nature sustainability

ARTICLES

https://doi.org/10.1038/s41893-022-00956-y

Lower test scores from wildfire smoke exposure

Jeff Wen[1] ✉ and Marshall Burke[1,2,3]

Wildfires have increased in frequency and severity over the past two decades, threatening to undo substantial air quality improvements. We investigate the relationship between wildfire smoke exposure and learning outcomes across the United States using standardized test scores from 2009–2016 for nearly 11,700 school districts and satellite-derived estimates of daily smoke exposure. Relative to a school year with no smoke, average cumulative smoke-attributable $PM_{2.5}$ (surface particulate matter <2.5 µg m^{-3}) exposure during the school year (~35 µg m^{-3}) reduces test scores by ~0.15% of a standard deviation. These impacts are more pronounced among younger students and are observed across differing levels of economic disadvantage and racial/ethnic composition. Additionally, we project that smoke $PM_{2.5}$ exposure in 2016 reduced discounted future earnings by nearly $1.7 billion ($111 per student). Roughly 80% of these costs are borne by disadvantaged districts. Our findings quantify a previously unaccounted for social cost of wildfire that is likely to worsen under a warming climate.

The frequency and severity of wildfires throughout the western United States have increased in recent decades and are expected to worsen as the climate continues to warm[1]. Literature has linked these wildfire events and the smoke they generate to a variety of social and economic impacts, in particular health and infrastructure related damages[2–4]. Yet emerging evidence from studies on non-wildfire air pollutants suggests that wildfire smoke exposure could have much wider-ranging impacts, including possible negative effects on human cognition[5–20]. Such effects would have broader implications for human capital formation and longer-term economic well-being, as well as for the social costs of a warming climate, but have not been documented in existing wildfire literature.

Recent studies have focused on the biological channels through which air pollution exposure might affect human health and have found that non-wildfire-related air pollution exposure is associated with higher likelihood of neuroinflammation[5,6,21] and increased risks for Alzheimer's, dementia and Parkinson's disease[22,23]. Epidemiological and social science studies have begun to draw links between air pollution exposure and cognitive performance on real-world tasks, including declining performance in chess tournaments[8], stock trading[9], call centre productivity[10], umpire decisions[11], cognitive assessments[15,17–19] and online brain games[12]. Similar to our setting, a handful of studies have assessed how student test scores have responded to variation in exposure to ambient (non-wildfire) air pollution[7,13,16,20,24,25]. A recent study investigates the association between long-term ambient air pollution exposure and student test performance in the United States and finds negative impacts of increased ambient air pollution[24]. Other studies found that short-term changes in air pollution on the day of the exam led to declines in student performance[7,13,16,20] and decreased future earnings[13]. While these studies focus on the impact of air pollution on student test performance, to our knowledge there are no studies that focus on wildfire smoke particulate matter, which recent studies suggest could potentially be more harmful to human health than other sources of particulate matter[4,26] and is likely the fastest growing source of air pollution in the United States[27].

As wildfire activity has dramatically increased in recent decades due to a rapidly warming climate and a century of fire suppression practices across the western United States, wildfire smoke has become an increasingly important contributor to surface particulate matter <2.5 µg m^{-3} ($PM_{2.5}$) concentrations[27]. Increasing wildfire-derived $PM_{2.5}$ threatens to undermine decades of progress in improving overall $PM_{2.5}$ concentrations—improvements brought about by changes in manufacturing practices, energy production and legislation[27–30]. Furthermore, while exposure to ambient smoke-derived $PM_{2.5}$ appears more evenly distributed across economic and racial/ethnic groups than other sources of $PM_{2.5}$[27], similar ambient exposures may differentially impact communities due to a variety of factors including differences in housing or school characteristics[31,32] or differences in knowledge of or ability to undertake protective behaviours. Ultimately, the differences in realized exposures could result in differential impacts across racial/ethnic and socioeconomic groups, as has now been documented for other environmental exposures[33–38].

Here we quantify the impact of wildfire smoke exposure on learning outcomes across the United States, as measured by standardized test scores, and estimate potential heterogeneous impacts of this exposure across demographic and socioeconomic groups. We first develop estimates of local-level wildfire-smoke-attributable $PM_{2.5}$ exposures across the United States and over time, using a combination of high-resolution predicted $PM_{2.5}$ data and satellite-derived wildfire smoke plumes[39,40] (Methods). We then study the effect of cumulative smoke exposure during the school year on student learning outcomes, as measured through harmonized national test score data for students from 3rd–8th grades collected across nearly ~11,700 school districts between 2009–2016. These comprehensive longitudinal data allow us to plausibly isolate the effect of wildfire-smoke-attributable $PM_{2.5}$ on student learning outcomes.

We model the effect of smoke exposure on student test performance using fixed-effects regression models that account for unobserved time-invariant differences in smoke exposure and test scores across districts as well as time-trending year-grade-specific differences common to all locations (Methods). As there has been an upward trend in both wildfire smoke exposure and test performance across our study period as well as large regional differences in average smoke exposure (Fig. 1), simple cross-sectional or time-series regressions could conflate overall trends or average differences in smoke with other factors that affect learning outcomes. Rather than comparing across districts, our approach compares

[1]Department of Earth System Science, Stanford University, Stanford, CA, USA. [2]Center on Food Security and the Environment, Stanford University, Stanford, CA, USA. [3]National Bureau of Economic Research, Cambridge, MA, USA. ✉e-mail: jlwen@stanford.edu

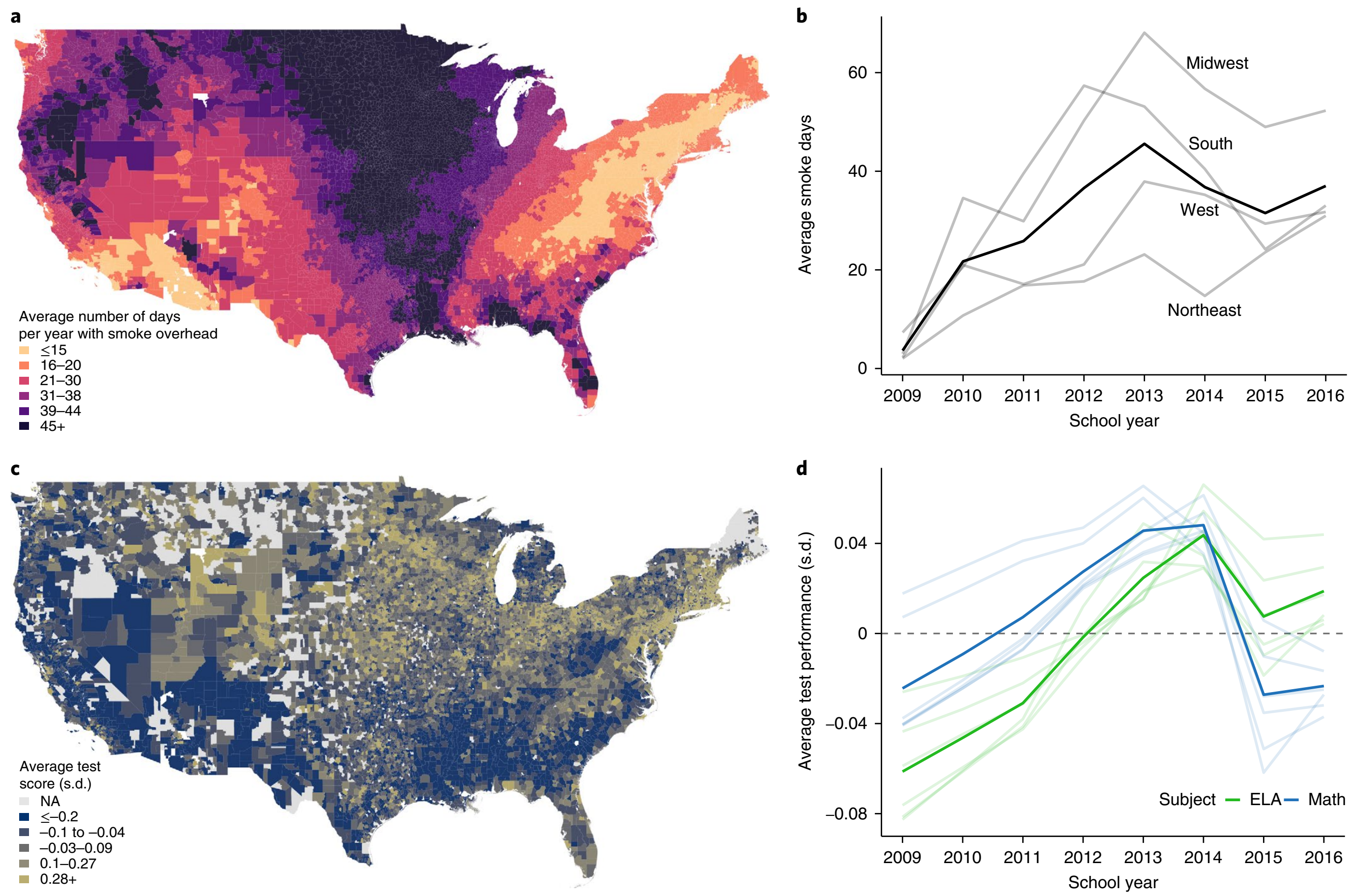

Fig. 1 | Spatiotemporal variation in wildfire smoke exposure and average test scores. **a**, Spatial distribution of the average number of days with smoke overhead from 2009–2016 for school districts[58] across the continental United States. **b**, Temporal variation in the average number of smoke days for various census regions. Black line represents the average over the entire United States. **c**, Spatial distribution of test scores. Math and ELA scores are averaged across the study period from 2009–2016 for each district[58] and are represented in standard deviations. **d**, Average test performance relative to the national reference cohort. Each state's standardized test results are scaled to the nationally comparable (National Assessment of Educational Progress) test. Faded lines represent grade-specific performance and darker lines represent the average over all grades.

