EXPOSURES AND BEHAVIORAL RESPONSES TO WILDFIRE SMOKE

Marshall Burke
Sam Heft-Neal
Jessica Li
Anne Driscoll
Patrick W. Baylis
Matthieu Stigler
Joakim Weill
Jennifer Burney
Jeff Wen
Marissa Childs
Carlos Gould

Working Paper 29380
http://www.nber.org/papers/w29380

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2021

We thank the Robert Wood Johnson Foundation for funding, SafeGraph for data access, and thank members of the ECHOLab and seminar participants at Cornell, Columbia, MIT, Stanford, UC Berkeley, and UC Santa Barbara for helpful comments. All errors are our own. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Marshall Burke, Sam Heft-Neal, Jessica Li, Anne Driscoll, Patrick W. Baylis, Matthieu Stigler, Joakim Weill, Jennifer Burney, Jeff Wen, Marissa Childs, and Carlos Gould. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.
Exposures and Behavioral Responses to Wildfire Smoke
Marshall Burke, Sam Heft-Neal, Jessica Li, Anne Driscoll, Patrick W. Baylis, Matthieu Stigler, Joakim Weill, Jennifer Burney, Jeff Wen, Marissa Childs, and Carlos Gould
NBER Working Paper No. 29380
October 2021
JEL No. Q5,Q53

ABSTRACT

The impacts of environmental change on human outcomes often depend on local exposures and behavioral responses that are challenging to observe with traditional administrative or sensor data. We show how data from private pollution sensors, cell phones, social media posts, and internet search activity yield new insights on exposures and behavioral responses during large wildfire smoke events across the US, a rapidly-growing environmental stressor. Health-protective behavior, mobility, and sentiment all respond to increasing ambient wildfire smoke concentrations, but responses differ by income. Indoor pollution monitors provide starkly different estimates of likely personal exposure during smoke events than would be inferred from traditional ambient outdoor sensors, with similar outdoor pollution levels generating >20x differences in average indoor PM2.5 concentrations. Our results suggest that the current policy reliance on self protection to mitigate health risks in the face of rising smoke exposure will result in modest and unequal benefits.

Marshall Burke
Department of Earth System Science Stanford University
Stanford, CA 94305
and NBER
mburke@stanford.edu

Anne Driscoll
Stanford University
616 Serra St
Stanford, CA 94305
annedriscoll@cmu.edu

Sam Heft-Neal
Center on Food Security and the Environment
Stanford University
473 Via Ortega
Stanford, CA 94305
sheftneal@stanford.edu

Patrick W. Baylis
Vancouver School of Economics
University of British Columbia
6000 Iona Drive
Vancouver, BC V6T1L4
Canada
patrick.baylis@ubc.ca

Jessica Li
Center on Food Security and the Environment
Stanford University
473 Via Ortega
Stanford, CA 94305
jbli@stanford.edu

Matthieu Stigler
Stanford University
Y2E2 Building
Stanford, CA 94305
mstigler@stanford.edu
Joakim Weill  
Department of Agricultural  
and Resource Economics  
University of California, Davis  
One Shields Ave  
Davis, CA 95616  
jweill@ucdavis.edu

Jennifer Burney  
University of California San Diego  
9500 Gilman Dr. #0519  
La Jolla, CA 92093  
jburney@ucsd.edu

Jeff Wen  
Department of Earth System Science  
Stanford University  
473 Via Ortega  
Stanford, CA 94305  
jlwen@stanford.edu

Marissa Childs  
Stanford University  
marissac@stanford.edu

Carlos Gould  
Department of Earth System Science  
Stanford University  
473 Via Ortega  
Stanford, CA 94305  
cfgould@stanford.edu
Introduction

A large body of scientific evidence documents how environmental exposures can substantially shape human outcomes. For instance, poor air quality is estimated to kill millions of people per year (1), warming temperatures lead to more crime and less economic output (2), and exposure to lead reduces educational achievement (3). Effects can vary substantially across subgroups: air pollution is more harmful to health in poorer US counties (4), warming temperatures have more negative effects on economic output in already hot locations (5), and lead exposures appear to have larger effects in already disadvantaged households (6). Standard models of behavior in both public health and economics suggest the magnitude and variation of these effects in part reflect choices that individuals make or are unable to make regarding their exposures, and that the choice sets available to individuals are in turn reflective of individuals’ knowledge, circumstances, and preferences (7, 8). Unfortunately, these decision-making components, as well as their behavioral outcomes, are typically hard to observe at scale. This makes it difficult to understand why a given environmental exposure generates the effect it does, why this effect might differ across groups, and whether and how policy should respond.

Here we show how combining traditional sensor and survey data with information from non-traditional distributed sensors – including data from private outdoor and indoor pollution sensors, cell phones, social media posts, and internet search activity – can generate population-scale insights into people’s knowledge, preferences, and choices regarding a changing environment, and into how economic circumstances shape their choice set. We focus on understanding responses to wildfire smoke, a rapidly growing environmental stressor throughout much of the US and internationally. Annual area burned by wildfires in the US has more than doubled in recent decades, a result of a century of fire suppression and a warming climate that has left the resulting abundant fuel much more flammable (9). This increase in fire activity has led to substantial increases in average smoke exposure across the continental US, potentially reversing decades of improvements in air quality (10). Absent substantial intervention, these trends are expected to continue and perhaps accelerate in a warming climate (11–13).

A growing literature has begun to document myriad health impacts of ambient wildfire smoke exposure (14–18). As with impacts of other environmental stressors, the magnitude of these health impacts likely depend on behaviors and individual-specific exposures that are often poorly measured and understood. In particular, recent wildfire case studies suggest that individuals vary in their knowledge and beliefs about their own exposures and about the risk that these exposures pose to their health (19, 20). Documented heterogeneity in the impacts of wildfire smoke expo-
sure or exposure to other particulates (4, 18, 21–23) suggest that both socioeconomic circumstances and previous exposures could also constrain behavior in important ways. Understanding which features matter most is consequential for policy design: impacts driven by a lack of awareness of one’s exposure, for instance, call for different interventions than impacts driven by an inability to protect oneself from a known exposure. Such understanding is particularly important for wildfire, given that current policy approaches to risk mitigation focus on the private provision of protection – i.e., asking individuals to stay indoors, limit infiltration, and purchase protective technologies (24).

To better understand exposures, behavioral responses, and outcomes in the face of rapidly changing wildfire risk, we first develop measures of daily exposure to ambient PM$_{2.5}$ from wildfire smoke, as this exposure itself is not directly measured by existing pollution sensors. To do so we combine long time series of regulatory ground monitor data on PM$_{2.5}$ concentrations from US Environmental Protection Agency (EPA) pollution monitors with satellite-derived estimates of smoke exposure. To isolate wildfire-smoke-derived PM$_{2.5}$ from other sources of PM$_{2.5}$, we define smoke PM$_{2.5}$ as location-, month-, and period-specific anomalous PM$_{2.5}$ on days in which satellites indicated that smoke was overhead (Methods). Resulting ambient smoke PM$_{2.5}$ exposure data cover 970 US counties that contain ~80% of US population (Fig S2), and display wide spatial and temporal variability (Fig 1). Satellite-based counts of the annual number of dense smoke plumes overhead have trended upward in the last decade throughout most of the continental US, particularly in the West (Fig 1a), which we estimate have helped drive rapid increases in estimated annual smoke PM$_{2.5}$ across the country and in days with extreme PM$_{2.5}$ due to smoke (Fig 1b-c).

We merge these daily wildfire-smoke-derived data with multiple high-frequency datasets that are measured at population scale and shed light on individuals’ knowledge, beliefs and behaviors during wildfires. To study individuals’ awareness of their exposure, we first analyze location-specific variation in search query behavior related to smoke exposure. We use public data on specific search queries (e.g. ‘air quality’) from Google Trends, which provides normalized data on search term popularity at the weekly level across “designated market areas” (roughly, groups of counties; see Methods). We interpret a purposeful search for information related to smoke exposure as evidence that an individual is aware she is being exposed and that her exposure level is worth learning about – what we call “salience”.

Second, we study individuals’ preferences and sentiments regarding wildfire smoke exposure. Such preferences underlie standard theoretical models of choice behavior but are hard to observe directly and at scale. Past work has shown that social media behavior can be a sensitive and ac-
accurate tool for understanding individuals’ preferences and sentiments toward what is happening around them (25), including a changing environment (26, 27). Following this earlier work, we analyze ~1.7 billion georeferenced Twitter updates (“tweets”) posted since 2016 using natural language processing algorithms that extract information on the sentiment revealed in each tweet (28) (Methods). This approach has been validated at population scale against self-reported measures of emotional state (25), and complements earlier work that used Twitter to directly measure wildfire activity (29) and infer smoke concentrations (30).

Third, we again use Google search queries to study whether individuals sought information regarding specific health-protective actions, analyzing item-specific search terms like “air filter”, “air purifier”, or “smoke mask”. While we can only observe whether individuals eventually purchased these items in the case of PurpleAir monitors as described below, such search behavior can be interpreted, at a minimum, as evidence of an individual’s belief that health-protective options exist. Evidence from other settings suggests that search activity is predictive of future behavior, including consumer purchases (31, 32).

Fourth, we use smartphone-derived location data to study whether individuals altered their physical movements during periods of smoke exposure. Short-term migration in response to other environmental stress (e.g. hurricanes) is common, and is a plausible avenue by which individuals or households could seek to limit exposure to wildfire smoke. We study both the share of people estimated to be completely at home, and the share estimated to be completely away from their home, on days or weeks of smoke exposure.

We combine each measure with our smoke PM$_{2.5}$ data and analyze the effect of smoke on each outcome using panel fixed effects estimators that exploit local temporal variation in both exposures and outcomes. While long-term exposure to wildfire smoke shows clear spatial patterns and temporal trends (Fig 1), local-level variation in daily exposure is highly random, and panel estimators – which are commonly employed in related environmental settings (2) – plausibly isolate the impact of variation in smoke exposure from other time-invariant and time-varying factors that could be correlated with both smoke exposure and outcomes, including potential confounding from COVID-19 (Methods). To ensure that we are measuring the impact of wildfire smoke and not simply proximity to wildfire itself, we develop measures of distance to the nearest active wildfire and analyze whether responses differ by fire proximity.

