacts of $PM_{2.5}$ exposure

Citation	Empirical approach	Results	N
Gilraine and Zheng (2022)	IV with Coal plant closures & Shift share	"We find that each one-unit increase in particulate pollution reduces test scores by 0.02 standard deviations."	700K
Martin et al. (2025)	Downwind of BRT lines (IV)	"$1\mu g/m^3$ additional $PM_{2.5}$ throughout the academic year decreases math scores by 0.0040 standard deviations, this is equivalent to 6% of the mean"	300K
Dang et al. (2025)	Wind direction IV by month	"However, air pollution concentration during the past 12 months does not have significant impacts on the test scores." But they do find adverse impacts of pollution in the 1 month and 3 months prior. They use a linear function with average $PM_{2.5}$.	40K
Wen and Burke (2022)	Cumulative $PM_{2.5}$ attributable to smoke, no IV	"An additional $10\mu g/m^3$ of cumulative smoke $PM_{2.5}$ on school days decreases average test scores by 0.044% of a standard deviation."	300K

A2.2 Appendix Figures

Table A2: IV - Robustness to Alternative Fire Exclusion Distances

	Fire Exclusion Distance		
	10 km	20 km	30 km
(A) Overall pollution			
Pollution days	-0.208***	-0.124**	-0.143***
	(0.051)	(0.053)	(0.054)
(B) By pollution intensity			
Moderate pollution days	-0.508***	-0.375***	-0.392***
	(0.077)	(0.073)	(0.071)
Heavy pollution days	-0.022	0.100	0.129*
	(0.054)	(0.062)	(0.070)
N	16,742,701	15,026,981	13,559,580

Notes: Heteroskedasticity robust standard errors clustered by student high school are in parentheses (* p<.10 ** p<.05 *** p<.01). Coefficients in each column and panel come from a two-stage least squares instrumental variables regression of thousandths of a standard deviation in PSAT (math plus reading) z-scores on the pollution measures shown, where pollution is instrumented by pollution from smoke. Pollution is measured as days in designated bin ranges in the 365 days preceding the PSAT take. In Panel A, polluted days are those with $PM_{2.5}$ of at least 8 $\mu g/m^3$. Panel B distinguishes between moderately polluted days ($8 \leq PM_{2.5} < 12$ $\mu g/m^3$) and heavily polluted days ($PM_{2.5} \geq 12$ $\mu g/m^3$). All regressions include student fixed effects and fixed effects for each combination of cohort, test date and take number, as well as controls for prior year temperature and precipitation, test day pollution levels, and test day temperature and precipitation, and Census division-specific linear time trends. Column 1 reports the preferred specification used throughout the manuscript, excluding schools located within and including 10 km of a fire. Columns 2 and 3 increase the fire exclusion distance while holding the remaining specification fixed. Schools located within and including 20 km and 30 km of a fire are excluded, respectively. All geographies used for linear time trends reflect the location of the school the student last attended. The sample comprises all students from the high school classes of 2007–16 who took the PSAT more than once and had available pollution and smoke data.

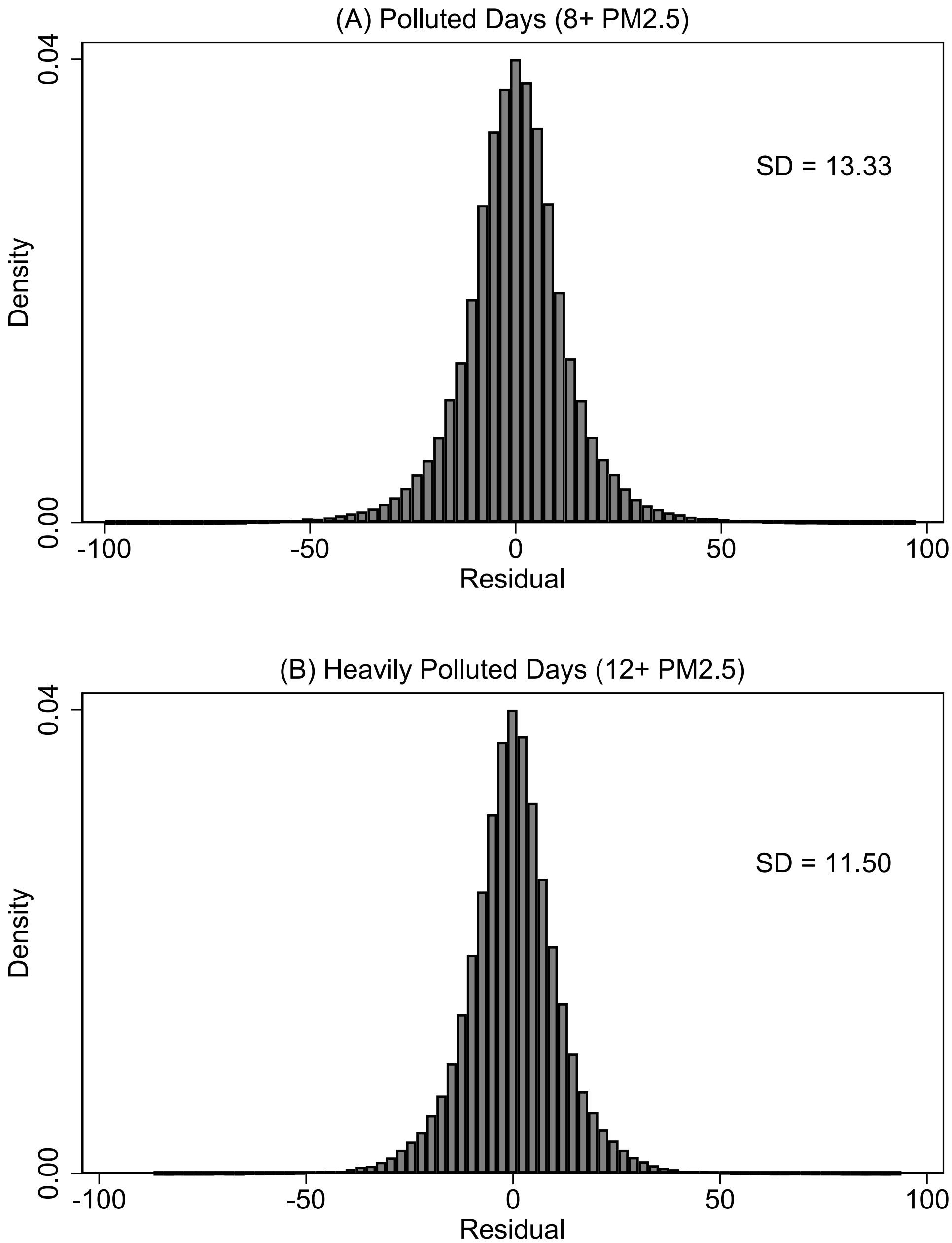


Figure A1: Residuals of Prior Year Pollution Days

Notes: The above figure shows the distribution of residuals resulting from regressions on student fixed effects of the number of polluted days (panel A) and number of heavily polluted days (panel B) experienced by students at their school location in the 365 days prior to taking the PSAT. All regressions include fixed effects for each combination of cohort, test date and take number. The figure excludes residuals with magnitude above 100 (panel A). The standard deviation of the full set of residuals is shown in each panel.

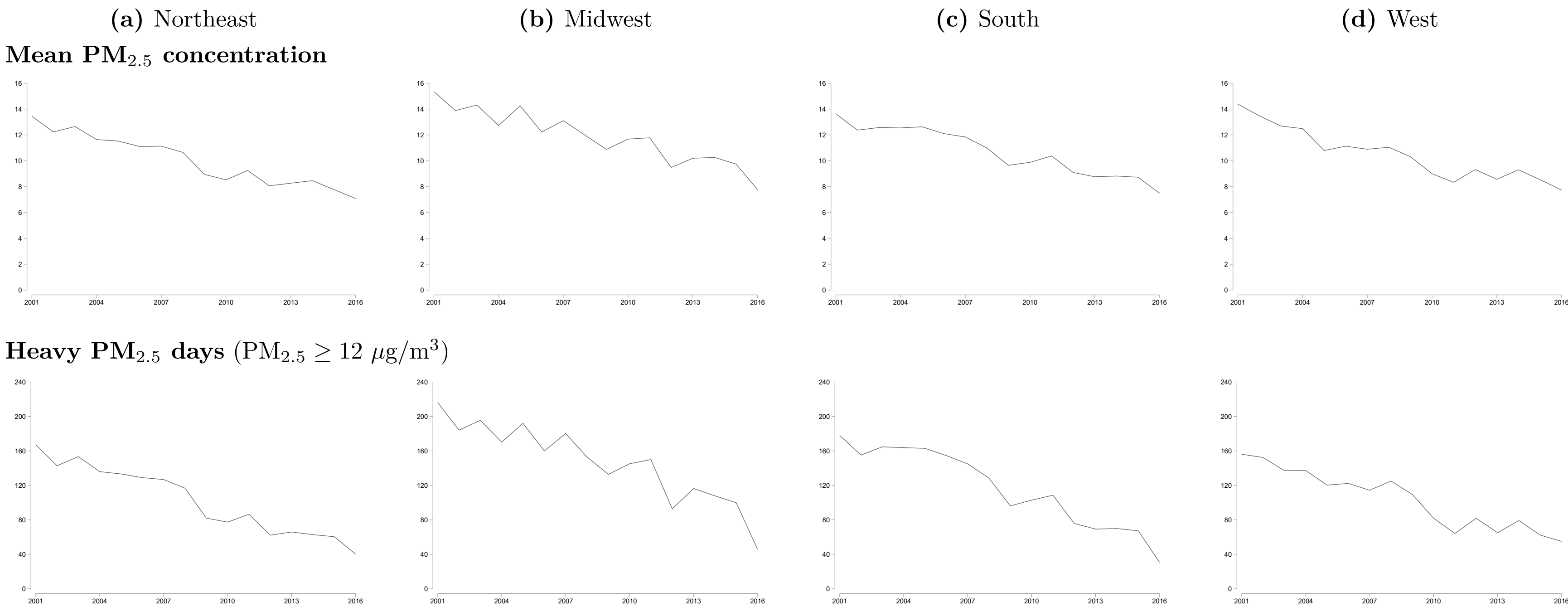


Figure A2: Regional trends in $PM_{2.5}$ exposure

Notes: The top row plots student-weighted mean $PM_{2.5}$ concentrations in the year prior to the PSAT exam. The bottom row plots the student-weighted mean number of heavy $PM_{2.5}$ days (days with $PM_{2.5} \geq 12\ \mu g/m^3$) in the prior year. Exposure measures are aggregated to Census region by exam year and weighted by the number of PSAT takers in each school–year. The sample includes exam years 2001–2016.