particular districts to themselves over time as smoke exposure fluctuates, after accounting for differences in grade-specific national averages between years. We control flexibly for other factors that may be correlated with wildfire activity and affect student test performance (Methods), such as temperature and precipitation[33,34,41]. For our results to have a causal interpretation, there must be no additional unobserved factors that vary over time within districts (with different trends across districts) and are correlated with both district-specific wildfire smoke variation and local variation in test scores but are uncorrelated with district-specific weather variation; given the randomness in both where and when fires start and where wind blows smoke on a given day, we believe the existence of such a factor to be unlikely.

We then examine the heterogeneous impacts of wildfire smoke exposure by estimating whether responses differ between school and non-school days or by student age groups, levels of economic disadvantage, and/or race and ethnicity—dimensions along which earlier research has suggested environmental exposures and impacts might differ. Finally, to quantify the economic magnitude of smoke-related impacts, we explore how learning outcomes differ between a less severe compared to a more severe smoke year and provide estimates of the impact of wildfire-smoke-attributable $PM_{2.5}$ in terms of students' lost future earnings, using literature-derived estimates of the relationship between test scores and earnings (Methods).

Results

We find that smoke exposure in the year leading up to the test negatively affects test scores (Fig. 2). An additional $10\,\mu g\,m^{-3}$ of cumulative smoke $PM_{2.5}$ in the year leading up to the exam decreases average test scores by 0.029% (95% confidence interval (CI): −0.045% to −0.013%) of a standard deviation. The effect is similar across subjects, with decreases in English language arts (ELA) scores of 0.035% (95% CI: −0.052% to −0.017%) and math scores of 0.024% (95% CI: −0.041% to −0.006%) of a standard deviation for school districts across the United States from 2009–2016 (Fig. 2b). These results are robust to flexible functional forms such as higher-order polynomials of the smoke $PM_{2.5}$ response relationship and are fairly linear (Fig. 2a). Random-effects estimates are qualitatively similar to our preferred fixed-effects estimates, although slightly smaller in magnitude (Supplementary Fig. 6).

Comparing school day vs non-school day exposure, we find that smoke exposure on school days has a statistically significant negative effect on test performance where an additional $10\,\mu g\,m^{-3}$ of cumulative smoke $PM_{2.5}$ on school days decreases average test scores by 0.044% (95% CI: −0.082% to −0.0006%) of a standard deviation. Exposure on non-school days results in a smaller statistically significant negative effect, although the estimates are not statistically distinguishable from one another (Wald test for equivalence of coefficients: $F_{1,5092} = 0.614$, $P = 0.433$). We focus on cumulative

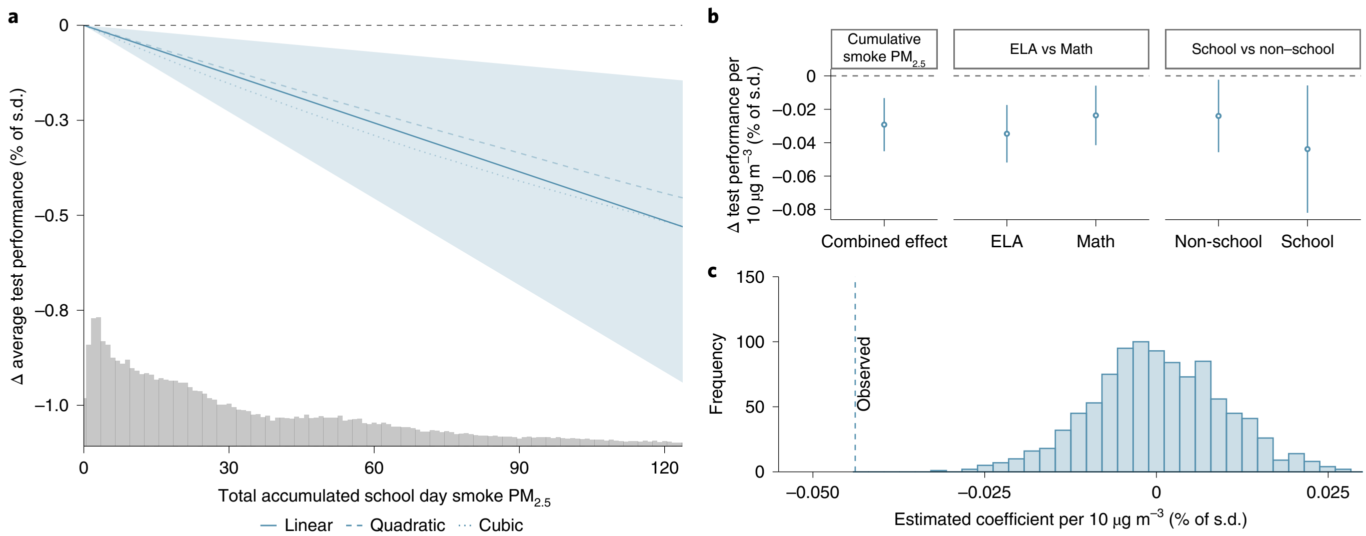

Fig. 2 | Effect of wildfire smoke exposure on student test scores. a, Test performance declines as a function of total accumulated daily smoke $PM_{2.5}$ during the school year before the test (only on school days). The line shows the regression point estimate of additional smoke $PM_{2.5}$ (μg m^{-3}) and the shaded area shows the bootstrapped 5th–95th percentile confidence intervals of 1,000 bootstrap samples. ΔTest performance is the change in test score relative to the national NAEP reference cohort, measured in percent of a standard deviation. **b**, Effect estimates of an additional 10 μg m^{-3} of smoke $PM_{2.5}$ in the year before the exam for school versus non-school day exposure, the combined average effect, and for ELA and Math. The circle markers represent the regression point estimates and the error bars show the 95% confidence intervals. **c**, Randomization inference test (1,000 permutations) showing the estimated effect size of an additional 10 μg m^{-3} of smoke $PM_{2.5}$ on school days when smoke $PM_{2.5}$ is randomly permuted across districts within each county. The observed effects are significantly different from the randomization test effects. For **a**, **b** and **c**, samples consist of $n = 438{,}613$ observations with 11,639 districts.

smoke $PM_{2.5}$ exposure on school days in the year before the exam as our measurement of exposure is at school locations and exposure on non-school days is uncertain.

While our main analysis clusters standard errors at the county level to account for correlation in errors across districts within the same county[42], we conduct additional analysis using a randomization inference approach to test the sharp null hypothesis of no effect for additional smoke $PM_{2.5}$ exposure on school days by randomly permuting test scores across districts within a county. This approach non-parametrically estimates statistical significance and is beneficial in the presence of fuzzy clustering where smoke exposure may be imperfectly correlated within the cluster[43,44]. We find that the estimated effect of school day smoke $PM_{2.5}$ exposure is significantly different from the distribution of permuted effect estimates (Fig. 2c). To additionally test whether regional time-trending unobservables could be driving our observed relationship between smoke exposure and test scores, we conduct an additional randomization inference test where we randomly shuffle district-level smoke $PM_{2.5}$ time series between districts within the same state; this is a demanding test, given the large within-state correlation in smoke exposures over time. Nevertheless, estimates when smoke time series are correctly matched to districts are in the tail (6th percentile) of the placebo treatment effect distribution (Supplementary Fig. 5), suggesting that state-level time-trending unobservables are unlikely driving our observed results (Supplementary Fig. 5).