Finally, we analyze how ambient outdoor smoke PM$_{2.5}$ infiltrates into the indoor home environment, and whether behaviors and circumstances shape this infiltration. Understanding indoor concentrations is critical, as individuals in the US spend the vast majority of their time indoors.
Using data from the American Time Use Survey, we calculate that Americans spend over 70% of their time indoors at home, with higher shares for lower-income and elderly individuals and overall shares trending up over time (Fig S1). Personal integrated exposure to variation in ambient exposure is then likely substantially mediated by characteristics of home and work environments that are hard to observe (33–35), and these differences could in turn affect outcomes (36–39). If socioeconomic or demographic variables shape indoor environments in ways that affect exposures, as has been hypothesized (37), then exposure levels or policy choices that appear equitable based on traditional outdoor measures could obscure large disparities in realized exposures.

We assemble and harmonize hourly data from 1,520 indoor PurpleAir air pollution monitors that individuals have put in their single-family homes across the US and use nearby outdoor PurpleAir monitors to construct outdoor PM$_{2.5}$ concentrations at each home (Methods). To estimate infiltration, we use distributed lag or lagged-dependent variable panel regression to estimate the marginal increase in indoor PM$_{2.5}$ when outdoor PM$_{2.5}$ increases by one unit (i.e., $\partial \text{Indoor PM}_{2.5} / \partial \text{Outdoor PM}_{2.5}$), controlling flexibly for time of day, day of week, and month of sample (Methods). We estimate models that pool all indoor monitors as well as monitor-specific models, and study how infiltration differs as a function of household and neighborhood characteristics. Our approach complements recent work using PurpleAir to study infiltration generally (40) and during wildfires specifically (33), though our estimation approach offers advantages relative to the latter such as robustness to indoor pollution sources and to diurnal patterns in infiltration-relevant behaviors (Methods).

The timeliness and granularity of novel passive distributed sensor data needs to be weighed against its potential non-representativeness, as the latter can bias population-scale inferences. Our search data and mobility data are likely our most representative, as the vast majority of Americans use the internet regularly and most own and use smartphones. Twitter users are less representative on average, but Twitter-derived sentiment measures have been shown to validate well against population emotional state, and related work shows that the response of sentiment to environmental stress mirrors that measured in representative survey data. PurpleAir data are least representative of our datasets, with wealthier and more educated households more likely to own monitors; however, as discussed below, socioeconomic and demographic information does not appear strongly predictive of infiltration. See Methods for more detailed discussion of sample representativeness.
Results

Increases in wildfire-derived ambient PM$_{2.5}$ exposure lead to an increase in popularity for air quality-related search terms (Fig 2a), with even small increases above zero exposure appearing salient (Fig 2a). Results are robust to alternate air-quality-related search terms and to using analogous search terms in Spanish (Table S3), are not driven by proximity to active wildfires, and are robust to inclusion of weather controls or alternate fixed effects (Table S4). Our results are consistent with interview-based evidence that found that individuals who perceived they were being exposed to smoke often used internet-based sources to confirm their perceptions (19). The salience of ambient exposure at low levels is also somewhat reassuring given recent evidence of health impacts for sensitive populations even at very low levels of ambient exposure (18, 41).

We find that exposure to ambient smoke PM$_{2.5}$ makes people unhappier, as measured in an automated sentiment analysis of 5 years of tweets across the US. Expressed sentiment in tweets declines roughly linearly above smoke PM$_{2.5}$ exposures of $\sim$20\(\mu g/m^3\) (Fig 2b). A very bad smoke day (average smoke PM$_{2.5}$ concentration of 100\(\mu g/m^3\)) is associated with a roughly 0.2 standard deviation decline in sentiment relative to the overall sample standard deviation. For context, the average difference between tweet sentiment on Wednesdays and Saturdays (respectively, the lowest- and highest-sentiment days of the week in our data) is about 0.007 in our data; one day of very bad smoke (100\(\mu g/m^3\)) is thus about 3 times worse in sentiment terms than replacing an average Saturday with an average Wednesday.

Negative effects of smoke on sentiment could occur through a variety of channels, including from fear or anxiety about proximate fires themselves or about what the fires represent (e.g. a changing climate), from unhappiness due to disruption in normal activities (e.g. school closure or inability to recreate), or from anticipation or experience of negative health impacts. While we cannot distinguish the latter channels in our data, effects of smoke on sentiment are not driven by proximity to active wildfire, and are robust to temperature and rainfall controls and to alternate fixed effects (Table S5).

Exposure to smoke PM$_{2.5}$ increases search activity related to protective behavior. Searches for technologies known to help limit exposure, including “air filter”, “air purifier”, "smoke mask", and "PurpleAir", all increase on days in which smoke exposure is higher (Fig 2c, Table S6). Some search queries in Spanish ("purificador de aire") respond similarly, although others ("filtro de aire") do not (Table S6).

Finally, smoke PM$_{2.5}$ exposure on average causes more people to not leave their home, with im-
mediate increases at low levels of exposure that flatten off at high levels (Fig 2d). A day of smoke exposure above 50µg/m³ leads to a roughly 3 percentage point increase in the proportion of people fully at home, which corresponds to about a 10% increase above the mean. Smoke PM$_{2.5}$ exposure has a limited effect on the proportion of people fully away from their home at low exposure levels, but an increasing effect at higher exposure levels (Fig S4). Both results can be interpreted as protective behavior: during heavy smoke days, many individuals shelter in their homes, and some leave the area when exposure gets severe. Both results are robust to controls and are not driven by proximity to active fires; the effect on percent of people at home is less robust to the addition of more stringent time controls (Table S7).

**Exposure and response heterogeneity** Individuals are likely to respond to environmental exposures in different ways, either because their personal exposure varies or because, for a given exposure, their knowledge of that exposure or their willingness or ability to respond to it could differ. We explore heterogeneous exposures and responses to wildfire smoke as a function of socioeconomic status (as measured by locality-specific median household income) and variation in average previous exposure to wildfire or other PM$_{2.5}$. Previous literature suggests both could moderate behavioral responses to environmental stress through a variety of mechanisms, including through differential access to information about exposure risk or differential ability, motivation, or knowledge of how to take protective action (8, 20, 42).

Consistent with earlier work (10), but in stark contrast to strong socioeconomic and ethnic/racial gradients in exposure to other key pollutants in the US (43, 44), we find that exposure to both average and acute smoke PM$_{2.5}$ is largely uncorrelated with income in the US (Fig S3). We also find no differences in salience of smoke exposure between lower- and higher-income counties, with similar responses of search query activity to a day of heavy smoke across income levels (Fig 3a).

Other behavioral measures show strong income gradients. For sentiment, wealthier counties respond much more negatively to a heavy smoke day than less wealthy counties (Fig 3b). This finding is not driven by average differences in sentiment between more and less wealthy counties, by higher overall variation in sentiment in wealthier versus poorer counties (temporal variation in sentiment is lower in wealthy counties than in less-wealthy counties in our sample), or by differences in average exposure to smoke PM$_{2.5}$ or other sources of PM$_{2.5}$ (Table S9).

Search activity related to protective behavior is also substantially higher in wealthier counties (Fig 3c), and not statistically different from zero in roughly the bottom third of the county income distribution. Finally, populations in wealthier counties are also substantially more likely to remain fully at home during a day of heavy wildfire smoke exposure than lower-income populations (Fig
3d). Results are robust to more or less restrictive time controls (Fig S5a). We find no meaningful difference across income groups in the proportion of individuals fully away from their house during days of heavy smoke exposure (Fig S5b).

Why do wealthier locations respond differentially to smoke exposure? Measured differences do not appear to reflect differences in exposure information or in overall internet activity, given the consistent response of air-quality related searches across income groups. Rather, responses are consistent with lower incomes constraining choice sets and behaviors, including less flexibility in working from home, fewer resources with which to consider purchasing protective technology, and – regarding the sentiment results – having other more pressing matters to worry about.

We find that behavioral measures are also affected by previous experience with smoke and with other PM$_{2.5}$ sources. An additional smoke day was less salient in locations with higher previous exposure to smoke PM$_{2.5}$, and locations with higher average PM$_{2.5}$ exposure prior to our study period showed smaller declines in sentiment during an additional high smoke day and fewer searches related to health-protective behavior, but were more likely to stay at home when smoke PM$_{2.5}$ was high (Tables S8-S9). Results are consistent with individuals adapting their behavior and beliefs based on repeated exposure, e.g. through investments in health-protective technologies.

**Smoke PM$_{2.5}$ infiltration into indoor environments** We find that Census tracts with PurpleAir monitors tend to be wealthier on average than tracts without monitors (Fig S6), a finding consistent with other analyses (45) and with the income-differentiated search activity for “PurpleAir” and related health-protective technologies found above. Nevertheless, the average income of locations owning indoor monitors varies by 10x across locations, enabling an exploration of the role of income and other demographic factors in shaping exposures among a population with identical access to information on their exposures.

Using a pooled model we estimate that a 1-$\mu$g/m$^3$ increase in outdoor PM$_{2.5}$ is associated with an 0.15 (95% CI: 0.135-0.153) $\mu$g/m$^3$ increase in indoor PM$_{2.5}$ over the next six hours. Estimates are robust to alternate regression models and alternate corrections to the monitor data (Fig S7) and are comparable in magnitude to recently published estimates (40). Estimated infiltration is substantially lower during periods of very high outdoor PM$_{2.5}$ (Fig 4a). However, for a given outdoor PM$_{2.5}$ level, infiltration is higher when at least some of the PM$_{2.5}$ is from smoke. Earlier findings of lower infiltration on smoke days (33) were likely capturing the effect of overall high PM$_{2.5}$ rather than the effect of smoke-derived PM$_{2.5}$ specifically.