Without access to data on evacuation orders, we test whether the identified effects are driven by smoke $PM_{2.5}$ exposure or by direct wildfire effects by dropping districts that are certain distances from the nearest fire perimeter provided by the National Interagency Fire Center in that year, then running a similar regression as the main specification (Methods). We find that the identified effect estimates remain fairly stable up to 6.2 miles (10 km), which provides evidence that the effects identified are probably not driven by direct wildfire effects but rather by the smoke $PM_{2.5}$ impacts (Supplementary Fig. 3).

We also test whether the effects of smoke exposure during previous school years carry over into test performance in the current year. While results are somewhat noisy, point estimates suggest that learning impacts can persist into future years (Supplementary Table 4).

In line with previous studies that find negative effects of air pollution exposure on younger children[7,45,46], we find that among primary school children, an additional 10 μg m^{-3} of cumulative smoke $PM_{2.5}$ on school days decreases scores by 0.122% (95% CI: −0.191% to −0.053%) of a standard deviation. However, these impacts are not apparent for secondary school students (6th–8th grade) (Fig. 3).

Consistent with previous work[27,47], we find that exposure to ambient $PM_{2.5}$ from wildfire smoke is largely similar across racial/ethnic subgroups (Supplementary Table 1) and across different levels of economic disadvantage (Supplementary Table 2). However, similar ambient exposures could result in very different impacts across subgroups due to potential differences in how pollutants infiltrate into indoor environments and/or differences in how increased wildfire smoke exposure interacts with baseline differences in other pollutant exposures or other determinants of learning outcomes. We thus explore differential responses across subgroups, with groups defined as being above or below the median value of each variable (Methods), to a given exposure across districts with varying levels of economic disadvantage and proportions of non-White students. We emphasize that the estimated moderating effect of economic disadvantage or racial/ethnic categories in this analysis should be understood to reflect the possible effect of racist and/or discriminatory policies or attitudes on outcomes, rather than as reflecting inherent characteristics of individuals or communities that fall into these categories.

We find that districts with high economic disadvantage and high proportion of non-White student population, as well as districts with low economic disadvantage and low proportion of non-White student population are more negatively affected by smoke $PM_{2.5}$ exposure compared with other subgroups (Fig. 3). For students

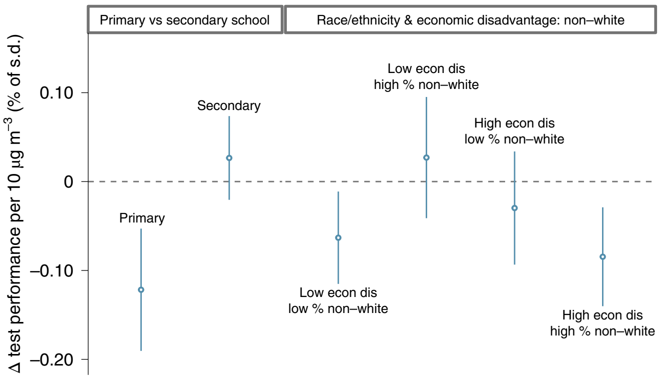

Fig. 3 | Heterogeneous effects of school day smoke $PM_{2.5}$ on test performance by grade, race/ethnicity and level of economic disadvantage. Left: effect estimate of an additional 10 μg m^{-3} of cumulative smoke $PM_{2.5}$ on school days for primary school (Grades 3–5) students and secondary school (Grades 6–8) students. Right: regression point estimates as circle markers and 95% confidence intervals for different intersecting levels of economic disadvantage and non-white racial/ethnic student population. In the primary vs secondary school analysis, the sample included $n = 438{,}613$ observations with 11,639 districts. In the race/ethnicity and economic disadvantage analysis, the sample had $n = 433{,}677$ observations with 11,623 districts due to some districts being without race/ethnicity and economic disadvantage information.

in districts with high economic disadvantage and a high proportion of non-White students, an additional 10 μg m^{-3} of cumulative smoke $PM_{2.5}$ on school days lowered test scores by 0.085% (95% CI: −0.140% to −0.029%) of a standard deviation. Districts with low economic disadvantage and low proportion of non-White students also appeared negatively impacted by additional smoke $PM_{2.5}$, with decreases of 0.063% (95% CI: −0.115% to −0.011%) of a standard deviation. When we stratify and run a separate regression for each racial/ethnic subgroup, we find that districts with a greater proportion of Asian, Black or Hispanic students exhibit responses to additional school day smoke $PM_{2.5}$ that are similar to each other but are different from districts with a greater proportion of White students (Supplementary Fig. 1).

To understand nationwide impacts of less severe versus more severe average smoke years on learning, we compared the 2011 versus 2016 school years, the former being the least smoky on average across districts and the latter being the most smoky year in our sample—albeit much less smoky than either 2018 or 2020, which are not in our sample. Taking into account heterogeneity in economic and racial/ethnic composition across school districts, we find substantial impacts of smoke exposure on learning across broad swaths of the United States (Fig. 4a). If smoke years continue to mirror the severity of 2011, we expect students to experience a decrease of 0.031% of a standard deviation in average test scores (median across districts), relative to a counterfactual of no smoke. However, if future wildfire events reflect severe smoke years, such as 2016, the median effect would be nearly an order of magnitude larger at 0.207% of a standard deviation decrease in average test scores.

As a rough estimate of the economic impact of cumulative smoke $PM_{2.5}$ exposure during the school year, we follow ref. [33] and calculate smoke impacts in terms of lost future earnings for students in our sample (Methods). We apply estimates from ref. [48] and estimate that district-average smoke $PM_{2.5}$ exposure led to a reduction in the net present value of lost future earnings of ~\$111 per student in 2016 compared with ~\$17 in 2011. The lost earnings of ~\$111 per student in 2016 totals nearly \$1.7 billion in potential lost future income from smoke $PM_{2.5}$ exposure when aggregating across all students in the United States. These impact estimates assume that increased future earnings due to teacher quality improvements are comparable to benefits of reducing smoke-attributable $PM_{2.5}$ in the classroom and could be overstated if teacher quality improvements result in other non-test performance related benefits that increase students' future earnings. However, impacts of this magnitude illustrate the potential benefits of reducing wildfire smoke $PM_{2.5}$ exposure.

When we consider the cumulative losses over all study years and across subgroups (Fig. 4b), we estimate the net present value of lost future income to be roughly \$544 million (95% CI: −\$999 million to −\$100 million) from smoke $PM_{2.5}$ exposure in 2016 for districts with low economic disadvantage and low proportion of non-White students. For districts with high economic disadvantage and high proportion of non-White students, we estimate cumulative impacts to be \$1.4 billion (95% CI: −\$2.3 billion to −\$477 million) from cumulative smoke $PM_{2.5}$ exposure in 2016. Thus, of the roughly \$1.7 billion in total costs during the smokiest year in our sample, 82% of the costs we estimate were borne by economically disadvantaged communities of colour. The larger total burden in these communities is a function of both the more negative effect size and the relatively larger total number of students who attend schools in economically disadvantaged communities of colour (around 50% of the exposed students in our sample). This suggests that additional increases in future wildfire smoke exposure due to climate change will likely disproportionately harm these communities.

Discussion

Our study quantifies the impact of wildfire-smoke-attributable $PM_{2.5}$ exposure, a rapidly growing source of particulate exposure throughout much of the United States[27], on student learning outcomes by leveraging a large sample of repeated observations. While test scores are an imperfect measure of student cognition, they are a common metric for evaluating student learning with relevance for long-term outcomes and opportunities[13,16,48]. We find that the negative impact of smoke exposure is present across test subjects, appears stronger on days when kids are in school, and affects communities with differing levels of economic disadvantage and racial/ethnic composition.

In a study of the effect of air pollution exposure on the day of test taking on test performance in Israel, it was found that a 1 standard deviation increase in $PM_{2.5}$ (~16.7 Air Quality Index) led to a 3.9% of a standard deviation decrease in test scores[13]. A 10 μg m^{-3} increase in $PM_{2.5}$ on the day of the National College Entrance Examination in China was estimated to reduce test scores by 4.6% of a standard deviation[16]. We find that a 1 standard deviation increase in the cumulative school day smoke $PM_{2.5}$ (32.5 μg m^{-3}) would result in a decrease in test scores of 0.14% (95% CI: −0.266% to −0.019%) of a standard deviation. These results suggest that contemporaneous air pollution exposure has at least an order of magnitude larger effect on test scores compared with smoke $PM_{2.5}$ exposure in the year before the exam. One explanation for this is that, for exposure during the school year, students can catch up on non-smoky days after suffering learning decrements on smoky days; such catch-up is not possible when the exposure is on test day. Although not directly comparable, we find that wildfire smoke $PM_{2.5}$ results in much larger declines in student test scores than a recent study that measured the effects of total PM on student learning using similar test score data[24] (see Supplementary Information for details). Larger effect sizes in our setting could be because wildfire smoke $PM_{2.5}$ is more harmful[4], or because variation in total ambient PM is correlated with other factors that affect learning in the opposite direction, as is commonly found in health studies[49].

While our data do not allow direct identification of the mechanism by which wildfire smoke affects test performance, existing literature points to multiple plausible pathways. Evidence from air pollution and non-smoke $PM_{2.5}$ suggests that exposures can have

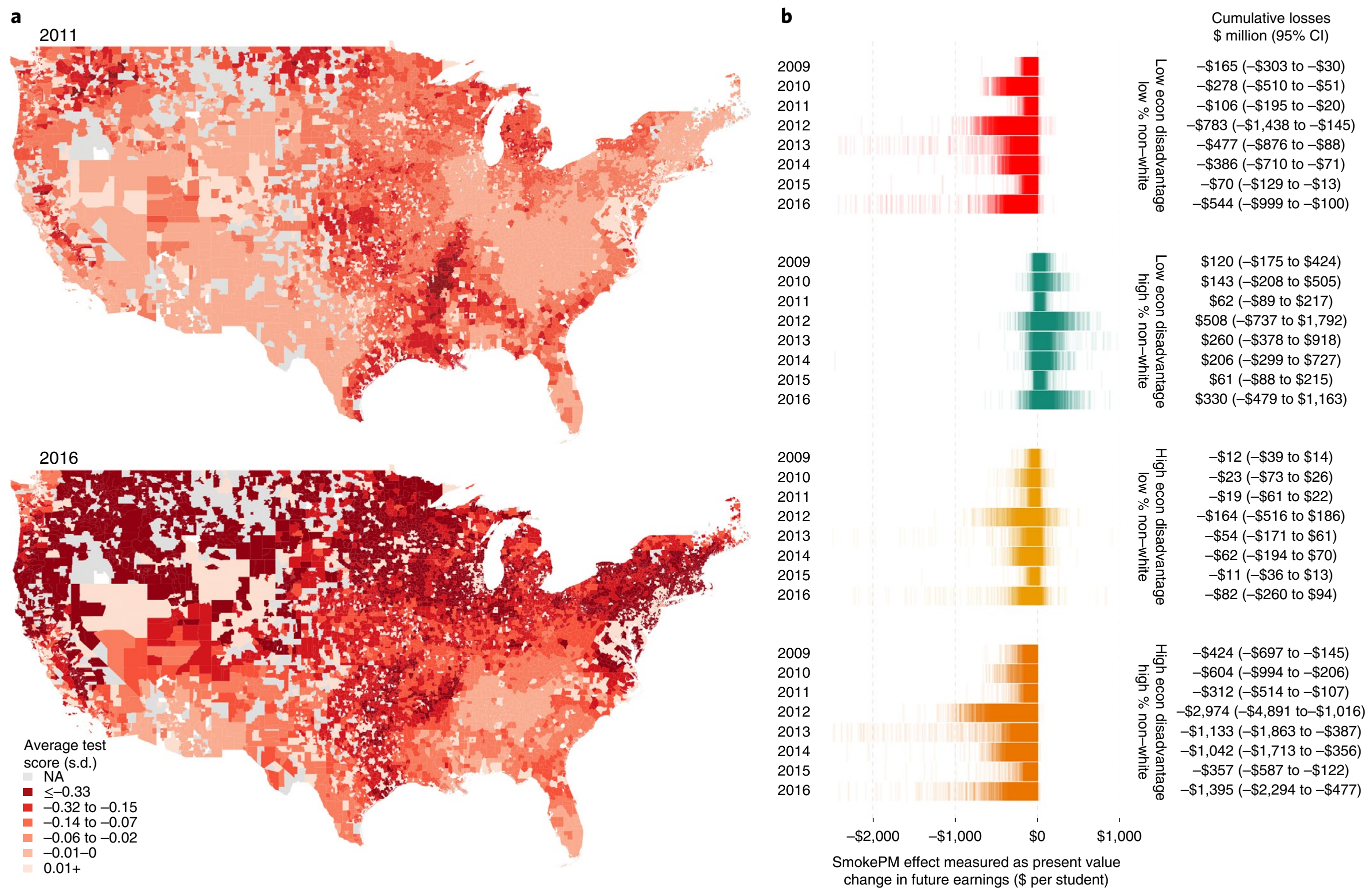

Fig. 4 | School day smoke $PM_{2.5}$ effect on average test scores by district, and total effect on lost future earnings by economic disadvantage and racial/ethnic subgroup over time. a, Predicted effect of cumulative smoke $PM_{2.5}$ exposure during school days on average student test scores in 2011 (top, a relatively low-smoke year in our sample) and 2016 (bottom, a relatively high-smoke year) for each school district[58]. Coefficients used to calculate the predicted effects are estimated with district-level heterogeneity by race/ethnicity and economic disadvantage. The sample has $n = 433,677$ observations with 11,623 districts. **b**, Effect of school day smoke $PM_{2.5}$ on the net present value of future earnings separated by year, economic disadvantage and racial/ethnic subgroup. Each tick mark represents a specific district in the matching year and subgroup (Methods). Cumulative net present value changes in future earnings are provided on the right and represent the total changes in future earnings across all students that fall into the matching year and subgroup, with 5th–95th percentile range across districts in parentheses.

direct biological impacts through neuroinflammation[21] and can be particularly harmful for younger children as the natural barriers in their lungs are still developing and they have a higher rate of breathing relative to their body size[46]. Additionally, particulate exposure can have near-term impacts on cognition[12,13,16] and attention in class[50], both of which could affect cumulative learning. As has been shown for other air pollutants[51], wildfire smoke $PM_{2.5}$ exposures could also increase student absenteeism, given the documented negative short-term effects of smoke exposure on child asthma and other respiratory outcomes[52] that could keep kids from school. Evidence from the broader educational literature documents how cumulative absences affect later test performance[14,53,54]. Both channels provide mechanisms through which repeated short-term exposure to wildfire smoke could have cumulative effects on learning and test performance over the school year.

In a study of the effect of heat on learning, an additional day with temperatures above 80 °F (26.7 °C) during the school year was found to decrease average test scores by 0.07% of a standard deviation[34], which is roughly three times our estimated impact of an average ($6\,\mu g\,m^{-3}$) smoky day. Because there were on average more school days across the United States with temperature above 80 °F (32 days in their sample) than average days with smoke in the air in our sample (7 days per school year), this suggests that the effects of heat are currently a more important determinant of learning outcomes than smoke. Nevertheless, the number of days with smoke in the air and the average concentration of smoke $PM_{2.5}$ on smoky days have both increased dramatically in the few years since the end of our study period[27,47], suggesting a growing influence of smoke in more recent years.

Perhaps surprisingly, we find that estimated effects of smoke on learning are larger in both the least and the most disadvantaged communities. Similar effects at different ends of the disadvantage spectrum could be a result of multiple sources of heterogeneity that each have independent effects across groups. For instance, in districts with high economic disadvantage and high proportion of non-White students, differences in housing or school characteristics—for example, a more permeable building envelope, differences in available filtration or lower access to air conditioning in disadvantaged schools—could allow more ambient pollution to infiltrate into and remain in indoor environments[31–33,47]. While the limited work on infiltration of wildfire smoke does suggest some role for factors such as income, race/ethnicity and housing quality in predicting infiltration into homes[32,47], more widespread measurement in schools will be needed to understand whether differential infiltration can help explain the heterogeneous results we find.

One potential explanation for the observed negative impacts in low-disadvantage communities is if the marginal effect of additional exposure declines at higher baseline $PM_{2.5}$ exposure. Such a nonlinear relationship has been documented in health impacts studies of wildfire specifically[55] and air pollution more broadly[49,56], and could be explained as the result of adaptive investments in communities accustomed to higher average exposures, or alternatively as the result of differences in the relative importance of other determinants of learning (for example, school funding or teacher quality) that happen to be correlated with baseline pollution exposure. Indeed, when we consider the interaction between baseline $PM_{2.5}$ terciles and smoke $PM_{2.5}$, we find that the effects of smoke $PM_{2.5}$ exposure are more negative for districts with lower average baseline $PM_{2.5}$ levels (Supplementary Fig. 2). Additionally, a majority of districts with low economic disadvantage and low proportion of non-White students within our sample are in the lowest baseline $PM_{2.5}$ bin (Supplementary Table 3). However, because other subgroups also appear to have many districts in the lowest baseline $PM_{2.5}$ bin, this explanation is unlikely to fully explain the heterogeneous effects we find.