Consistent with our other behavioral measures, declining infiltration at high outdoor PM$_{2.5}$ levels
suggests that salient ambient exposures induce behavioral responses, which could include closing windows or doors and/or using mechanical filtration. However, in contrast with our other behavioral measures, we find only a modest relationship between neighborhood average income and infiltration, with households in much wealthier Census blocks experiencing only slightly lower average infiltration than households in areas with one quarter the average income regardless of whether the PM$_{2.5}$ was smoke-derived (Fig 4b).

To further explore predictors of infiltration, we estimate infiltration separately for each of the 1,520 indoor monitors in our dataset, match each monitor to a wide range of house- and neighborhood-specific socioeconomic, demographic, environmental, and housing covariates, and fit flexible machine learning models relating infiltration to these covariates (Methods). Consistent with other work (40), we find many-fold differences in household-specific infiltration rates (mean = 0.19, sd = 0.16; see Fig 4c), and confirm using a Bayesian hierarchical model that this variation is largely due to "true" underlying variation between households rather than to sampling noise in household-level estimates (Fig S8a). Estimates are only modestly correlated with traditional indoor/outdoor ratio estimates (Fig S9), perhaps due to the difficulty in accounting for indoor sources of emissions or diurnal behavioral patterns in the traditional I/O approach (Methods).

While racial/ethnic, socioeconomic, environmental, and housing variables are associated with infiltration on held-out data, their individual explanatory power is very modest, and our rich set of predictors and flexible models are surprisingly poor predictors of overall variation in infiltration, explaining only ~5% of variation across indoor monitors in our data (Fig S8b-c). This lack of predictive ability of socioeconomic factors is also apparent on individual smoke days, where even among relatively socioeconomically advantaged households, very similar outdoor PM$_{2.5}$ concentrations during a given smoke day are associated with widely varying indoor PM$_{2.5}$ concentrations (Fig S10).

To further investigate the differential influence of behavior versus housing characteristics (and associated socioeconomic factors), we re-estimated infiltration for individual households during periods when windows were likely to be closed and indoor filtration not running (Methods). While average infiltration during these periods was relatively similar to infiltration during all periods (Fig S8d), infiltration varies much more strongly with both income (Fig S8e) and housing age (S8f) under these conditions. Taken together, and consistent with previous smaller-scale work (39), our results indicate that the poor explanatory power of socioeconomic and housing characteristics is not driven by poor measurement of these characteristics but by the dominant effect of idiosyncratic household-specific behaviors that are not correlated with these characteristics.
Finally, using indoor monitors across the Bay Area and data prior to Aug 2020, we divide moni-
tors into low (bottom quartile) and high (top quartile) infiltration groups (Fig 4d) and study out-
door and indoor PM$_{2.5}$ levels across these groups during the extreme wildfire smoke event the
area experienced in Aug-Sep 2020. High and low infiltration households experienced nearly identi-
tical daily outdoor concentrations during the many-week event (Fig 4e), but these ambient lev-
els led to starkly different indoor concentrations. On the worst smoke days, daily average indoor
concentrations across all high infiltration homes exceeded 65$\mu$g/m$^3$ and in some houses it ex-
cceeded 100$\mu$g/m$^3$, well above the World Health Organization 24-hour PM$_{2.5}$ exposure guideline
of 15$\mu$g/m$^3$. Low infiltration households were on average able to maintain indoor PM$_{2.5}$ concen-
trations near 5$\mu$g/m$^3$. Across the duration of the smoke event, daily mean indoor concentrations
were on average 3.5x higher in the highest quartile versus the lowest quartile of infiltration house-
holds. Differences were even larger when looking across all Bay Area monitors: households with
average outdoor PM$_{2.5}$ levels within 5$\mu$g/m$^3$ of each other experienced >20x differences in aver-
age indoor PM$_{2.5}$ concentrations during the smoke event (Fig S11).

Discussion

A growing literature documents the large and often disparate impacts of wildfire smoke on a
range of health outcomes (14–18, 21, 22). Our results show how novel sensor data can provide
policy-relevant insight into why the magnitude and incidence of these impacts might vary. Mul-
tiple lines of evidence indicate that awareness of smoke concentrations does not appear to be a
primary constraint on individual behavior in the face of wildfire smoke exposure: even small in-
creases in ambient exposure cause individuals to seek air quality information, become unhappier,
and stay in their homes. But while awareness appears to be broadly shared, it does not lead to ad-
equate health protection. Even among populations that own indoor monitors and who thus have
access to accurate, real-time measures of their indoor concentrations, information is not enough
to limit dangerous indoor exposures to these pollutants. This suggests policies that target infor-
mation provision about smoke are insufficient, and perhaps not central, to enabling protective
behavior.

Socioeconomic status is not correlated with outdoor smoke levels but does appear to mediate be-
havioral responses to such pollution. Wealthier households in our sample can more easily stay
home, are more likely to seek information on protective technology, and are more likely to own
indoor pollution monitors. Such differential behavior is consistent with a broader literature that
shows how socioeconomic status constrains households’ abilities to invest in environmental qual-
ity and health protection (8, 46). Yet, at least among our sample of monitor-owning households, income is only weakly correlated with the infiltration of ambient smoke into indoor environments, and we observe many households in wealthy neighborhoods experiencing exceedingly high levels of indoor smoke exposure.

Our results suggest this is likely because present infiltration rates are dominated by actions like opening windows and doors, not housing materials or quality that might be reflected in prices. Infiltration patterns thus point to the importance of behavior that remains unobserved, a fact which is both encouraging and troubling. If simple but difficult-to-observe behaviors such as closing windows and doors explain the vast majority of variation in smoke infiltration, then reducing infiltration at population scale could be much easier in theory than if infiltration was largely determined by income or housing quality, as changing these latter factors requires addressing deeper societal problems of inequality and structural racism.

On the other hand, current policy approaches to addressing smoke exposure focus on behavioral recommendations to stay at home and close windows and doors (24), but these policies alone are difficult to comply with and may still be inadequate: many households’ indoor environments remain highly exposed, and our mobility results suggest adherence might be difficult for lower-income households. If such behaviors are indeed hard to adopt, then the policy approach of promoting private provision of protection could be biased against disadvantaged groups. This policy approach also stands in stark contrast to the approach of public provision of protection used for other sources of PM$_{2.5}$, which has sought to reduce emissions of pollutants at their source and which has successfully reduced overall ambient exposure inequalities (47). Further understanding the variation and causes of the behaviors that can protect indoor environments will be key to designing policy that can both lower indoor concentrations and not disadvantage certain groups.
Figure 1: **Trends in smoke exposure across the US.** a Measurements from satellites indicate rapidly growing exposure to “heavy” smoke plumes across much of the US. Estimates are shown on a 10km grid and indicate the estimated annual increase between 2011 and 2020 in the number of smoke plumes that NOAA analysts designate as “heavy”, their densest plume classification. Dots indicate EPA ground-based pollution monitors. b Distribution of estimated annual average smoke PM$_{2.5}$ across EPA ground monitors reporting in each year. c Distribution of estimated daily smoke PM$_{2.5}$ across the same monitors. Recent increases in extreme annual exposure are being driven by increases in extreme daily exposure.
Figure 2: **Behavioral responses to wildfire smoke exposure.**

a. Searches for ‘air quality’ on Google at the US designated marketing area by week level, 2016-2020. Search index is normalized such that zero indicates no searches and 100 indicates the maximum number of weekly searches over the period. 

b. Average sentiment in Twitter updates (‘tweets’) at the county-day level, 2017-2020. 

c. Searches for ‘air filter’ on Google at the US marketing area by week level, 2016-2020. Search index normalized as in (a). 

d. Percent of mobile phones estimated to be completely at home on a given day at the US county level, 2019-2020. Lines are spline fits conditional on fixed effects, with shaded areas showing bootstrapped 95% confidence intervals. Number of observations in each regression is shown in upper left corner of each panel. Histograms at the bottom show the log distribution of smoke PM$_{2.5}$ exposure in each sample.
Figure 3: **Smoke salience does not differ, but other responses do differ, among more and less wealthy populations.**

- **Salience**
  - The effect of a heavy smoke exposure (50μg smoke PM$_{2.5}$ on that day) on searches for ‘air quality’ on Google do not differ by income.

- **Sentiment**
  - A heavy smoke week has a stronger negative effect on sentiment among wealthy populations.

- **Health protection**
  - Wealthier populations are substantially more likely to search for ‘air filter’ on Google during a heavy smoke day.

- **Mobility**
  - Wealthier populations are more likely to be completely at home during a week of heavy smoke.

Lines show the marginal effect of a 50μg smoke day on the outcome at different income levels, with shaded areas showing bootstrapped 95% confidence intervals. The slope of each line is shown in upper left, and the histograms at bottom show the distribution of county incomes in each sample. Sample in each panel is same as corresponding panels in Fig 2.
Infiltration rates decline strongly with outdoor PM$_{2.5}$ concentration during both fire and non-fire periods, but not with income, and differ greatly across households resulting in extreme differences in indoor exposure during wildfires. a Infiltration rates, measured as the integrated hourly change in indoor PM$_{2.5}$ per unit increase in outdoor PM$_{2.5}$, are lower during periods of high outdoor PM$_{2.5}$ regardless of whether PM$_{2.5}$ comes from smoke or other sources. b Infiltration rates are only weakly correlated with census block median income. c The distribution of infiltration rates estimated separately for each household in the sample is wide. d Residents from the highest and lowest infiltration quartiles in the Bay Area are mapped and colored according to group, showing the groups are geographically intermixed. e Daily outdoor PM$_{2.5}$ concentrations during unprecedented Aug/Sep 2020 wildfire smoke event were highly similar between high and low infiltration households. Transparent lines show concentrations at each residence and thick lines show the averages within each group. The left panel shows daily mean concentrations and the right panel shows averages across all days during the event. f Same as e but for indoor PM$_{2.5}$ concentrations, showing very large (>100μg) daily differences during peak outdoor concentrations (left) and 3.5x differences in average indoor exposure between high and low infiltration groups (right).
Methods

**Estimating ground PM$_{2.5}$ concentrations from smoke**  We develop a generic, tractable method for estimating ground PM$_{2.5}$ attributable to smoke at the daily level. The method requires a credible estimate of whether there is smoke in the air on a given day, and a daily time series of PM$_{2.5}$ from which location- and period-specific anomalies can be constructed. In principle, any available (accurate) daily PM$_{2.5}$ estimates could be used, including promising recent machine-learning-based efforts at generating high-resolution gridded time series of PM$_{2.5}$ concentrations (48, 49). However, existing gridded data are not available for recent years, and so we instead use station-based daily PM$_{2.5}$ measures from the network of thousands of EPA stations across the continental US (Fig 1a).