Estimates of the present value of lost future earnings due to decreased learning outcomes resulting from smoke exposure suggest that in very smoky years, wildfire-attributable-smoke $PM_{2.5}$ would effectively decrease the net present value of future earnings by $55,500 per school (for a school with 500 students) and by nearly $1.7 billion across the United States, with ~80% of these impacts being in disadvantaged communities of colour. While these are rough calculations, as wildfires and the associated smoke events increasingly affect school districts across the United States, estimates such as these can inform cost-benefit analyses of investments aimed at reducing smoke $PM_{2.5}$ exposures. If we consider the combined effect of smoke exposure (aggregated total smoke $PM_{2.5}$ combining school and non-school days smoke $PM_{2.5}$) and conduct a similar calculation as shown in Fig. 4b, we find much larger estimated impacts across all years and subgroups (Supplementary Fig. 4) due to the greater amount of smoke $PM_{2.5}$ on non-school days (includes summer months) compared with school days. As we cannot be sure that students remain around their school districts during summer months and other non-school days, our analysis primarily focuses on school day exposures.

Compared with using satellite-derived smoke plume annotations alone, our approach provides improved estimates of smoke-attributable $PM_{2.5}$ by combining annotations with ground-level predicted $PM_{2.5}$ data to separate smoke $PM_{2.5}$ from background $PM_{2.5}$. However, the smoke plume annotations could be noisy because they are drawn over multiple hours and usually only a couple of times per day[57]. Furthermore, while smoke plume annotations are able to capture relatively new smoke plumes that are visible through satellite imagery, plumes that remain in the atmosphere for multiple days after a wildfire event could be dispersed and difficult to identify through satellite imagery. This could result in underestimation of smoke $PM_{2.5}$ present, due to non-discernible smoke plumes and a lack of smoke plume annotations on these days. Future work to improve the precision of the smoke annotations could lead to more precise estimates of smoke-attributable $PM_{2.5}$. Additionally, we currently do not account for the specific district test taking dates and instead remove any smoke observations between March and May. The exposure calculation could be improved by compiling district-specific testing dates, which would allow us to more precisely measure exposures before the exam. We also note that as our unit of observation is at the district, year and grade level, we are unable to investigate more granular levels of heterogeneity, such as at the individual level.

Our work contributes to a growing body of evidence demonstrating the cognitive, health and social harms of air pollution in general, and wildfires specifically, and shows how disparities in these impacts across socioeconomic and racial/ethnic groups can emerge even when there are negligible differences across groups in ambient exposures. Recent literature has suggested that wildfire smoke particulate matter is potentially more harmful than other sources of particulate matter and impacts to student learning could have an out-sized influence on future outcomes and long-term economic well-being. Hence, we provide rough estimates of the cost of wildfire smoke exposure to illustrate the magnitude of potential benefits that protective investments beneficial to student learning can provide, such as improving classroom filtration or creating clean-air shelters in districts where clean indoor air at home may be inaccessible. Our estimates uncover yet another substantial cost of a warming climate, with future warming-driven increases in wildfire activity likely to worsen learning outcomes.

Methods

Measuring wildfire-smoke-attributable $PM_{2.5}$. To generate estimates of wildfire-smoke-attributable $PM_{2.5}$ across all school districts for all study years, we merged satellite-derived smoke plume data from the National Environmental Satellite, Data, and Information Service (NESDIS) Hazard Mapping System (HMS) with gridded estimates of daily $PM_{2.5}$ concentrations from refs. [39,40]. We then estimated smoke-attributable $PM_{2.5}$ as location- and period-specific anomalous $PM_{2.5}$ on days in which the plume data indicated that smoke was overhead. Plume data were derived from manual annotations by trained analysts, using a variety of remote sensing products including visible-band imagery from the GOES satellites of the National Oceanic and Atmospheric Administration, multiple times per day across the United States[57]. In total, we used nearly 200,000 individual smoke plumes between 2008–2016.

The predicted $PM_{2.5}$ data[39,40] were provided as daily $PM_{2.5}$ concentrations for all-source $PM_{2.5}$ (not just wildfire $PM_{2.5}$) for the contiguous United States in a 1 km grid from 2000–2016. The predictions were made using an ensemble of three machine learning models including neural networks, random forests and gradient boosted trees. Each of the models comprised multiple explanatory variables, including satellite observations, land-use variables, chemical transport predictions and other variables. The authors note that the ensemble model achieved a performance of $r^2 = 0.86$ for daily $PM_{2.5}$ predictions.

To isolate $PM_{2.5}$ from wildfires, we followed ref. [47] and calculated smoke-attributable $PM_{2.5}$ as the deviation from location-specific median $PM_{2.5}$ on non-smoke days in the same month, with the median calculated over a 3-year window centred on the current year. Specifically, the smoke-attributable $PM_{2.5}$ anomaly was calculated by subtracting the month-specific 3-year non-smoke day median estimated from the predicted $PM_{2.5}$ at each school district on days with a smoke plume overhead. After we obtained the smoke $PM_{2.5}$ anomalies, we set this smoke $PM_{2.5}$ variable to 0 for non-smoke days and the positive anomaly for days with a plume overhead. Smoke days with negative anomaly values were also set to 0. The resulting measure of smoke $PM_{2.5}$ isolates the smoke component from overall $PM_{2.5}$ as long as, on average, other $PM_{2.5}$ sources are not also anomalously high on days when smoke is in the air—a plausible assumption given the large degree of temporal and spatial randomness in when and where fires start and where plumes go.

Assigning smoke $PM_{2.5}$ exposure to school districts. We calculated a student-population weighted average of school level exposure to estimate aggregate exposure at the district level. We further delineated school day exposure versus non-school day exposure, specifying non-school days as weekends and federal bank holidays throughout the year and all days from June 15 to August 15. Because standardized testing in the United States is conducted at various points throughout the spring, usually between March and May, our analysis focused on exposures from the previous June to February. For this analysis, we focused on school years between 2009–2016 as the predicted $PM_{2.5}$ data are only available between 2000 and 2016.

Outcome and covariate data. Test score data were compiled by Stanford University and made available through the Stanford Education Data Archive (SEDA)[58]. The SEDA data are derived from state-level standardized accountability tests for math and English Language Arts (ELA) that are administered to all public-school students. These tests are typically taken between March and May and raw scores are provided in aggregated form by the US Department of Education. The SEDA team constructed the dataset by converting state-specific proficiency data to a nationally comparable dataset by scaling the state results using a nationally representative sample from the National Assessment of Educational Progress (NAEP). NAEP is a test taken in every state by a random sample of students in Grades 4 and 8 in math and ELA in odd years (for example, 2009, 2011, 2013, 2015, 2017 and 2019). SEDA interpolates for grades and years where the NAEP was not administered and uses each states' grade, year and subject results on the NAEP to rescale the state-standardized accountability tests[58].

For additional details about how the dataset was created and the calculations involved in scaling the state scores to a national dataset, please see ref. [58]. The SEDA data contain nationally comparable test scores for students in Grades 3–8 from 2009–2018. These test scores are broken down into district-level results for both math and ELA subjects. Rather than represent an absolute score, the metric provided in the dataset is a standardized score within subject and grade, relative to representative cohorts who had taken the NAEP assessments[59]. Therefore, a score of 0.25 for math means that an average student in that district performed 0.25 of a standard deviation higher than the reference cohort that took the NAEP assessment. The primary outcome we considered is the average test score for ELA and math at the district-year-grade level.

In addition to calculating an average treatment effect across aggregated data in our main model specification, we also investigated heterogeneous effects using district-level racial/ethnic and economic disadvantage covariates. Specifically, the level of economic disadvantage is measured by the Federal EdFacts data system and is typically defined using the proportion of students eligible for free or reduced-price lunch[60]. The proportion of non-White students in a district was calculated by subtracting the proportion of White students from 1, where the proportion of White students is collected by the Common Core of Data and aggregated by the SEDA team to the district level[59].

We used gridded (4 km × 4 km) temperature and precipitation data produced by the Parameter elevation Regressions on Independent Slopes Model (PRISM) Climate Group at Oregon State University[41]. We extracted the maximum daily temperature at each school and took a weighted average using the student population at the schools that belong to that district. As with the weighting for smoke exposure, the student population data were collected from the National Center for Education Statistics. We then created bins by counting the number of days in the year before the exam with maximum temperatures from 0 °F (−17.7 °C) to 80 °F (26.7 °C), in 10 °F (5.5 °C) increments. All days less than or equal to 0 °F (−17.7 °C) were grouped into 1 bin and all days greater than 80 °F (26.7 °C) were grouped into another bin. We processed the daily precipitation data similarly and extracted the daily precipitation measurements at each school, then calculated the total annual precipitation at the district level.

Estimating the effect of smoke $PM_{2.5}$ on student performance. Our main regression specification is as shown in equation (1):

$$\begin{aligned}\text{Score}_{igy} &= \beta_1 \text{SmokePM}_{iy}^{\text{school}} + \beta_2 \text{SmokePM}_{iy}^{\text{non-school}} \\ &\quad + f(\mathbf{X}_{iy}) + \eta_i + \gamma_{yg} + \epsilon_{igy}\end{aligned} \tag{1}$$

Here, 'Score' represents the scaled standardized score for each district i in grade g and year y. The district fixed effect η_i is a separate intercept (dummy variable) for each district (that is, 11,639 dummies with one for each district) that accounts for any average differences in smoke exposure or test scores across districts. This empirical approach ensures that we are not comparing districts that might inherently be very different from each other. The year-grade fixed effects γ_{yg} (that is, 6 dummies for each of the grades in each of the 8 years for a total of 48 dummies) account for differences in year-grade-specific exposures or outcomes that affect all districts across the United States, such as overall trends in test scores or in average differences between test scores across grades. $f(\mathbf{X})$ represents a vector of controls including the number of days with maximum temperatures in each of the 10 °F (5.5 °C) bins, which controls for potential nonlinear effects of temperature within the district for the year preceding the test, and total annual precipitation in the year before the exam. 'SmokePM' is defined as the total amount of smoke-attributable $PM_{2.5}$ in the year before the exam in our primary specification. We define $\text{SmokePM}^{\text{school}}$ and $\text{SmokePM}^{\text{non-school}}$ as the total cumulative smoke $PM_{2.5}$ in a preceding year y within district i between June and February on school and non-school days, respectively. β_1 represents the average effect of an additional $\mu g\,m^{-3}$ of cumulative smoke $PM_{2.5}$ on school days on test performance. β_2 represents the average effect of an additional $\mu g\,m^{-3}$ of cumulative smoke $PM_{2.5}$ on non-school days on test performance. To estimate the subject-specific coefficient estimates (Fig. 2b), we replaced mean score Score_{igy} with the subject-specific scores and considered the cumulative 'SmokePM' (both school and non-school) in the year before the exam. We clustered standard errors by county to account for arbitrary within-unit autocorrelation in ϵ_{igy}[42] and weighted districts by the total number of students who took the test provided in the SEDA dataset. Estimation of the main model was carried out using the fixest package (0.8.4)[61] in the R programming language (4.0.4). Both the level of clustering and student population weights were provided to the fixest model object when fitting the regression model. In all statistical tests, we assumed normality but did not formally test for it and all tests of significance were two-tailed.

To investigate whether the observed effect is due to random noise, we conducted a randomization inference test and took the observations within a county and randomly permuted the school day smoke $PM_{2.5}$ variable of district, year and grade observations without replacement while keeping other variables constant. Then we ran the same regression as in our main specification and recorded the school day smoke $PM_{2.5}$ coefficient estimate. We repeated this procedure 1,000 times and plotted out the distribution of estimated effects.

We also conducted secondary analyses (equation 2) looking at the heterogeneous effects of smoke exposure on test outcomes. To examine these effects, we studied whether the effects of smoke $PM_{2.5}$ differed across different grade levels and a combination of economic disadvantage and race/ethnicity, using the following specification:

$$\begin{aligned}\text{Score}_{igy} &= \sum_n \beta_n (\mathbb{1}_n * \text{SmokePM}_{iy}^{\text{school}}) \\ &\quad + \sum_n \beta_n (\mathbb{1}_n * \text{SmokePM}_{iy}^{\text{non-school}}) \\ &\quad + f(\mathbf{X}_{iy}) + \eta_i + \gamma_{yg} + \epsilon_{igy}\end{aligned} \tag{2}$$

Here, $\mathbb{1}_n$ represents an indicator function for whether or not the observation i falls into a specific bin n. In the primary vs secondary school-aged student analysis, grades were divided into grade buckets with Grades 3–5 grouped into 'primary' and Grades 6–8 into 'secondary'. Then as shown in the equation above, we considered the interaction between the grade bucket and smoke $PM_{2.5}$ experienced in the year before the exam to estimate the effect on student test scores in that district. Similarly, in the race/ethnicity and economic disadvantage heterogeneity analysis, we divided districts into 'High' or 'Low' categories on the basis of thresholding at the median value for race/ethnicity and economic disadvantage variables. We then estimated heterogeneous effects using a regression model that considers how the interaction of the district's level of non-white population, level of economic disadvantage and smoke $PM_{2.5}$ experienced in the year before the exam is associated with student test scores in that district. The remainder of the equation is similar to equation (1).

In Supplementary Fig. 1, we stratified by racial/ethnic group and ran separate regressions for Asian, Black, Hispanic and White subgroups. These regressions are similar to equation (2) above but instead consider how the interaction between 'High' or 'Low' levels of Asian, Black, Hispanic or White population, 'High' or 'Low' levels of economic disadvantage and smoke $PM_{2.5}$ experienced in the year before the exam is associated with student test scores in that district. Again, these racial/ethnic group stratified regressions also consider the same fixed-effects (dummy variables) and covariates as in the main specification.

Projecting the effect of smoke $PM_{2.5}$ as lost future income. To translate the effect estimates into the net present value of lost future earnings, we followed the approach used in ref. [33]. Specifically, we assumed that the relationship found in ref. [48] holds, which estimated that a 1 standard deviation increase in teacher quality raised average test scores by 0.13 standard deviations and resulted in a net present value of \$7,000 in future increased earnings for 12-year-old students. Therefore, if the estimated effect of an additional $\mu g\,m^{-3}$ of smoke $PM_{2.5}$ is a decrease of 0.01% of a standard deviation and the average smoke $PM_{2.5}$ experienced in a year is $10\,\mu g\,m^{-3}$, then we calculate the average effect as $0.01\% \times 10 = 0.1\%$. We can then apply the conversion in ref. [48] and calculate that $\frac{0.001 \times 7000}{0.13} = \$53.85$ on average per student for that year of smoke $PM_{2.5}$ exposure.

In Fig. 4b, we plot the average net present value change in future earnings for each district as an individual tick mark. For each of the four economic disadvantage and racial/ethnic subgroups, we drew 3,000 samples from a normal distribution, with the mean centred at the matching subgroup coefficient (equation 2) and the standard deviation set to the estimated standard error. We then merged this with district information by matching on the districts' subgroup for each year. From this data, we estimated the district-specific average impacts by year and we sampled 1 observation out of the 3,000 samples to show as a tick mark. Additionally, we used the sampled data to estimate 95% interval estimates for the cumulative changes in net present value of future earnings.

Calculating distance to nearest fire perimeter. We calculated distances from schools to the nearest fire perimeter in each year provided by the National Interagency Fire Center[62]. Then, we took an average of the minimum school-to-fire distances to get estimates of the average distance to the nearest fire for each district. While evacuation zone distances vary, recent studies of wildfire evacuations in California suggest that short-distance evacuations are much more common than longer distance evacuations to destinations outside of the county of residence[63]. Additionally, a 1.5 mile (2.4 km) distance is often cited as the distance that forest fire embers can travel and ignite flammable materials at distant locations beyond the fire front[64]. Given this, we iteratively dropped school districts that are 1 km, 3 km, 5 km, 10 km and 20 km away from the nearest fire perimeter and for each drop distance, we ran a similar regression as our main specification without the dropped school districts. We found that the identified effects are consistent up to dropping districts within 10 km of the nearest fire perimeter (Supplementary Fig. 3).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The weather data used in this study are available through the Parameter elevation Regressions on Independent Slopes Model (PRISM) Climate Group at Oregon

State University (https://prism.oregonstate.edu). Student test performance and district-level covariate data are available through the Stanford Education Data Archive (SEDA) (https://purl.stanford.edu/db586ns4974). School location and student population data are available through the National Center for Education Statistics (NCES). Fire perimeter data used to calculate the distance of school districts to fire perimeters are available through the National Interagency Fire Center (NIFC) (https://data-nifc.opendata.arcgis.com/datasets/nifc::interagency-fire-perimeter-history-all-years). Smoke plume annotations are available through the National Environmental Satellite, Data and Information Service (NESDIS) Hazard Mapping System (HMS) (https://www.ospo.noaa.gov/Products/land/hms.html#data). Daily gridded estimates of $PM_{2.5}$ concentrations are available from Di et al. (2021) (https://doi.org/10.7927/0rvr-4538; https://doi.org/10.1016/j.envint.2019.104909). Processed data to replicate the results in the main text and Supplementary Information are available at https://github.com/jeffwen/wildfire_smoke_education_public.