To construct our daily measures of smoke PM$_{2.5}$, we define $PM_{idmy}$ as the PM$_{2.5}$ concentration recorded by the EPA monitor at location $i$ on day $d$, month $m$, and year $y$. From this time series we construct location- and month-specific anomalies $PManom_{idmy} = PM_{idmy} - PM_{mmy}$, where $PM_{mmy}$ is monthly median PM$_{2.5}$ on non-smoke days at that location, and where median is defined over the three years surrounding the year of interest. So for example, a PM$_{2.5}$ anomaly for the Redwood City CA EPA station on Jan 10, 2019 is calculated as the value on Jan 10, 2019 minus the median PM$_{2.5}$ value on all January days in 2018, 2019, and 2020 in Redwood City when smoke was not overhead. Our measure of whether smoke was overhead, $plume_{idmy}$, is derived from the National Oceanic and Atmospheric Administration (NOAA) Hazard Mapping System (HMS) satellite estimates of smoke plume boundaries. We define $plume_{idmy} = 1$ if there was a smoke plume of any thickness over location $i$ during any time on day $d$, and zero otherwise. Finally, from these data we can construct $SmokePM_{idmy} = PManom_{idm} * plume_{idmy}$. $SmokePM_{idmy}$ will thus equal zero when there is no plume overhead, and will equal the anomaly value when there is smoke overhead. Our approach thus provides a continuous measure of smoke exposure intensity.

Our approach is similar to recent work (50) using interpolated station data and plumes to estimate smoke PM$_{2.5}$. However, given the high spatial variation in smoke exposure and often large distance between EPA stations, we chose not to interpolate EPA stations. Given the secular decline in background PM$_{2.5}$ across the country, we also define anomalies relative to a moving (3-year) monthly median rather than a fixed seasonal mean.

**Measuring salience and health-protective behavior**  We measure salience and health-protective behavior using public search query data from Google Trends. Data are accessed using the R package gtrendsR, and are provided as location-, term- and period-normalized indices ranging
from 0 to 100, where 0 is the lowest search volume for that term in that location during the chosen period, and 100 is the highest search volume. Data are available at the "Designated Market Area" level (referred to as "Metro" areas by Google Trends), which are geographic regions encompassing television media markets as defined by Nielsen.

We study searches in both English and Spanish, which together are the primary languages spoken by 92% of US households (51). We use weekly data on designated market areas, the native spatial resolution of the public Trends data, between Jan 2016 and Dec 2020, and analyze data on terms related to smoke exposure (including “air quality”, "smoke", and "wildfire smoke").

**Measuring sentiment**  We measure online sentiment for a county-day using the text of Twitter posts (”tweets”) created in that county on that day. Specifically, we collect nearly all of the geolocated tweets for the continental United States between December 2016 and February 2021 through the Twitter Streaming API. To compute sentiment for each tweet, we apply the VADER sentiment analysis model (28), a natural language processing algorithm tuned specifically for estimating sentiment from online language. We take the average of the “compound” scores (ranging between -1 and 1) computed by VADER for all tweets in a county-day as our measurement of sentiment. Our approach builds on the computation of expressed sentiment described in ref (26). Readers may refer to that article for additional details on the general approach to collecting and processing tweets for use in empirical analysis. On average, mean sentiment for a county-day is 0.17, computed from 455.4 tweets.

**Measuring mobility**  We assembled a daily dataset of mobility measures at the county-level collected between Jan 2019-Dec 2020, the period over which mobility data were generously made available to researchers by Safegraph. These data measure the aggregate activity of anonymized device signals, or “pings”, at the census block group-level. Signals are collected from smartphones, not all cell phones. We focus on two measures constructed from these anonymized signals: percentage of individuals completely at home on that day, and percentage of individuals completely away from home on that day. We construct the *completely away from home* variable by counting the percentage of devices on a given day that were not observed in their respective home location. Safegraph assigns a *home* location to each device based on their mobility pattern observed over the previous six weeks. We aggregate these data to the county-day level by taking means weighted by the number of devices in each census block. Data processing details are discussed further in refs (52–54).

**Measuring distance to fire**  To distinguish the effects of exposure to wildfire smoke from potentially correlated effects of being near an active wildfire, we develop daily measures of proximity to active wildfires and control for them in the regression analysis. Our measures of active fires
are derived from NOAA HMS fire points. Because these fire detections often indicate very small and isolated fires that likely do not represent threatening wildfires, and large active fires typically have many nearby fire detections, we build on earlier work (10) and denote active fires by constructing "fire clusters", or locations with multiple active nearby fire points. To do this, we buffer each HMS fire point by 3km square, and take the union of existing overlapping squares over a given day and previous three days. We then define an active fire as a location where this union is >10km². This does not mean 10km² are burning, but within that 10km² there are multiple fire points over a three day period, representing an active and potentially growing fire. We then compute "distance to fire" as the distance from a geography centroid (e.g. county) to the nearest active fire, setting this distance to 0 if the active fire cluster is inside the county on that day.

**Estimating ambient smoke impact**  We combine the above behavioral measures with our smoke PM_{2.5} estimates and analyze their correspondence using panel fixed effects estimators, with the goal of isolating the impact of variation in smoke exposure from other time-invariant and time-varying factors that could be correlated with both smoke exposure and outcomes. Specifically, we estimate econometric models of the form:

\[ y_{isdmy} = f(SmokePM_{ismy}) + Z_{ismy} + \alpha_{im} + \eta_d + \epsilon_{isdmy} \]  

where \( y_{isdmy} \) is outcome of interest in unit \( i \), state \( s \), day \( d \), month \( m \), and year \( y \), \( SmokePM_{ismy} \) is our smoke PM_{2.5} measure on the same day and location, and \( Z_{ismy} \) are additional time-varying controls. Our preferred model includes location-by-month FE to account for local seasonality in either outcomes or exposures (i.e. one intercept for each of the twelve months in Santa Clara County, CA), day-of-sample FE (e.g. a dummy for Jan 1, another for Jan 2 2016, etc) to account for common trends or shocks to outcomes or exposures on a given day. Our date FE implicitly also account for any average differences in outcomes between weekends and weekdays. We estimate \( f() \) using either linear models or more flexible cubic splines to capture potential non-linearities. In all analyses using search query data, Twitter data, or mobility data, smoke PM_{2.5} is measured using EPA station data, as described above.

In these models, the effect of smoke exposure on outcome \( y \) is estimated by relating, e.g., outcomes in Santa Clara County on Aug 30 2020 versus Sep 1 2020 to differences in smoke exposure on those days, after accounting for any common difference across counties in exposure or outcomes between the two days, and any average differences in smoke exposure or outcomes in August vs September in Santa Clara County. A confounding variable would have to be a local time-trending unobservable correlated with both smoke exposure and the outcome. Possible candidates include weather variables and the presence of an active wildfire nearby, and we addi-
tionally control for these variables ($Z_{ismy}$ in Equation 1) in robustness tests, or split the sample between locations nearby and further from an active wildfire.

Another potential threat to identification is the COVID epidemic, which near the end of our sample period had demonstrated effects on mobility (55) and sentiment (56), and likely enhanced awareness about the importance of air filtration (57); 2020 was also a year of severe smoke exposure throughout much of the US West. Because we exploit daily variation in smoke exposure over time at particular locations, and because such variation depends largely on stochastic factors such as exactly where fires ignite and which way the wind is blowing, we believe daily variation in COVID outcomes or behaviors are unlikely correlated with wildfire smoke exposures. However, to further address this confounding risk, we test robustness to even more stringent time controls, including county by month of sample FE, and state by day of sample FE; these further account for any state-specific differences or trends in COVID severity and/or policy intervention that happened to coincide with wildfire risk. We note that any changes in our observed behavioral outcomes due to wildfire-specific effects on health outcomes, including wildfire’s potential effects on COVID itself (17), are not confounding and would constitute part of the overall "effect" that we wish to understand.

To study whether the effects of smoke on outcomes vary across locations, we interact smoke exposure with time-invariant covariates of interest:

$$y_{isdmy} = \beta_1 SmokePM_{ismy} + \beta_2 SmokePM_{ismy} * X_i + Z_{ismy} + \alpha_{im} + \eta_d + \epsilon_{isdmy}$$

(2)

where $X_i$ in our analysis include median household income, average previous exposure to PM$_{2.5}$, and average smoke PM$_{2.5}$ exposure, either included individually or jointly.

**Measuring indoor and outdoor household PM$_{2.5}$ using Purple Air** To estimate household infiltration of outdoor PM$_{2.5}$ into indoor environments, we utilize data collected by low-cost Purple Air monitors measuring indoor and outdoor PM$_{2.5}$ concentrations. Raw 10-minute observations were downloaded for all available indoor and outdoor Purple Air monitors in the contiguous United States and then cleaned according to the procedures developed by recent studies (38, 40) in order to produce hourly indoor and outdoor PM$_{2.5}$ concentrations. Namely, for each monitor, we download 10-minute observations from the earliest available date through the end of 2020 or the last available date, whichever is later.