Code availability

The code to replicate the results and figures in the main text and supplementary material are available at https://github.com/jeffwen/wildfire_smoke_education_public.

Received: 24 February 2022; Accepted: 4 August 2022;
Published online: 29 September 2022

References

1. Kitzberger, T., Falk, D. A., Westerling, A. L. & Swetnam, T. W. Direct and indirect climate controls predict heterogeneous early-mid 21st century wildfire burned area across western and boreal North America. *PloS ONE* **12**, e0188486 (2017).
2. Dennison, P. E., Brewer, S. C., Arnold, J. D. & Moritz, M. A. Large wildfire trends in the western United States, 1984–2011. *Geophys. Res. Lett.* **41**, 2928–2933 (2014).
3. Reid, C. E. et al. Critical review of health impacts of wildfire smoke exposure. *Environ. Health Perspect.* **124**, 1334–1343 (2016).
4. Aguilera, R., Corringham, T., Gershunov, A. & Benmarhnia, T. Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. *Nat. Commun.* **12**, 1493 (2021).
5. Calderón-Garcidueñas, L. et al. Air pollution, cognitive deficits and brain abnormalities: a pilot study with children and dogs. *Brain Cogn.* **68**, 117–127 (2008).
6. Block, M. L. & Calderón-Garcidueñas, L. Air pollution: mechanisms of neuroinflammation and CNS disease. *Trends Neurosci.* **32**, 506–516 (2009).
7. Marcotte, D. E. Something in the air? Air quality and children's educational outcomes. *Econ. Educ. Rev.* **56**, 141–151 (2017).
8. Künn, S., Palacios Temprano, J. F. & Pestel, N. *Indoor Air Quality and Cognitive Performance* Discussion Paper No. 12632 (IZA, 2019).
9. Huang, J., Xu, N. & Yu, H. Pollution and performance: do investors make worse trades on hazy days? *Manage. Sci.* **66**, 4455–4476 (2020).
10. Chang, T. Y., Graff Zivin, J., Gross, T. & Neidell, M. The effect of pollution on worker productivity: evidence from call center workers in China. *Am. Econ. J. Appl. Econ.* **11**, 151–172 (2019).
11. Archsmith, J., Heyes, A. & Saberian, S. Air quality and error quantity: pollution and performance in a high-skilled, quality-focused occupation. *J. Assoc. Environ. Resour. Econ.* **5**, 827–863 (2018).
12. La Nauze, A. & Severnini, E. *Air Pollution and Adult Cognition: Evidence from Brain Training* Working Paper 28785 (NBER, 2021).
13. Ebenstein, A., Lavy, V. & Roth, S. The long-run economic consequences of high-stakes examinations: evidence from transitory variation in pollution. *Am. Econ. J. Appl. Econ.* **8**, 36–65 (2016).
14. Shier, V., Nicosia, N., Shih, R. & Datar, A. Ambient air pollution and children's cognitive outcomes. *Popul. Environ.* **40**, 347–367 (2019).
15. Zhang, X., Chen, X. & Zhang, X. The impact of exposure to air pollution on cognitive performance. *Proc. Natl Acad. Sci. USA* **115**, 9193–9197 (2018).
16. Zivin, J. G., Liu, T., Song, Y., Tang, Q. & Zhang, P. The unintended impacts of agricultural fires: human capital in China. *J. Dev. Econ.* **147**, 102560 (2020).
17. Lai, W., Li, S., Li, Y. & Tian, X. Air pollution and cognitive functions: evidence from straw burning in China. *Am. J. Agric. Econ.* **104**, 190–208 (2022).
18. Laurent, J. G. C. et al. Associations between acute exposures to $PM_{2.5}$ and carbon dioxide indoors and cognitive function in office workers: a multicountry longitudinal prospective observational study. *Environ. Res. Lett.* **16**, 094047 (2021).
19. Shehab, M. & Pope, F. Effects of short-term exposure to particulate matter air pollution on cognitive performance. *Sci. Rep.* **9**, 8237 (2019).
20. Carneiro, J., Cole, M. A. & Strobl, E. The effects of air pollution on students' cognitive performance: evidence from Brazilian university entrance tests. *J. Assoc. Environ. Resour. Econ.* **8**, 1051–1077 (2021).
21. Gao, X. et al. Short-term air pollution, cognitive performance and nonsteroidal anti-inflammatory drug use in the veterans affairs normative aging study. *Nat. Aging* **1**, 430–437 (2021).
22. Calderón-Garcidueñas, L. et al. Long-term air pollution exposure is associated with neuroinflammation, an altered innate immune response, disruption of the blood-brain barrier, ultrafine particulate deposition, and accumulation of amyloid β-42 and α-synuclein in children and young adults. *Toxicol. Pathol.* **36**, 289–310 (2008).
23. Fu, P., Guo, X., Cheung, F. M. H. & Yung, K. K. L. The association between $PM_{2.5}$ exposure and neurological disorders: a systematic review and meta-analysis. *Sci. Total Environ.* **655**, 1240–1248 (2019).
24. Lu, W., Hackman, D. A. & Schwartz, J. Ambient air pollution associated with lower academic achievement among US children: a nationwide panel study of school districts. *Environ. Epidemiol.* **5**, e174 (2021).
25. Gilraine, M. *Air Filters, Pollution and Student Achievement* EdWorkingPapers (Annenberg Institute at Brown University, 2020); http://www.edworkingpapers.com/ai20-188
26. Liu, J. C. et al. Wildfire-specific fine particulate matter and risk of hospital admissions in urban and rural counties. *Epidemiology* **28**, 77-85 (2017).
27. Burke, M. et al. The changing risk and burden of wildfire in the United States. *Proc. Natl Acad. Sci. USA* **118**, e2011048118 (2021).
28. Marlon, J. R. et al. Long-term perspective on wildfires in the western USA. *Proc. Natl Acad. Sci. USA* **109**, E535–E543 (2012).
29. Westerling, A. L., Hidalgo, H. G., Cayan, D. R. & Swetnam, T. W. Warming and earlier spring increase western U.S. forest wildfire activity. *Science* **313**, 940–943 (2006).
30. Colmer, J., Hardman, I., Shimshack, J. & Voorheis, J. Disparities in $PM_{2.5}$ air pollution in the United States. *Science* **369**, 575–578 (2020).
31. Chan, W. R., Joh, J. & Sherman, M. H. Analysis of air leakage measurements of US houses. *Energy Build.* **66**, 616–625 (2013).
32. Liang, Y. et al. Wildfire smoke impacts on indoor air quality assessed using crowdsourced data in California. *Proc. Natl Acad. Sci. USA* https://www.pnas.org/content/118/36/e2106478118 (2021). https://www.pnas.org/content/118/36/e2106478118.full.pdf
33. Park, R. J., Goodman, J., Hurwitz, M. & Smith, J. Heat and learning. *Am. Econ. J. Econ. Policy* **12**, 306–339 (2020).
34. Park, R. J., Behrer, A. P. & Goodman, J. Learning is inhibited by heat exposure, both internationally and within the United States. *Nat. Hum. Behav.* **5**, 19–27 (2021).
35. Tessum, C. W. et al. Inequity in consumption of goods and services adds to racial–ethnic disparities in air pollution exposure. *Proc. Natl Acad. Sci. USA* **116**, 6001–6006 (2019).
36. Su, J. G., Jerrett, M., Morello-Frosch, R., Jesdale, B. M. & Kyle, A. D. Inequalities in cumulative environmental burdens among three urbanized counties in California. *Environ. Int.* **40**, 79–87 (2012).
37. Lipfert, F. Air pollution and poverty: does the sword cut both ways? *J. Epidemiol. Community Health* **58**, 2-3 (2004).
38. Legot, C., London, B. & Shandra, J. Environmental ascription: high-volume polluters, schools, and human capital. *Organ. Environ.* **23**, 271–290 (2010).
39. Di, Q. et al. An ensemble-based model of $PM_{2.5}$ concentration across the contiguous United States with high spatiotemporal resolution. *Environ. Int.* **130**, 104909 (2019).
40. Di, Q. et al. *Daily and Annual PM2.5 Concentrations for the Contiguous United States, 1-km Grids*, v1 (2000–2016) (NASA, 2021); https://sedac.ciesin.columbia.edu/data/set/aqdh-pm2-5-concentrations-contiguous-us-1-km-2000-2016
41. *PRISM Gridded Climate Data 2019* (Oregon State University PRISM Climate Group, accessed 2 March 2020); https://prism.oregonstate.edu
42. Bertrand, M., Duflo, E. & Mullainathan, S. How much should we trust differences-in-differences estimates? *Q. J. Econ.* **119**, 249–275 (2004).
43. Abadie, A., Athey, S., Imbens, G. W. & Wooldridge, J. *Clustering as a Design Problem* Working paper (MIT Economics, 2016); http://economics.mit.edu/files/11355
44. Blattman, C., Green, D. P., Ortega, D. & Tobón, S. Place-based interventions at scale: the direct and spillover effects of policing and city services on crime. *J. Eur. Econ. Assoc.* **19**, 2022–2051 (2021).
45. Mullen, C., Grineski, S. E., Collins, T. W. & Mendoza, D. L. Effects of PM2.5 on third grade students' proficiency in math and english language arts. *Int. J. Environ. Res. Public Health* **17**, 6931 (2020).
46. Brockmeyer, S. & d'Angiulli, A. How air pollution alters brain development: the role of neuroinflammation. *Transl. Neurosci.* **7**, 24–30 (2016).
47. Burke, M. et al. *Exposures and Behavioral Responses to Wildfire Smoke* Working Paper 29380 (NBER, 2021).
48. Chetty, R., Friedman, J. N. & Rockoff, J. E. Measuring the impacts of teachers II: teacher value-added and student outcomes in adulthood. *Am. Econ. Rev.* **104**, 2633–79 (2014).
49. Deryugina, T., Heutel, G., Miller, N. H., Molitor, D. & Reif, J. The mortality and medical costs of air pollution: evidence from changes in wind direction. *Am. Econ. Rev.* **109**, 4178–4219 (2019).