Following existing literature, we use multiple approaches for estimating PM$_{2.5}$ concentrations from Purple Air data. In a “particle count” approach, we derive 10-min PM$_{2.5}$ concentrations based on the reported particle counts reported by each sensor (i.e., for dual sensor monitors two
different PM$_{2.5}$ concentrations were derived for each 10-min period, one for each sensor). PM$_{2.5}$ concentrations were derived from particle counts using equations 1 and 2 from Bi et al 2021 (38). For the other approaches that relied directly on Purple Air’s reported concentrations we record the 10-min average PM$_{2.5}$ reported in the cf$_1$ channel. We choose the cf$_1$ channel rather than the cf$_{atm}$ channel because previous work has shown it correlates better with concentrations measured by EPA monitors (58). Furthermore, because Purple Air previously had the cf$_1$ and cf$_{atm}$ columns incorrectly labeled in their data, we use the updated labels for each channel to ensure we have correctly identified the cf$_1$ observation. Then, for all approaches the following steps were used to go from the 10-minute Purple Air PM$_{2.5}$ concentrations to the final cleaned hourly data used in the analysis:

1. For monitors with A and B sensors, we compare the two sensors and set any 10-min observation to NA if the difference between sensors exceeds a certain threshold. Specifically, when PM$_{2.5}$ $\leq 100$ $\mu g/m^3$ we set the observation to NA if the A and B sensors differ by more than 10 $\mu g/m^3$ and when PM$_{2.5}$ $> 100$ we set the observation to NA if the A and B channels differ by more than 10%. We then take the average across the two sensors. Observations where either sensor has a missing value are set to missing.

2. Next, we examine all adjacent three period runs to see if there are ‘spikes’ defined as an increase and then immediate decrease in the subsequent period or a decrease and then subsequent increase in the next period. If the mirror jumps in subsequent periods from t-1 to t and from t to t+1 are both over 100 then we set observation in time t to NA.

3. To get hourly observations we take the 60-min mean, setting the hourly average to NA if any 10-min period was missing.

This process resulted in two separate measures of PM$_{2.5}$ concentrations for each monitor, one based on particle counts and one based on Purple Air’s reported PM$_{2.5}$ concentrations. We then derived a third measure of PM$_{2.5}$ concentrations by using the Purple Air reported concentrations (i.e., not those derived from the particle counts) and applying the correction developed in Barkjohn et al 2021 (58) to improve agreement between Purple Air PM$_{2.5}$ concentrations and EPA reference monitor concentrations.

Hourly ambient exposures were estimated at each indoor monitor site by first identifying all outdoor monitors within 5km and then taking the inverse distance weighted average of hourly PM$_{2.5}$ concentration across the (up to) 10 nearest monitors. Monitors with less than 720 non-missing
hourly indoor and outdoor PM$_{2.5}$ measurements (i.e., 30 days of hourly data) were excluded from the analysis.

Finally, indoor Purple Air monitors are deployed in many different types of buildings. We used a combination of information from monitor labels and manually checking of geolocations to determine which buildings with indoor Purple Air monitors were single family residences (SFR). All other types of buildings were removed from the sample. In total, there were 1,520 monitors reporting in our sample of single family residences.

In order to validate the three versions of Purple Air derived PM$_{2.5}$ concentrations we identify the nearest outdoor Purple Air monitor to each available EPA monitor that reports hourly data and we estimate the following regression:

$$PM_{ihm}^{EPA} = \beta PM_{jhm}^{PA} + \alpha_i + \theta_h + \eta_m + \epsilon_{ihm}$$ (3)

where $PM_{ihm}^{EPA}$ is the EPA reported hourly PM$_{2.5}$ concentration for monitor $i$ in hour-of-day $h$ and for month-of-sample $m$ and $PM_{jhm}^{PA}$ is one of the three versions of hourly PM$_{2.5}$ concentration estimated at the nearest Purple Air monitor. Our infiltration measure derived from these data is a regression based estimate that relates anomalies (i.e. deviations from averages along several dimensions) in indoor PM$_{2.5}$ to anomalies in outdoor PM$_{2.5}$. Therefore when evaluating performance of the Purple Air PM$_{2.5}$ measures with respect to EPA monitors, we include monitor, hour-of-day, and month-of-sample fixed effects in the comparison regression.

Results from estimation of the EPA monitor comparison regressions are shown in Table S1. Anomalies in all three versions of Purple Air PM$_{2.5}$ measures are strongly correlated with anomalies in hourly measures from nearby EPA monitors with R$^2$ values in each case $>$0.65. The uncorrected Purple Air cf$_{-1}$ PM$_{2.5}$ measures overestimate EPA PM$_{2.5}$ and explain the smallest share of the variation in anomalies. Unsurprisingly, given the Barkjohn et al correction was designed specifically to calibrate Purple Air PM$_{2.5}$ to EPA PM$_{2.5}$ levels, it is the best of the three measures at predicting the level of EPA monitor PM$_{2.5}$. However, while the particle count derived PM$_{2.5}$ concentrations underestimate EPA monitor concentrations, they explains the largest share of the variation (R$^2$=0.72). This is consistent with a recent study on Purple Air monitor PM$_{2.5}$ calibration that found PM$_{2.5}$ concentrations derived from Purple Air particle counts best explained EPA measured PM$_{2.5}$ anomalies both overall and specifically during wildfire events (59). In light of these findings, we follow several recent publications (38, 40, 59) and present our main infiltration results with PM$_{2.5}$ concentrations derived from Purple Air particle counts. However, we also

21
present infiltration estimates for all three versions of PM$_{2.5}$ concentrations (in each case using the same measure for indoor and outdoor concentrations) and find estimates are highly correlated across choice of measure (Fig S7).

**Estimating infiltration rates** In order to estimate the average indoor infiltration rate, which we define as the increase in indoor PM$_{2.5}$ concentration per unit increase in local outdoor PM$_{2.5}$ concentration (i.e., $\partial \text{Indoor PM}_{2.5}/\partial \text{Outdoor PM}_{2.5}$), we estimate a regression at the monitor-hour level. Namely, for each residence $i$ in hour $h$ on day-of-week $d$ and month-of-sample $m$, we estimate how indoor PM$_{2.5}$ varies with contemporaneous and previous hour measurements of outdoor PM$_{2.5}$:

$$PM_{ihdm}^in = \sum_{k=0}^{6} \beta_k PM_{i;dh-k;dm}^{out} + \gamma_i + \delta_h + \eta_d + \theta_m + \varepsilon_{ihdm}$$

(4)

To isolate the contribution of outdoor PM$_{2.5}$ to indoor PM$_{2.5}$ from other time varying PM$_{2.5}$ sources (most notably, indoor-sourced PM$_{2.5}$), we use fixed effects to flexibly control for monthly trends in PM$_{2.5}$ over the sample ($\theta_m$) and household-specific average variation in PM$_{2.5}$ within the day ($\delta_h$). Day-of-week fixed effects ($\eta_d$) control for differences in patterns between weekdays and weekends.

We include six lags here (outdoor PM$_{2.5}$ at each of the previous 6 hours) to account for lingering effects of outdoor concentrations in previous hours on contemporaneous indoor concentrations, although results are robust to the inclusion of additional lags. From this regression we derive an estimate for outdoor-indoor infiltration by calculating the cumulative effect of 1 $\mu g/m^3$ increase in outdoor concentrations on indoor concentrations:

$$\text{infiltration} = \sum_{k=0}^{6} \beta_k$$

(5)

In order to assess the importance of modeling structure we re-estimated Eq. 4 with four different lag structures: a distributed lag model with lags for outdoor PM$_{2.5}$ only (shown above), a lagged dependent variable model with a lag for indoor PM$_{2.5}$ only, a model with both indoor and outdoor PM$_{2.5}$ lags, and finally a model with no lag terms (Table S2). Infiltration rate estimates derived from each of the models are highly similar (Fig S7) and models with more than six lags produce indistinguishable estimates of infiltration rates.
In order to examine heterogeneity in infiltration rates across hourly outdoor pollution levels and by smoke presence, we first estimate a non-linear version of Equation 4. Namely, we model indoor PM$_{2.5}$ as a 4th degree polynomial of outdoor PM$_{2.5}$ (and its lags) and interact it with a dummy variable indicating whether smoke was present. The smoke dummy is defined as 1 when a NOAA HMS plume reported a heavy plume over the PurpleAir monitor on that day and 0 otherwise:

$$PM_{ih}^{in} = \left[ \sum_{k=0}^{6} \beta_k PM_{i,h-k;dm}^{out} + \sum_{k=0}^{6} \alpha_k (PM_{i,h-k;dm}^{out})^2 + \right. \left. \sum_{k=0}^{6} \nu_k (PM_{i,h-k;dm}^{out})^3 + \sum_{k=0}^{6} \lambda_k (PM_{i,h-k;dm}^{out})^4 \right] \cdot 1(\text{smoke} = 1) + \left[ \sum_{k=0}^{6} \beta_k PM_{i,h-k;dm}^{out} + \sum_{k=0}^{6} \alpha_k (PM_{i,h-k;dm}^{out})^2 + \right. \left. \sum_{k=0}^{6} \nu_k (PM_{i,h-k;dm}^{out})^3 + \sum_{k=0}^{6} \lambda_k (PM_{i,h-k;dm}^{out})^4 \right] \cdot 1(\text{smoke} = 0) + \gamma_i + \delta_h + \eta_d + \theta_m + \varepsilon_{ihdm}$$

To measure the infiltration rate we then calculate the derivative of indoor PM$_{2.5}$ with respect to outdoor PM$_{2.5}$ estimated in Equation 6 and evaluate across the 1st-99th percentile of observed hourly outdoor PM$_{2.5}$ concentrations as well as the indicator for whether or not smoke was present. Responses are plotted in Figure 4a.

We also estimate infiltration rates as a function of median census tract income by estimating Equation 4 with an additional income interaction term:

$$PM_{ih}^{in} = \left[ \sum_{k=0}^{6} \beta_k PM_{i,h-k;dm}^{out} \right] + \left[ \sum_{k=0}^{6} \beta_k PM_{i,h-k;dm}^{out} \right] \cdot income + \gamma_i + \delta_h + \eta_d + \theta_m + \varepsilon_{ihdm}$$

Median income data come from the American Community Survey. Each indoor monitor was matched to a census tract and median income was pulled for the most recent available year and updated to 2020 dollars.

Finally, for each indoor monitor we estimated a separate distributed lag model analogous to the
pooled model in Equation 4:

\[
PM_{ihdm} = \sum_{k=0}^{6} \beta_{ik}PM_{j,h-k;m} + \delta_{h} + \eta_{m} + \theta_{d} + \varepsilon_{ihdm}
\]  

(8)

where PM$_{2.5}$ at indoor monitor $i$ in hour $h$ on day-of-week $d$ and month-of-sample $m$ is modeled as a function of outdoor PM$_{2.5}$ in that location in the contemporaneous period and for each of the previous 6 hours. Our estimate of the overall infiltration rate for each monitor, which we denote $\beta_i$ is then the sum of coefficients over time from the regression for that monitor (i.e., $\beta_i = \sum_{k=0}^{6} \beta_{ik}$).

A large number of previous studies [e.g., (33, 35, 38, 39) among many others] have focused on average indoor-to-outdoor concentration ratios (I/O) as a metric of infiltration. If measured over long time periods, this average I/O ratio does represent the net atmospheric barrier provided by a building, but is inclusive of emissions from indoor sources (which may be either trapped or ventilated) (60). Because indoor-sourced emissions can be large, approaches using I/O to understand infiltration have often focused on periods of the day (typically the middle of the night) when indoor sources are likely to be minimal. However, if households alter other infiltration-relevant behaviors at night (e.g., either keeping windows open, or keeping them closed), then period-specific I/O estimates will not be representative of infiltration during the rest of the day. These differences could be correlated with demographics. For instance, households in higher-crime areas might close their windows at night but open them during the day, and wealthy households with air conditioning in low crime areas could do the reverse; in this case, I/O estimates would underestimate true average infiltration for the first set of houses and overstate for the second.

Our regression-based approach to infiltration measurement is less likely to suffer from this problem. We characterize infiltration as $\partial I/\partial O$, or the change in indoor concentrations when outdoor concentrations change, and use a regression framework to help isolate the contribution of outdoor PM$_{2.5}$ to indoor PM$_{2.5}$ from time-varying indoor PM$_{2.5}$ sources. Indoor emissions sources will not bias $\partial I/\partial O$ estimates so long as variation in these sources (after conditioning out our rich set of within- and across-day fixed effects) is uncorrelated with outdoor sources, which is plausible. As a result, our regression based estimates represent average infiltration rates throughout the whole day, and are less likely to be biased by remaining sources of indoor emissions.

Our analysis suggests this distinction in measurements is especially important over shorter time scales, and during major perturbations in outdoor pollution, like wildfires (Fig S9). A simple takeaway of this study is thus that I/O ratios might not adequately capture concentrations when
populations are trying to stay indoors during a heavy smoke event.

**Understanding variation in household infiltration rates**  Monitor-specific estimates suggest large variation in infiltration across households (Fig 4c), consistent with earlier work (40). However, since monitor-specific infiltration values are themselves estimates from data, observed variation across monitors could reflect either "true" underlying heterogeneity in infiltration, or could simply reflect sampling variation (or some combination of the two).

To distinguish sampling variation from underlying heterogeneity, we estimate a Bayesian hierarchical model (61, 62) that models monitor-specific infiltration estimates as being distributed normally about true monitor-specific infiltration values with estimated monitor-specific sampling variance $\hat{\sigma}_i^2$ and true monitor-specific infiltration values as drawn from an underlying normal distribution with unknown mean and variance $\beta_i \sim N(\beta, \sigma^2)$.

We obtain posterior monitor-specific infiltration values $\beta_i$ by assuming uniform prior distributions for $\beta$ and $\sigma$ and applying Bayes’ Rule to update our estimates of $\beta$, $\sigma$, and the $\beta_i$’s. We use 4 chains each with 2,000 total iterations and 1,000 warm-up iterations (resulting in 4,000 post-warmup draws) to fit the model and characterize the posterior distributions of these variables using R package *rstan* (63). For this analysis, we limit our eligible sample to the 1,484 indoor monitors associated with infiltration estimates between 0 and 1 across all PurpleAir PM$_{2.5}$ corrections and model lag structures. Posterior estimates of the $\beta_i$’s are very similar to “raw” regression based estimates from equation 8, suggesting nearly all the variation is true underlying variation in infiltration, and not due to sampling variation.

To understand underlying heterogeneity in infiltration, we then match each indoor monitor to a range of house- and neighborhood-specific covariates measuring demographics, neighborhood conditions, environmental conditions, and housing characteristics. We download demographic and neighborhood condition covariates at the Census tract level from the 2019 American Community Survey 5-year estimates via the R package *tidycensus*. We match each monitor to its overlapping or nearest neighboring Census tract.

Among environmental covariates, mean annual heating degree days and cooling degree days over 2000-2020 are obtained via Google Earth Engine (64) from the PRISM Daily Spatial Climate Dataset (AN81d) (65) at tract level and matched with monitors by overlap or nearest neighbor. Each household’s 10-year mean outdoor PM$_{2.5}$ concentration is calculated from gridded annual PM$_{2.5}$ estimates (66).

We interpolate housing covariates from georeferenced CoreLogic data on household character-
istics spanning the universe of residential homes in the US (67). Because individuals geolocate their monitors manually on the PurpleAir platform and might not choose to input a location at their precise residence, we match monitors to CoreLogic single-family homes in four ways, taking the inverse distance weighted average of housing characteristics within 100, 250, or 500 m, or by using values at the nearest neighbor within 1 km (i.e. the residence in CoreLogic closest to the monitor lat/lon). For consistency, the sample utilized in the analysis is limited to the 1,383 indoor monitors that match with a CoreLogic house within 100 m, regardless of interpolation method. Because housing characteristics tend to be highly correlated, we construct a housing index by averaging standardized values of home value, number of stories, number of baths, number of bedrooms, height, and area, and use the index as a predictor.

We then train flexible ML-based models to predict monitor-specific infiltration rates from matched covariates. We divide our sample into a 75% training dataset and a 25% held-out test dataset, splitting train and test within 13 disjoint geographic regions covering the contiguous US to ensure a geographically balanced split. We train random forest and gradient boosted trees models with manually tuned forest and boosting hyperparameters, respectively, and tree parameters tuned using random search with 3-fold cross-validation repeated 5 times. We conduct tuning and training for each method of matching monitors and CoreLogic houses for robustness.

We report performance statistics ($R^2$) on held-out test data, and compute the marginal effect of each predictor by evaluating the predicted effect in the test data of moving from the 5th percentile to the 95th percentile of the predictor, with all other variables fixed at their mean value. We repeat this evaluation for both random forest and gradient boosted trees models, and for all four ways of spatially matching to housing characteristics.

**Understanding representativeness of study samples** Our novel social media, search trends, and sensor data data are not necessarily derived from representative samples of individuals. Our inferences about behavioral responses to smoke that we draw from these data will thus also not be representative of the population at large if the behavior of those represented in the data in response to smoke exposure differs from the behavior of those not in the data. While the latter is by definition hard to observe, some progress can be made by understanding which individuals and groups are contributing information to these datasets, and by studying other efforts to validate derived behavioral measures against representative data.

Recent surveys of overall internet use suggest that the vast majority (>90%) of adult Americans use the internet, with limited differences by race and only moderate differences by income (68). While we are unaware of comprehensive demographic data on search activity, previous work has found that search-based measures of issue salience correlate reasonably well with survey-based
measures (69) and are predictive of a range of economic phenomena (70). In our data, the fact that we see higher income populations more likely to both search for and purchase (in the case of PurpleAir) health protective technologies is additional validation that income-differentiated search behavior predicts real-world behavior.

Regarding our sentiment analysis, roughly a quarter of American adults are estimated to use Twitter, with limited differences in use by race (slightly higher use by non-White populations) but higher use among younger Americans and higher-income Americans (71). Nevertheless, temporal variation in Twitter-derived sentiment measures appears to validate well against measures from independent survey data (25) and to correlate well with survey-derived measures of behavior (72). Finally, previous work on climate stress demonstrates that Twitter-derived sentiment measures responds similarly to such stress as compared with responses measured in representative vital statistics or survey data (73, 74).

Our mobility data are derived from smartphones. Recent data suggest that ~85% of adult Americans own smartphones (75). As with internet usage, smartphone ownership shows limited differences by race but some differences by income, with wealthier populations somewhat more likely to own a smartphone. Thus our mobility measures will be slightly skewed towards wealthier populations. As shown in Fig S1, wealthier populations on average report spending less time at home in representative survey data.

Our PurpleAir sample, among our many datasets, is likely the most non-representative of the overall American populace, with Census tracts that have PurpleAir monitors in them being wealthier, better educated, and slightly more White than tracts without monitors (45) (Fig S6). This finding is again consistent with our broader finding of income-differentiated responses to wildfire smoke. However, it also means that the variation in infiltration we measure might not accurately describe infiltration rates for lower-income and less educated populations, and better measurement of infiltration among these populations is an urgent research priority.

**Data availability** Data and code to replicate all results in the paper will be made available on a Github repo at the time of publication.

**References**

[1] Philip J Landrigan, Richard Fuller, Nereus JR Acosta, Olusoji Adeyi, Robert Arnold, Abdoulaye Bibi Baldé, Roberto Bertollini, Stephan Bose-O’Reilly, Jo Ivey Boufford, Patrick N


[52] Ryan Fox Squire. Measuring and correcting sampling bias in SafeGraph patterns for more accurate demographic analysis.

[53] Ryan Fox Squire. Quantifying sampling bias in SafeGraph patterns.

[54] Ryan Fox Squire. What about bias in the SafeGraph dataset?


Appendix

Figure S1: **Time spent indoors at home in America.** Data are from repeated rounds of the American Time Use Survey. Top panels show data by average income, age, season and race/ethnicity. Bottom map shows averages by state.
Figure S2: **Counties included in analyses that use EPA pollution monitors.** Counties in red are those with EPA pollution monitors from which we construct smoke PM$_{2.5}$ measures for the behavioral analyses.
Figure S3: **Exposure to average and acute smoke PM$_{2.5}$ at the county level does not differ systematically by income.** Daily smoke PM$_{2.5}$ exposures by income decile across US counties, 2006-2020. Dots represent daily observations where smoke PM$_{2.5}$ was non-zero. Plot is truncated at 300ug for clarity; not plotted are 71 days (0.001% of the sample) in which smoke PM$_{2.5}$ exceeded 300. Statistics at right show the percent of observations across the study period with daily smoke PM$_{2.5}$ observations above the listed value.
Figure S4: Effect of smoke PM$_{2.5}$ on percent fully home versus percent fully away. Left plot is as in Fig 2d, right plot is same but for % fully away from home on that day.
Figure S5: **Effect of smoke PM$_{2.5}$ on mobility as a function of income.** Lines show the marginal effect of a heavy smoke exposure (50ug smoke PM$_{2.5}$ on that day) on percent of individuals completely at home on that day (left plot) or completely away from home (right plot), as a function of median household income in that county. Colors represent models run with either date fixed effects (blue) or state-by-date fixed effects (orange). Shaded area is bootstrapped 95% CI.
Figure S6: **Higher income US census tracts are more likely to have PurpleAir monitors.**
Grey bars shows the distribution of tract-level median household income across all US census tracts in the contiguous US, red bars the income in tracts with at least one outdoor PurpleAir sensor, and blue bars the income in tracts with at least one indoor PurpleAir sensor. Vertical lines give the median of each distribution.
Infiltration estimates are highly correlated across alternate statistical models and methods of deriving PM$_{2.5}$ concentrations from Purple Air data. Correlation between infiltration estimates from statistical models with different lag structures (see Table S2 for model equations) and different PM$_{2.5}$ concentration estimates (see Methods for details). 1a is our preferred specification presented in the main results.
Figure S8: Understanding variation in household-specific infiltration estimates. a. Posterior estimates of monitor-specific infiltration rates from a Bayesian hierarchical model are very similar to “raw” estimates from our monitor-specific time-series regressions, indicating that true heterogeneity rather than sampling noise is what is driving observed differences in estimated infiltration. b. Ability of random forest (RF) or gradient boosted trees (GBT) model to explain variation ($R^2$) in infiltration across monitors remains low; models use predictors in (c). c. For each predictor, we calculate the effect on infiltration of moving from the 5th to the 95th percentile of that predictor in the test dataset, holding the other predictors constant at their average value in the test dataset; estimates are shown for RF and GBT models and for four alternate spatial buffers used to construct housing predictors. Housing Index is constructed by averaging standardized values of home value, number of stories, number of baths, number of bedrooms, height, and area. A/C measures the inverse distance weighted proportion of matched CoreLogic houses that have air conditioning. Median Income is the median household income in the Census tract population. Race variables (i.e. all demographic covariates except Hispanic) are measured among the non-Hispanic/Latino population. AI/AN stands for American Indian and Alaska Native. NHPI stands for Native Hawaiian and other Pacific Islander. HDD and CDD stand for heating degree days and cooling degree days, respectively. d Range of household-level infiltration estimates for the full sample and for sub-samples when behavior (i.e opening/closing of doors, use of air purifier) is expected to matter less: when it’s raining, nighttime, and periods when it is cold (<10°C) and low-PM$_{2.5}$ (<30µg/m$^3$). e-f during periods when behavioral factors are more likely minimized, infiltration varies more strongly with income and housing age.
Monitor-specific infiltration estimates using indoor/outdoor ratios versus regression-based approaches. I/O estimates are only modestly correlated with our preferred regression-based estimates that measure the marginal effect on indoor PM$_{2.5}$ concentrations of a unit increase in outdoor concentrations. For each monitor I/O ratio was calculated across all observations with hourly indoor PM$_{2.5} \leq$ outdoor PM$_{2.5}$. $\partial I/\partial O$ was estimated using Eq. 8 as described in Methods.
Figure S10: **Outdoor and indoor PM$_{2.5}$ concentrations on a smoke day in CA.** Very similar outdoor PM$_{2.5}$ concentrations during a smoke event on Aug 20th, 2021 over a high-income area of the peninsular Bay Area were associated with widely varying contemporaneous indoor PM$_{2.5}$ concentrations.
Figure S11: **Variation in indoor PM$_{2.5}$ across monitors with similar outdoor PM$_{2.5}$ during the Aug/Sep 2020 smoke event in the Bay Area.** Each dot is average outdoor PM$_{2.5}$ and average indoor PM$_{2.5}$ for an individual monitor in the Bay Area over the Aug/Sep 2020 smoke event, with monitors grouped into 5µg bins based on outdoor exposure. Numbers at top show the number of monitors in each bin (black), and the ratio of maximum to minimum indoor PM$_{2.5}$ across monitors within each outdoor PM$_{2.5}$ bin (red). Monitors with average outdoor PM$_{2.5}$ exposures within 5µg of each other experienced >20x differences in indoor PM$_{2.5}$ exposures.
Table S1: **Hourly EPA reference monitor PM$_{2.5}$ concentrations versus different constructions of Purple Air PM$_{2.5}$ concentrations.** Each column is a separate linear regression with hourly PM$_{2.5}$ concentration at an EPA reference monitor as the dependent variable and a different version of Purple Air derived PM$_{2.5}$ concentration as the independent variable. All models include EPA monitor fixed effects, hour-of-day fixed effects, and month-of-sample fixed effects.

<table>
<thead>
<tr>
<th>Purple Air PM$_{2.5}$ Construction:</th>
<th>CF_1</th>
<th>Barkjohn et al correction</th>
<th>particle count derived</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purple Air PM$_{2.5}$</td>
<td>0.600***</td>
<td>1.064***</td>
<td>2.196***</td>
</tr>
<tr>
<td></td>
<td>(0.035)</td>
<td>(0.069)</td>
<td>(0.109)</td>
</tr>
<tr>
<td>R$^2$</td>
<td>0.67</td>
<td>0.68</td>
<td>0.72</td>
</tr>
<tr>
<td>Num.Obs.</td>
<td>199,394</td>
<td>199,394</td>
<td>199,394</td>
</tr>
<tr>
<td>FE: EPA monitor</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FE: hour-of-day</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FE: month-of-sample</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

* p < 0.05, ** p < 0.01, *** p < 0.001
### Table S2: Models of infiltration

<table>
<thead>
<tr>
<th>Model description</th>
<th>Regression</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Distributed Lag</td>
<td>Regression: ( PM_{i,hdm}^{in} = \sum_{k=0}^{6} \alpha_k PM_{i,h-k,dm}^{out} + \gamma_i + \delta_h + \eta_d + \theta_m + \varepsilon_{i,hdm} )</td>
<td>( \sum_{k=0}^{6} \alpha_k )</td>
</tr>
<tr>
<td>2 Lagged Dependent Variable</td>
<td>Regression: ( PM_{i,hdm}^{in} = \alpha PM_{i,hdm}^{out} + \beta PM_{i,h-1,dm}^{in} + \gamma_i + \delta_h + \eta_d + \theta_m + \varepsilon_{i,hdm} )</td>
<td>( \alpha \cdot \left( \sum_{k=0}^{\infty} \beta^k \right) )</td>
</tr>
<tr>
<td>3 No Lags</td>
<td>Regression: ( PM_{i,hdm}^{in} = \alpha PM_{i,hdm}^{out} + \gamma_i + \delta_h + \eta_d + \theta_m + \varepsilon_{i,hdm} )</td>
<td>( \alpha )</td>
</tr>
<tr>
<td>4 All Lags</td>
<td>Regression: ( PM_{i,hdm}^{in} = \sum_{k=0}^{6} \alpha_k PM_{i,h-k,dm}^{out} + \beta PM_{i,h-1,dm}^{in} + \gamma_i + \delta_h + \eta_d + \theta_m + \varepsilon_{i,hdm} )</td>
<td>( \left( \sum_{k=0}^{6} \alpha_k \right) \cdot \left( \sum_{k=0}^{\infty} \beta^k \right) )</td>
</tr>
</tbody>
</table>
Table S3: **Salience results are robust to alternate search words in English and Spanish.** Each column is a separate linear regression using the phrase in the column header as the search term, and smoke PM$_{2.5}$ as the regressor. All models include the same fixed effects but different numbers of observations, as Google Trends does not return information for areas where search volume was low. Observations are marketing area by week. Standard errors are shown in parentheses, clustered at marketing area month.

<table>
<thead>
<tr>
<th>Outcome</th>
<th>air quality</th>
<th>smoke</th>
<th>calidad del aire</th>
<th>humo</th>
</tr>
</thead>
<tbody>
<tr>
<td>smoke PM$_{2.5}$</td>
<td>0.689***</td>
<td>0.567***</td>
<td>0.622***</td>
<td>0.118*</td>
</tr>
<tr>
<td></td>
<td>(0.095)</td>
<td>(0.052)</td>
<td>(0.114)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Num.Obs.</td>
<td>50600</td>
<td>51900</td>
<td>5220</td>
<td>22707</td>
</tr>
<tr>
<td>FE: date</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FE: DMA x month</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

$+ p < 0.1$, $* p < 0.05$, $** p < 0.01$, $*** p < 0.001$
Table S4: **Salience and health protection measures are robust to alternate samples, FE, and controls.** Each column is a separate linear regression of either salience or health protection on smoke PM$_{2.5}$, with dependent variable noted in the top line. Column 1 is baseline specification, column 2 includes controls for daily temperature and precipitation, columns 3-4 restrict panel to county-days < or > 50km from an active wildfire fire, column 5-6 uses more restrictive time FE; m.o.s = month of sample, m.o.y = month of year. Column 7 uses average smoke PM$_{2.5}$ over the previous week as the regressor. Last four columns repeat similar regression for % fully away. Smoke PM$_{2.5}$ is measures in ug/m3, with coefficients multiplied by 1000 to make them legible. Dependent variable is average sentiment on a county-day. Included fixed effects in each regression are shown at bottom. Standard errors are shown in parenthesis, clustered at the county-month.

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>Searches for 'air quality'</th>
<th>Searches for 'air filter'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model:</td>
<td>baseline</td>
<td>+controls</td>
</tr>
<tr>
<td>smoke PM$_{2.5}$</td>
<td>0.689*** (0.095)</td>
<td>0.681*** (0.118)</td>
</tr>
<tr>
<td>temperature</td>
<td>0.277*** (0.077)</td>
<td></td>
</tr>
<tr>
<td>precipitation</td>
<td>-6.808* (2.798)</td>
<td></td>
</tr>
<tr>
<td>Num.Obs.</td>
<td>50600</td>
<td>43814</td>
</tr>
<tr>
<td>FE: date</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FE: DMA x m.o.y</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FE: DMA x m.o.s</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001
Table S5: **Sentiment results are robust to alternate samples, FE, and controls.** Each column is a separate linear regression of sentiment on smoke PM$_{2.5}$. Column 1 is baseline specification, column 2 includes controls for daily temperature and precipitation, columns 3-4 restrict panel to county-days < or > 50km from an active wildfire fire, columns 5-6 uses more restrictive time FE; m.o.s = month of sample, m.o.y = month of year. Smoke PM$_{2.5}$ is measures in ug/m3, with coefficients multiplied by 1000 to make them legible. Dependent variable is average sentiment on a county-day. Included fixed effects in each regression are shown at bottom. Standard errors are shown in parenthesis, clustered at the county-month.

<table>
<thead>
<tr>
<th>Model:</th>
<th>baseline</th>
<th>+controls</th>
<th>distance &gt; 50</th>
<th>distance &lt; 50</th>
<th>alt FE</th>
<th>alt FE</th>
</tr>
</thead>
<tbody>
<tr>
<td>smoke PM$_{2.5}$</td>
<td>-0.087*** (0.011)</td>
<td>-0.088*** (0.011)</td>
<td>-0.085*** (0.013)</td>
<td>-0.081*** (0.017)</td>
<td>-0.075*** (0.016)</td>
<td>-0.082*** (0.009)</td>
</tr>
<tr>
<td>temperature</td>
<td>0.000</td>
<td>(0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>precipitation</td>
<td>-0.011*** (0.000)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num.Obs.</td>
<td>1239533</td>
<td>1175623</td>
<td>1009505</td>
<td>211485</td>
<td>1239533</td>
<td>1239533</td>
</tr>
<tr>
<td>FE: date</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FE: date x state</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>FE: county x m.o.y.</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FE: county x m.o.s.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001
Table S6: **Health protection results are largely robust to alternate search words in English and Spanish.** Each column is a separate regression using the word in the column header as the search term. All models include the same fixed effects but different numbers of observations, as Google Trends does not return information for areas where search volume was low. Observations are marketing area by week. Standard errors are shown in parentheses, clustered at marketing area month.

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>air filter</th>
<th>air purifier</th>
<th>purple air</th>
<th>smoke mask</th>
<th>purificador de aire</th>
<th>filtro de aire</th>
</tr>
</thead>
<tbody>
<tr>
<td>smokePM</td>
<td>0.453***</td>
<td>0.540***</td>
<td>0.535***</td>
<td>0.350***</td>
<td>0.485**</td>
<td>0.020</td>
</tr>
<tr>
<td></td>
<td>(0.078)</td>
<td>(0.069)</td>
<td>(0.103)</td>
<td>(0.089)</td>
<td>(0.163)</td>
<td>(0.102)</td>
</tr>
<tr>
<td>Num.Obs.</td>
<td>51644</td>
<td>50343</td>
<td>29493</td>
<td>8091</td>
<td>2088</td>
<td>4959</td>
</tr>
<tr>
<td>FE: date</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FE: DMA x month</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001
Table S7: **Mobility results are largely robust to alternate samples, FE, and controls.** Each column is a separate linear regression of sentiment on smoke PM$_{2.5}$, with either % of devices fully at home or % of devices fully away as the dependent variable, as noted in the top line. Column 1 is baseline specification, column 2 includes controls for daily temperature and precipitation, columns 3-4 restrict panel to county-days < or > 50km from an active wildfire fire, columns 5-6 uses more restrictive time FE; m.o.s = month of sample, m.o.y = month of year. Column 7 uses average smoke PM$_{2.5}$ over the previous week as the regressor. Last four columns repeat similar regression for % fully away. Smoke PM$_{2.5}$ is measures in ug/m$^3$, with coefficients multiplied by 1000 to make them legible. Dependent variable is average sentiment on a county-day. Included fixed effects in each regression are shown at bottom. Standard errors are shown in parenthesis, clustered at the county-month.

<table>
<thead>
<tr>
<th>Outcome:</th>
<th>Model:</th>
<th>baseline</th>
<th>+controls</th>
<th>distance &gt; 50</th>
<th>distance &lt; 50</th>
<th>alt FE</th>
<th>alt FE</th>
<th>smokeweek</th>
<th>% fully away</th>
<th>baseline</th>
<th>smokeweek</th>
<th>alt FE</th>
<th>alt FE</th>
</tr>
</thead>
<tbody>
<tr>
<td>smoke PM$_{2.5}$</td>
<td></td>
<td>0.023*** (0.004)</td>
<td>0.023*** (0.003)</td>
<td>0.026*** (0.004)</td>
<td>0.018*** (0.003)</td>
<td>0.001</td>
<td>0.012*** (0.001)</td>
<td></td>
<td>0.005** (0.001)</td>
<td>0.005* (0.002)</td>
<td>0.002* (0.001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>temperature</td>
<td></td>
<td>-0.069*** (0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>precipitation</td>
<td></td>
<td>6.012*** (0.112)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>smoke PM$_{2.5}$ (week)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.031*** (0.006)</td>
<td></td>
<td>0.005* (0.002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Num.Obs.</td>
<td></td>
<td>678970</td>
<td>645141</td>
<td>538707</td>
<td>132504</td>
<td>678970</td>
<td>678970</td>
<td>608771</td>
<td></td>
<td>678970</td>
<td>608771</td>
<td>678970</td>
<td>678970</td>
</tr>
<tr>
<td>FE: date</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>FE: date x state</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FE: county x m.o.y.</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>FE: county x m.o.s.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001
Table S8: **Heterogeneity in salience and health protection measures.** Each column is a separate regression using the word in the column header as the moderator (interaction term in the regression), with searches for ‘air quality’ in the first four columns and searches for ‘air filter’ in the last four. Moderators are centered at their mean value. Columns marked “all” include all three moderators interacted with smoke PM$_{2.5}$
. Observations are marketing area by week. Standard errors are shown in parentheses, clustered at marketing area month.

<table>
<thead>
<tr>
<th>Interaction variable</th>
<th>Outcome:</th>
<th>searches for ‘air quality’</th>
<th>searches for ‘air filter’</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>income</td>
<td>avg smoke PM$_{2.5}$</td>
<td>avg PM$_{2.5}$</td>
</tr>
<tr>
<td>smoke PM$_{2.5}$</td>
<td>0.701*** (0.124)</td>
<td>0.873*** (0.116)</td>
<td>0.659*** (0.073)</td>
</tr>
<tr>
<td>smoke PM$_{2.5} \times$ income</td>
<td>-0.001 (0.006)</td>
<td>-0.009 (0.006)</td>
<td>-0.018*** (0.003)</td>
</tr>
<tr>
<td>smoke PM$<em>{2.5} \times$ avg smoke PM$</em>{2.5}$</td>
<td>-0.722*** (0.202)</td>
<td>-0.752*** (0.228)</td>
<td>-0.382+ (0.211)</td>
</tr>
<tr>
<td>smoke PM$<em>{2.5} \times$ avg PM$</em>{2.5}$</td>
<td>0.041+ (0.022)</td>
<td>0.024 (0.023)</td>
<td>-0.024+ (0.013)</td>
</tr>
</tbody>
</table>

Num.Obs. | 50600 | 50600 | 49573 | 49573 | 50343 | 50343 | 49316 | 49316 |
FE: date | X | X | X | X | X | X | X | X |
FE: DMA x month | X | X | X | X | X | X | X | X |

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001
Table S9: **Heterogeneity in sentiment and mobility measures.** Each column is a separate regression using the word in the column header as the moderator (interaction term in the regression). Sentiment is the outcome in the first four columns and mobility (% of devices fully at home) the outcome in the last four. Moderators are centered at their mean value. Columns marked “all” include all three moderators interacted with smoke PM$_{2.5}$. Observations are county-day. Standard errors are shown in parentheses, clustered at county-month.

<table>
<thead>
<tr>
<th>Interaction variable:</th>
<th>Outcome:</th>
<th>sentiment</th>
<th>% completely at home</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>income</td>
<td>avg smoke PM$_{2.5}$</td>
<td>avg PM$_{2.5}$</td>
</tr>
<tr>
<td>smoke PM$_{2.5}$</td>
<td>-0.070***</td>
<td>-0.097***</td>
<td>-0.083***</td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.012)</td>
<td>(0.012)</td>
</tr>
<tr>
<td>smoke PM$_{2.5}$ × income</td>
<td>-0.001**</td>
<td>-0.002**</td>
<td>(0.000)</td>
</tr>
<tr>
<td>smoke PM$<em>{2.5}$ × avg smoke PM$</em>{2.5}$</td>
<td>0.050</td>
<td>-0.026</td>
<td>(0.042)</td>
</tr>
<tr>
<td>smoke PM$<em>{2.5}$ × avg PM$</em>{2.5}$</td>
<td>0.016***</td>
<td>0.016***</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Num.Obs.</td>
<td>1239533</td>
<td>1239533</td>
<td>1117019</td>
</tr>
<tr>
<td>FE: date</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>FE: county x month</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

+ p < 0.1, * p < 0.05, ** p < 0.01, *** p < 0.001