50. Sunyer, J. et al. Traffic-related air pollution and attention in primary school children: short-term association. *Epidemiology* **28**, 181–189 (2017).
51. Currie, J., Hanushek, E. A., Kahn, E. M., Neidell, M. & Rivkin, S. G. Does pollution increase school absences? *Rev. Econ. Stat.* **91**, 682–694 (2009).
52. Holm, S. M., Miller, M. D. & Balmes, J. R. Health effects of wildfire smoke in children and public health tools: a narrative review. *J. Expo. Sci. Environ. Epidemiol.* **31**, 1–20 (2021).
53. Mohai, P., Kweon, B.-S., Lee, S. & Ard, K. Air pollution around schools is linked to poorer student health and academic performance. *Health Aff.* **30**, 852–862 (2011).
54. Klein, M., Sosu, E. M. & Dare, S. School absenteeism and academic achievement: does the reason for absence matter? *AERA Open* **8**, 23328584211071115 (2022).
55. Heft-Neal, S., Driscoll, A., Yang, W., Shaw, G. & Burke, M. Associations between wildfire smoke exposure during pregnancy and risk of preterm birth in California. *Environ. Res.* **203**, 111872 (2022).
56. Burnett, R. et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc. Natl Acad. Sci. USA* **115**, 9592–9597 (2018).
57. Brey, S. J., Ruminski, M., Atwood, S. A. & Fischer, E. V. Connecting smoke plumes to sources using hazard mapping system (hms) smoke and fire location data over North America. *Atmos. Chem. Phys.* **18**, 1745–1761 (2018).
58. Reardon, S. F. et al. *Stanford Education Data Archive* (version 4.0) (Stanford Libraries, 2021); http://purl.stanford.edu/db586ns4974
59. Fahle, E. M. et al. *Stanford Education Data Archive Technical Documentation* Version 4.0 (Stanford Libraries, 2021).
60. Reardon, S. F. Educational opportunity in early and middle childhood: using full population administrative data to study variation by place and age. *RSF* **5**, 40–68 (2019).
61. Bergé, L. *Efficient Estimation of Maximum Likelihood Models with Multiple Fixed-effects: the R Package FENmlm* Discussion Papers (CREA, University of Luxembourg, 2018).
62. Wildland Fire Management Research, Development, & Application program data team. *Interagency Fire Perimeter History - All Years* (National Interagency Fire Center, accessed 30 January 2021); https://data-nifc.opendata.arcgis.com/datasets/nifc::interagency-fire-perimeter-history-all-years/about
63. Wong, S. D., Broader, J. C. & Shaheen, S. A. *Review of California Wildfire Evacuations from 2017 to 2019* (UC Berkeley Institute of Transportation Studies, 2020).
64. Radeloff, V. C. et al. The wildland–urban interface in the United States. *Ecol. Appl.* **15**, 799–805 (2005).

Acknowledgements

We thank the ECHOLab at Stanford University and the Exploring the 'Hidden Burden' of Climate Change and Pollution on Mental Health and Conflict session at American Geophysical Union 2020 for helpful discussions and comments. M.B. thanks the Robert Wood Johnson Foundation (ID No. 76555) for funding. J.W. gratefully acknowledges partial funding from Stanford Data Science for this work. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

J.W. and M.B. contributed to the conception and design of the study. J.W. conducted data extraction and econometric analysis. J.W. and M.B. analysed the results and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41893-022-00956-y.

Correspondence and requests for materials should be addressed to Jeff Wen.

Peer review information *Nature Sustainability* thanks Tarik Benmarhnia, Caroline Gao, Fay Johnston and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Jeff Wen

Last updated by author(s): July 26, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used to collect data.
Data analysis	R version 4.0.4 (2021-02-15) was used for the analysis and figure generation. Code to replicate the results in the main text and supplementary information are available at https://github.com/jeffwen/wildfire_smoke_education_public

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The weather data used in this study is available through the Parameter elevation Regressions on Independent Slopes Model (PRISM) Climate Group at Oregon State University (https://prism.oregonstate.edu). Student test performance and district level covariate data are available through the Stanford Education Data Archive (SEDA) (https://purl.stanford.edu/db586ns4974). School location and student population data are available through the National Center for Education Statistics

(NCES). Fire perimeter data used to calculate the distance of school districts to fire perimeters is available through the National Interagency Fire Center (NIFC) (https://data-nifc.opendata.arcgis.com/datasets/nifc::interagency-fire-perimeter-history-all-years). Smoke plume annotations are available through the National Environmental Satellite, Data, and Information Service (NESDIS) Hazard Mapping System (HMS) (https://www.ospo.noaa.gov/Products/land/hms.html#data). Daily gridded estimates of PM2.5 concentrations are available from Di et al. (2021) (https://doi.org/10.7927/0rvr-4538; https://doi.org/10.1016/j.envint.2019.104909). Processed data to replicate the results in the main text and supplementary information are available at https://github.com/jeffwen/wildfire_smoke_education_public.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	Test scores are aggregated at the district-grade level in the SEDA dataset. Information on sex and/or gender are not included in the SEDA dataset.
Population characteristics	Test scores are aggregated at the district-grade level in the SEDA dataset. District-grade level covariates including race/ethnicity and economic disadvantage were used to estimate heterogeneous effects.
Recruitment	SEDA data is based on aggregated standardized testing data from all public-school students in grades 3-8. Data was provided to SEDA in aggregate from by the U.S. Department of Education.
Ethics oversight	Only aggregated data was used for this analysis. Approval of study protocol was not required.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Quantitative quasi-experimental method using observational data.
Research sample	The data sample used in this analysis consists of public school student test scores for students in US school districts from 3-8th grades from 2009-2016 in 11,639 districts. These data are compiled by the Stanford Education Data Archive (SEDA).
Sampling strategy	All US public school districts included in SEDA were used for analysis.
Data collection	The student test performance and district level covariate data used in this study are downloaded from the Stanford Education Data Archive (SEDA). School location and student population data is from the National Center for Education Statistics (NCES). Weather data is downloaded from the Parameter elevation Regressions on Independent Slopes Model (PRISM) Climate Group at Oregon State University. Fire perimeter data used to calculate the distance of school districts to fire perimeters is downloaded from the National Interagency Fire Center (NIFC). The smoke plume annotations are downloaded from the National Environmental Satellite, Data, and Information Service (NESDIS) Hazard Mapping System (HMS). Daily gridded estimates of PM2.5 concentrations are downloaded from Di et al. (2021).
Timing and spatial scale	Measurements used for this analysis were collected from 2009-2016.
Data exclusions	- 2 years of SEDA data (n=128,588 observations) were excluded because the gridded PM2.5 data used to estimate wildfire attributable PM2.5 extended until 2016 while SEDA extended to 2018. - n=100,700 observations dropped out due to missing English language arts and/or math scores. - n=7 observations dropped out because of missing average temperature data.
Reproducibility	Experiments were not conducted for this study.
Randomization	The authors did not conduct randomization. Instead, a quasi-experimental approach was used to analyze observational data.
Blinding	Blinding was not used during the data acquisition or analysis steps of this study.

Did the study involve field work? ☐ Yes ☒ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging