Environmental Science & Technology

pubs.acs.org/est

Article

Using Crowd-Sourced Data to Assess the Temporal and Spatial Relationship between Indoor and Outdoor Particulate Matter

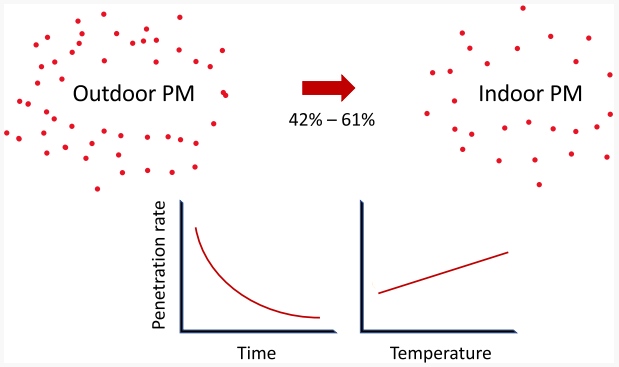
Benjamin Krebs,* Jennifer Burney, Joshua Graff Zivin, and Matthew Neidell

Cite This: *Environ. Sci. Technol.* 2021, 55, 6107−6115

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Using hourly measures across a full year of crowd-sourced data from over 1000 indoor and outdoor pollution monitors in the state of California, we explore the temporal and spatial relationship between outdoor and indoor particulate matter (PM) concentrations for different particle sizes. The scale of this study offers new insight into both average penetration rates and drivers of heterogeneity in the outdoor−indoor relationship. We find that an increase in the daily outdoor PM concentration of 10% leads to an average increase of 4.2−6.1% in indoor concentrations. The penetration of outdoor particles to the indoor environment occurs rapidly and almost entirely within 5 h. We also provide evidence showing that penetration rates are associated with building age and climatic conditions in the vicinity of the monitor. Since people spend a substantial amount of each day indoors, our findings fill a critical knowledge gap and have significant implications for government policies to improve public health through reductions in exposure to ambient air pollution.



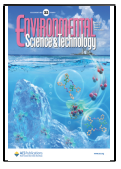
INTRODUCTION

Particulate matter (PM) is the largest environmental cause of mortality around the globe, with fine particulate matter—particles with a diameter less than 2.5 μm—estimated to cause millions of premature deaths annually.[1,2] Although PM is produced by a vast array of sources, most of the dominant contributors, especially in regions that do not depend on biomass for cooking and heating, tend to be outdoors (e.g., transportation, electric power generation, industry, forest fires, dust). While the built environment provides shelter from the elements, it is an incomplete filter; PM can penetrate buildings through open doors and windows, and localized studies have shown that smaller particles can enter through the small cracks and porous materials that comprise the outer shells of nearly all structures.[3,4] Given that Americans spend approximately 85−90% of their time indoors, with the vast majority in a private residence,[5,6] any effort to understand and limit damages from PM will fundamentally hinge on the degree to which outdoor air pollution penetrates indoors.

Despite the preponderance of evidence on the outdoor−indoor relationship, critical gaps in our understanding remain. Most evidence has relied on laboratory simulations or studies that measured indoor and outdoor conditions either in a small sample of locations or over a short period of time. (See ref 7 for an overview of this study and refs 8−19 for some of the most prominent contributions.) Estimates from these studies are wide-ranging, and it is unclear whether the differences across studies reflect true differences in PM penetration, differences in methodology, or a combination of these factors. Moreover, the vast majority of prior studies have focused on average indoor/outdoor ratios (or correlations) that provide a steady-state measure of the relationship, with limited evidence on the dynamic process that connects outdoor conditions to the indoor environment.

Here, we provide a large-scale analysis across a broad spatial range of the relationship between outdoor and indoor aerosol particulate matter concentrations based on unique, crowd-sourced data from the PurpleAir Real Time Air Quality Monitoring Network (PA hereafter). Our dataset includes over 14.5 million observations and builds upon pioneering large-scale analyses of data from Beijing, China.[20] It allows us to more precisely estimate this relationship and explore temporal and spatial relationships not feasible in earlier work. Using a full year of hourly data from these monitors, placed by consumers both indoors and outdoors at thousands of locations across the state of California, we directly explore the dynamic relationship between outdoor and indoor PM concentrations of various particle sizes. We further examine heterogeneity in this relationship by time of day, season, and across monitor locations (see the Data section for details). Our analysis thus fills a critical

Received: December 15, 2020
Revised: April 8, 2021
Accepted: April 8, 2021
Published: April 21, 2021



https://doi.org/10.1021/acs.est.0c08469
Environ. Sci. Technol. 2021, 55, 6107−6115

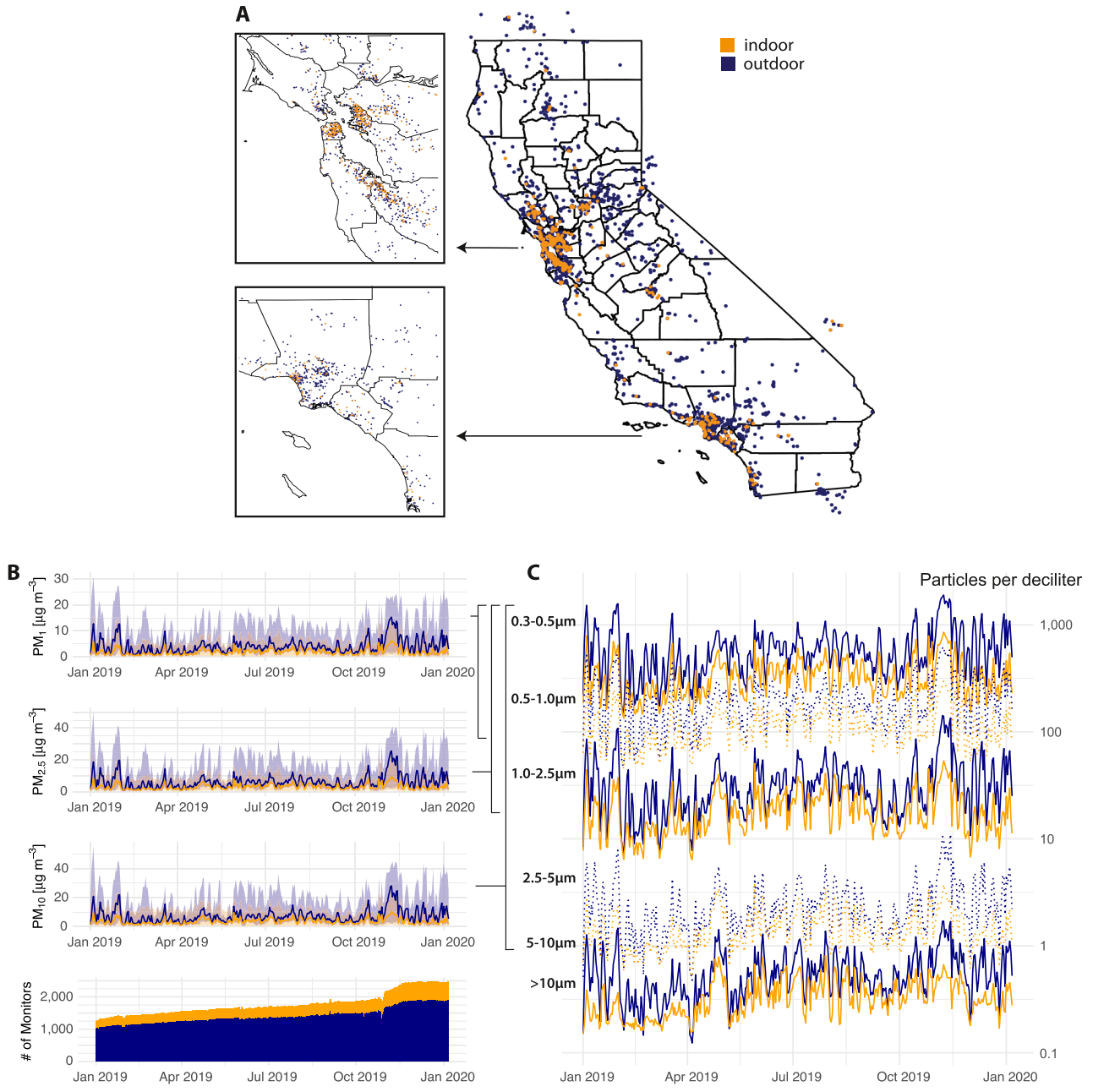


Figure 1. Summary of study parameters and data. (A) Locations of indoor (orange) and outdoor (blue) PA monitors operating for some or all of the time between January 1, 2019, and December 31, 2019. The insets are shown for the Bay Area (top) and Los Angeles area, where a majority of monitors have been installed. All monitors within the state of California were included. (B) From top to bottom, median $PM_{1.0}$, $PM_{2.5}$, and PM_{10} concentrations over the study period, and total number of monitors online by date. Monitor installation and concentrations increased appreciably during the November wildfire period. (C) Concentration data are derived from underlying particle count data shown here. Since the conversion to concentrations rests on assumptions about particle size distributions (see Figure S1), we use the underlying particle count data by size bin in this study.

knowledge gap about the magnitudes and variability of indoor particulate matter exposures, stemming from outdoor sources, with direct implications for government policies to improve public health through reductions in exposure to ambient air pollution.

DATA AND METHODS

Data. The data for our analysis comes from the PurpleAir Real Time Air Quality Monitoring Network (https://www2.purpleair.com/). PA produces a retail, residential air pollution monitor that measures airborne particulate matter (PM) at 120 s intervals. PA sensors use laser particle counters to measure particle number concentrations (number of particles per deciliter, dl) under different size thresholds (0.3, 0.5, 1, 2.5, 5, and 10 μm diameter cutoffs). PA reports direct number concentrations as well as conversions to the more standard $PM_{1.0}$, $PM_{2.5}$, and PM_{10} mass-based metrics ($\mu g\ m^{-3}$). Since these conversions rely on assumptions about the particulate mix that may not be uniformly valid across our study region, we use the raw particle counts to assess variation in outdoor−indoor relationships for different size particles.

Although PA monitors are not officially approved by the Environmental Protection Agency (EPA) for regulatory monitoring purposes, they have been tested by the Air Quality

https://doi.org/10.1021/acs.est.0c08469
Environ. Sci. Technol. 2021, 55, 6107−6115

Sensor Performance Evaluation Center maintained by the South Coast Air Quality Management District, the regulatory agency responsible for improving air quality for more than 17 million residents in Southern California. The Agency concluded that the PA-II sensor data correlated highly with standard reference measures, with an R^2 of 0.96 for particles smaller than 1 μm and 0.93 for particles smaller than 2.5 μm.[21] Studies have shown that the reliability of these monitors declines for particles over 2.5 μm.[22] We include such data in our analysis while noting the limitations that this measurement error presents. The PA monitors, which retail for $179−259, have proliferated throughout the U.S. and abroad in a short period of time, with approximately 7000 monitors sold from mid-2016 through the end of 2019.

Once installed at a particular location, monitor readings are automatically sent to the PA servers and are available for download from https://thingspeak.com/ via JSON. We downloaded hourly data from all monitors in CA in 2019, shown in Figure 1A. After the quality check described below, these raw data contain approximately 14.5 million hourly readings from 596 indoor and 2006 outdoor monitors. We then match each indoor monitor with all outdoor monitors within a specific radius (see below for details). We exclude indoor monitors beyond 500 m of an outdoor monitor. This provides our analysis sample of approximately 1.4 million merged hourly readings from 349 indoor monitors and 708 outdoor monitors residing in 29 of the 57 counties in CA. This compares with 52 $PM_{2.5}$ monitors and 41 PM_{10} monitors for the entire state as maintained by the California Air Resources Board (https://ww3.arb.ca.gov/qaweb/site.php). In addition to number concentrations, reported data include an indicator for whether the purchaser placed the monitor indoors or outdoors, along with the monitor GPS coordinates.

We took several steps to prepare the raw data for analysis. First, we cleaned the data as suggested by PA.[23] Nearly all outdoor monitors as well as approximately one-third of our indoor monitor sample contain two laser particle counters. For these monitors, we dropped all observations for which the two measurements for $PM_{2.5}$ differed by more than 10 units if both values fell below 100 $\mu g\ m^{-3}$, or by more than 10% if one of the values exceeded 100 $\mu g\ m^{-3}$. We then calculated the mean of the two values and removed all observations with values higher than 500 $\mu g\ m^{-3}$, the specified accuracy threshold for PA. For monitors with a single particle counter, we only performed the latter of these two cleaning steps. These steps remove 955 075 observations in total. The concentration data are shown in Figure 1B.

Second, we transformed the particle number concentrations under different size cutoffs to number concentrations for five nonoverlapping size bins: 0.3−0.5, 0.5−1, 1−2.5, 2.5−5, and 5−10 μm. (To create, e.g., the 0.3−0.5 μm size bin, we subtracted the particle concentration in the 0.5 μm cutoff from the 0.3 μm cutoff.) We then dropped any resulting negative or zero observations for each of those bins and trimmed the top percentile of observations. We also dropped monitors with less than 720 observations (30 days). We lose 379 599 observations by dropping negatives and zeros, 307 013 by trimming, and 43 424 by removing monitors with less than 720 observations. The measures for these individual size bins (and their mapping to calculated mass-based concentrations) are shown in Figure 1C.

Finally, from the cleaned individual monitor data, we used the geolocation for each monitor to calculate physical distances between indoor and outdoor monitors. From this, we matched each indoor monitor to all outdoor monitors within 500 m (we assess robustness to alternative distance assumptions). When multiple outdoor monitors matched to a particular indoor monitor, we calculated the mean of all PM measures from the matched outdoor monitors, weighted by the inverse squared distance to the indoor monitor.

To control for local meteorological conditions that might influence both indoor and outdoor concentrations (through either direct physical effects or by induced behavior changes), we obtained hourly weather data from the National Centers for Environmental Information's Integrated Surface Database (https://www.ncdc.noaa.gov/isd, accessed August 24, 2020). We calculate county-level averages from the network's hourly readings of air temperature, dew point temperature, precipitation, and wind speed. We restricted our calculations to include only stations that were in operation for all of 2019 and had no more than 25% of observations missing, and we interpolated any remaining missing values with the average of all surrounding stations within 50 km, weighted by the inverse squared distance. Finally, we calculated county-level metrics by taking the average of all stations within the respective county, or, in the one case no monitor is present, by taking the values of the closest monitor to the county's centroid.

Methods. For comparison with previous studies, we first estimate the raw correlation between daily outdoor and indoor measurements (and calculated their ratios). We measure this static relationship both overall, across days, and across seasons, separately for each size bin (Figure 2).

To understand the dynamic relationship between outdoor and indoor particulate matter levels, we estimate hourly distributed-lag regression models with indoor pollution as the dependent variable and outdoor pollution as the key independent variable

$$I_{it} = \sum_j \beta_j \bar{O}_{it-j} + \gamma X_{ct} + \theta_i + f(t) + \varepsilon_{it} \quad (1)$$

Given the right-skewed nature of the pollution data (Figure 2A, note the log scale), we specify our measures of indoor and outdoor pollution in logs. We relate measured (log) indoor particle number concentrations (I) at monitor i at hour t to nearby (log) outdoor particle number concentrations ($\bar{O}$, the inverse-distance weighted average of measurements at time t within a radius of 500 m of indoor monitor i) as shown in eq 1. To explore temporal dynamics, we include 12 hourly lags of outdoor pollution. As we demonstrate below, 12 lags are sufficient for understanding the entire dynamic process; adding additional lags has no material impact on estimates. We account for local meteorology by including a vector of weather variables (X_{ct}) (i.e., air temperature, air temperature squared, dew point temperature, wind speed, and precipitation) in county c at the time t. The coefficient vector γ thus captures the effects of weather on indoor pollution. To account for other potential confounds, we allow each indoor monitor to have its own intercept (or fixed effect) (θ_i), thereby adjusting for overall average differences in indoor and outdoor PM concentrations specific to each location. To account for potential temporal confounds, we include $f(t)$, a vector of binary time variables, representing month, day-of-week, and hour-of-day fixed effects. These fixed effects control for all common characteristics specific to each month, day of the week or hour of the day across all locations. We assume that any remaining errors (ε) are serially correlated over time within each monitor.

https://doi.org/10.1021/acs.est.0c08469
Environ. Sci. Technol. 2021, 55, 6107−6115

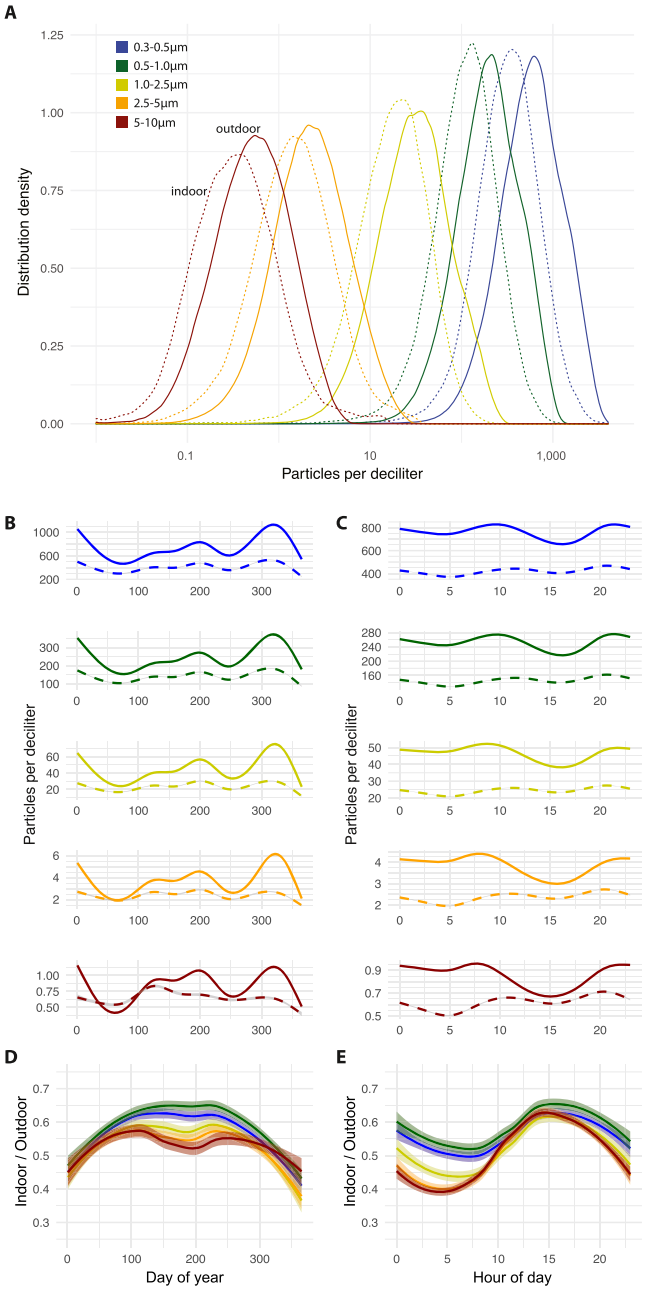


Figure 2. Distributions of outdoor and indoor particle number concentrations as measured by PA monitors and unconditional average relationships between them. (A) Distributions of indoor (dashed) and outdoor (solid) particle number concentrations, by size bin. (B) Average daily particle number concentrations across all sites outdoors (solid lines) and indoors (dashed lines), shown for each size bin over the course of the year. (C) Average hourly particle number concentrations across all sites outdoors (solid lines) and indoors (dashed lines), shown for each size bin by hour of day (PST). (D) Median ratios of indoor/outdoor particle concentrations by day of year. (E) Median ratios of indoor/outdoor particle concentrations by hour of day (PST). On average, particle concentrations are higher outdoors than indoors across all size bins, although this ratio varies and is sometimes inverted, particularly for the largest particles. The ratio of indoor to outdoor concentrations is closer to 1 between March and November, and smaller during winter months; the ratio is also higher during afternoon and evening hours. Uncertainty ranges in (D) and (E) show the standard errors around loss fits to median indoor/outdoor ratios.

We fit this model using ordinary least squares and focus on the best-fit estimated coefficient $\hat{\beta}_j$, or $\langle \partial I / \partial \bar{O}_j \rangle$, as the main policy parameter of interest. Given the log−log regression specification, $\hat{\beta}_j$ is thus interpretable as elasticity, or the % change in indoor concentrations from a 1% change in average nearby outdoor concentrations. We calculate standard errors for $\hat{\beta}_j$ clustered on the indoor monitor to allow for arbitrary serial correlation within each monitor. We perform separate analyses for each particle size bin to understand whether the outdoor−indoor relationship varies by size. We examine heterogeneity in this relationship by interacting outdoor concentrations (with $\bar{O}_{it}$) with county-level temperature metrics (Figure 3C and Table S5) to test the hypothesis that indoor PM concentrations might more closely mirror outdoor PM concentrations at times when ambient temperatures are more conducive to the opening of doors and windows.

To explore heterogeneity at the individual monitor level, we interact $\bar{O}_j$ with θ_i to estimate penetration rates specific to each indoor pollution monitor (β_i). Since we find that nearly all penetration occurs within a few hours (as shown below), we modify eq 1 to the daily level. After obtaining these estimates, we probe sources of heterogeneity using the following procedure. We recast the estimated β_i's as a random variable (z_i) and regress this against fixed factors at each location according to eq 2

$$z_i = \delta_1 V_z + \delta_2 W_c + u_i \tag{2}$$

where V_z is a vector that includes the percent of buildings constructed after 1960 and the percent of owner-occupied buildings at the zip code (obtained from the 2018 American Community Survey) and W_c is a vector of seasonal climate measured at the county level, including means of winter temperature, summer temperature, annual dew point, and annual precipitation. Given the cross-sectional nature of this regression and our ecological rather than individual measures of the characteristics of the environment where the monitor resides, we interpret the coefficients δ_1 and δ_2 with caution, noting that they represent suggestive evidence on monitor-level sources of heterogeneity.

RESULTS

The data for our analysis comes from the PurpleAir Real Time Air Quality Monitoring Network (https://www2.purpleair.com/), with monitor locations shown in Figure 1A. We calculate hourly, raw particle counts under different size bins: 0.3−0.5, 0.5−1, 1−2.5, 2.5−5, and 5−10 μm, with the concentration data shown in Figure 1B. For comparison with previous studies, we first estimate the raw correlation between daily outdoor and indoor measurements (and calculated their ratios) overall, across days, and across seasons, separately for each size bin (Figure 2). To understand the dynamic relationship, we estimate hourly distributed-lag regression models with indoor pollution as the dependent variable and outdoor pollution as the key independent variable, controlling for numerous sources of potential confounding. Given the right-skewed nature of the pollution data (Figure 2A, note the log scale), we specify our measures of indoor and outdoor pollution in logs, and interpret estimates as the % change in indoor concentrations from a 1% change in average nearby outdoor concentrations (see the Methods section for more details).

Consistent with the existing literature, we find that the indoor environment is only marginally protective of PM: on average, outdoor concentrations are higher than indoor concentrations for all size bins (Figure 2A), and this is true both over the course

https://doi.org/10.1021/acs.est.0c08469
Environ. Sci. Technol. 2021, 55, 6107−6115

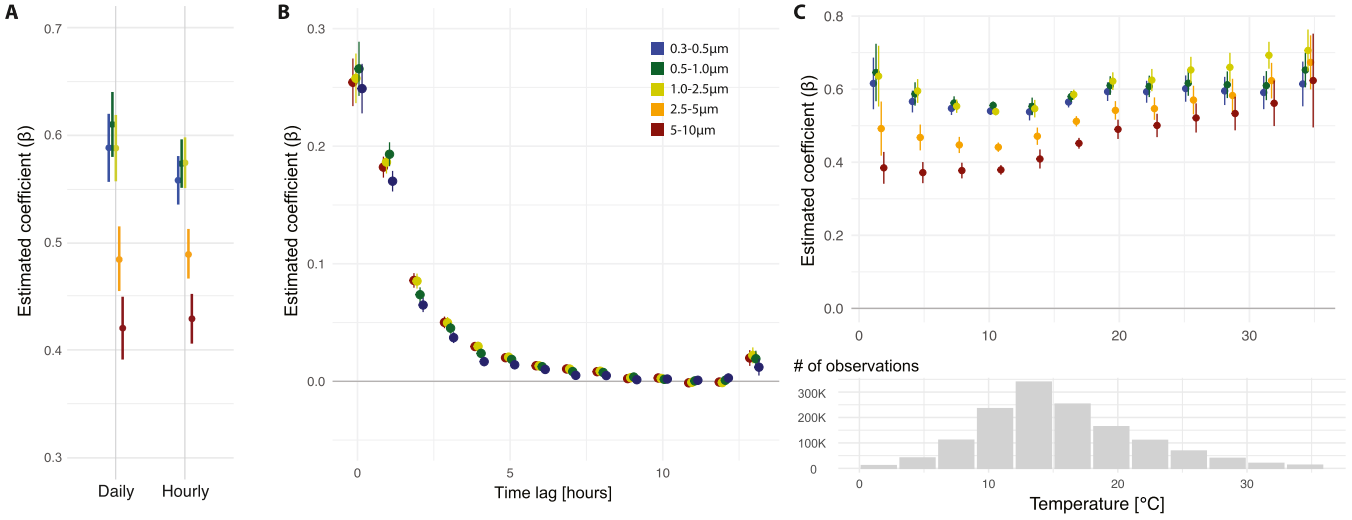


Figure 3. Relationship between changes in outdoor concentrations and changes in indoor concentrations, as measured by the inverse-distance-weighted average concentration measured by outdoor monitors within 500 m of an indoor monitor. (A) Basic contemporaneous daily and hourly relationships. (B) Breakdown of lags comprising the total daily relationship. Almost all of the influence of the outdoor environment on the indoor environment takes place within 5 h. (C) Breakdown of penetration rates by outdoor temperature. All statistical relationships include adjustment for monitor, day of year, and hour of day (for hourly); in all cases, error bars show 2× standard error of the coefficient (and are too small to be seen for many estimates).

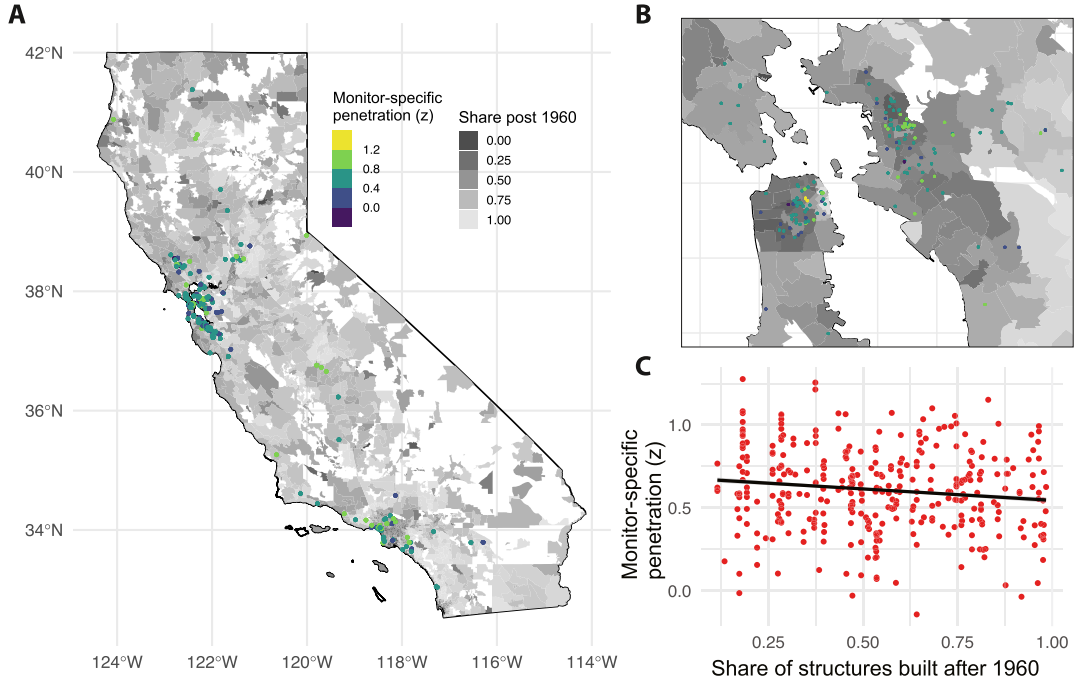


Figure 4. Variation in penetration rates across indoor monitoring locations. (A) Median building construction year across California zip code tabulation areas (ZCTA), and monitor-specific relationship between changes in average outdoor particle number concentrations and changes in indoor particle number concentrations (shown for the 0.3−0.5 μm size bin). (B) Zoomed-in view of the San Francisco Bay Area, site of the largest density of PurpleAir monitors. (C) At the ZCTA level, the age of building stock plays a significant role in explaining this heterogeneity.

of the year (Figure 2B,C) and over the course of a day (Figure 2D,E). Average indoor particulate levels are 40−70% of nearby (500 m) outdoor levels, but that ratio varies in time and space, and by particle size: smaller (more dangerous) particles have higher average indoor/outdoor ratios than larger particles, though these differences may be, at least in part, attributable to measurement error.[22] These raw ratios are lowest in the winter for all size bins, and highest in the summer for smaller particles (<1.0 μm), and in late spring and early fall for the larger size bins (>2.5 μm). The ratio of indoor/outdoor is smaller at night than during the day for all size bins, but with spread between them: the largest particles reach a minimum indoor/outdoor ratio of <0.4 in the predawn hours, while the smaller bins never drop below 0.5. The average indoor/outdoor ratio peaked in early afternoon, with similar ratios across size bins. These findings are consistent with increased air exchange during waking hours (when doors and windows are more like to be opened).[24,25]

The PA data also offer a unique opportunity to examine particulate size distributions below the 2.5 μm threshold. This view is important as particle size influences subsequent health

https://doi.org/10.1021/acs.est.0c08469
Environ. Sci. Technol. 2021, 55, 6107−6115

effects[26−28] and comparable data are not available from the EPA nor could it be detected using mass-based measurement (Figure S1). In our sample, indoor and outdoor PM distributions have a similar fraction of total particles in the 1.0−2.5 μm size range, but a larger fraction in the 0.5−1.0 μm range are present in indoor air and a slightly smaller fraction of the smallest particles (0.3−0.5 μm) are present in indoor air (Figure S2).

Estimates from our dynamic regression models that relate outdoor and indoor particle concentrations, while adjusting for weather, location, and temporal effects (Methods, eq 1), show that a 10% increase in daily outdoor concentrations leads to a 4.2−6.1% increase in daily indoor concentrations, with the estimates higher for smaller particles (Figure 3A and Table S1). Turning to an hourly analysis shows comparable size estimates (also Figure 3A and Table S2). When we extend the hourly regression to include 12 lags of outdoor concentrations, we find that penetration is nearly complete within 5 h for all size bins (Figure 3B and Table S3) (based on *t*-tests for whether the sum of the hourly lags coefficients equals the daily coefficient, separately for each particle size bin (Table S4)). The accumulated effect over the course of the day, obtained by summing the individual estimates of β_j, ranged from 0.559 for the 5−10 μm size bin to 0.684 for the 0.5−1.0 μm size bin. The penetration occurs quite rapidly, with 37-42% of the effect occurring in the first hour. As with the static results, the coefficients for smaller particles were larger in magnitude than for bigger particles, with the 0.3−2.5 μm bins having the strongest relationship. Our finding that transfer rates significantly increased during moderate temperature days, particularly for larger sizes particles, suggests a potentially important role for open windows and doors (Figure 3C and Table S5). To address selection concerns regarding the spike in monitor ownership in response to the California wildfires, we exclude monitors purchased after November 1, 2019, and find very similar estimates (Table S8).

To explore heterogeneity in this relationship across monitors, we estimate penetration rates specific to each indoor pollution monitor. Looking across our study sites, the monitor-specific coefficients show considerable heterogeneity in the outdoor−indoor relationship for all particle sizes (Figures S3 and 4A). We probe the source of this heterogeneity by regressing the monitor-specific penetration rates against fixed factors at each location (Table S6). We find that local weather conditions and the age of building stock play a significant role in explaining this heterogeneity. Lower winter temperatures, higher summer temperatures, and higher dew point temperatures all increase the degree of penetration. The summer and dew point temperature results likely reflect the role of open windows, while the winter temperature may reflect an increased role for heating. It may also reflect the varied topography of California, where inland and mountain regions tend to have hotter summers and colder winters than the coastal region. In terms of magnitude, however, the building stock result dominates. Buildings constructed before 1960, which were subject to quite lax building codes, exhibit more indoor penetration. To put this relationship in context, it is important to note that the average percent of buildings constructed after 1960 in a zip code is 54%. As such, the coefficient on building stock suggests that it explains 40% of a standard deviation of the heterogeneity in estimates for the 0.3−0.5 μm size bin and 33% for the 5−10 μm size bin.

DISCUSSION

Our study utilizes crowd-sourced data from across the state of California to analyze the relationship between outdoor and indoor particulate matter pollution. We find that average indoor particulate levels are 40−70% of nearby outdoor levels, consistent with the lower end of estimates found in prior studies that more narrowly targeted a small number of study sites and/or shorter time periods as well as one large-scale study conducted in China where the climate and heating fuel sources differ considerably from our study.[20] Outdoor particulate matter rapidly penetrates indoors. Nearly half of the penetration occurs within the first hour, with the balance slowly accruing over the next 5 h.

This pattern holds across particle sizes, with the majority of penetration occurring within the first few hours, but overall penetration is larger for fine particles than for coarse ones (i.e., less than or more than 2.5 μm, respectively). This latter finding is particularly concerning from a health perspective because smaller particles pose significantly more harm to human health.[26,27] As temperature increases, so does the rate of penetration, particularly for the largest particles. A leading explanation is the opening of household windows, especially because of the pattern by size where larger particles need larger openings to penetrate inside.[29]

Despite the strength of our design, there are several important limitations in our analysis. First, we lack information on pollution-generating activities within the home. This will be especially important when considering use and interpretation of PA monitor data in places with high indoor primary PM generation (e.g., from cooking, gas and wood stoves, fireplaces, fuel oil boilers).[30−34] While this matters for understanding all factors that contribute to indoor pollution, it should only bias our estimates of the outdoor−indoor relationship if indoor emissions correlate with outdoor pollution levels, a threat limited by our use of multiway fixed effects. In particular, monitor fixed effects control for all time-invariant characteristics of the monitor location, such as proximity to a highway and indoor cooking source. Month fixed effects control for all unobserved factors constant for each particular month, such as the occurrence of a wildfire. Hourly fixed effects control for all unobserved factors constant for each particular hour across all locations. Hourly weather variables also control for time-varying changes within a day that vary by location. Therefore, the time-varying unobserved factor that our model does not account for is one that varies within a particular month, on a particular hour, and for a particular location that is not perfectly correlated with weather. For example, if people alter the way they use indoor stoves or the frequency with which they open and close windows at certain times of the day when the wildfires occurred, then this would be a source of bias. While we cannot rule out such changes, we believe this threat is limited in our setting.

Second, PurpleAir monitors are purchased directly by consumers, and thus our sample of observations may not paint a representative picture of the outdoor−indoor relationship across the State of California. This reflects a common conundrum with the use of crowd-sourced data, though we contend these potential limitations are compensated for by the increased sample size and ability to extend our analysis relative to prior studies, which also have their own issues of representativeness.

Third, in contrast to results from other outdoor−indoor studies, our analysis is based on indoor and outdoor monitors

https://doi.org/10.1021/acs.est.0c08469
Environ. Sci. Technol. 2021, 55, 6107−6115

within 500 m of each other, rather than co-located at a particular site. While it reduces our sample size, we are reassured by comparable results obtained when we restrict to co-located monitors (within 50 m of each other) (see appendix Table S7). Fourth, PA uses a different technology to measure particulate matter from the much more expensive (but regularly calibrated) equipment deployed by the EPA monitoring system. PA uses laser particle counters to measure a number density of particles in each size bin, and then makes assumptions about particle chemistry and density to convert number densities to the more familiar mass-based measurements. EPA monitors directly measure particulate mass under the 2.5 μm size threshold accumulated on filters over a given time period. We report direct bin-by-bin number density relationships to avoid making assumptions about the particulate mix, but work incorporating monitors of different types will need to address known discrepancies.[35] Fortunately, the measures from PurpleAir monitors are highly correlated with those from EPA monitors.[35−38]

Our results have several important implications for understanding the public health benefits from improvements in air quality. First, the primary regulatory tool to protect citizens from the harmful impacts of air pollution is through restrictions on ambient pollution levels. While emissions from particular indoor sources can be controlled through product standards, total indoor levels cannot because they will depend on the accumulated decisions within a residence as well as the ways in which consumers use those products. In contrast, restrictions on ambient pollution levels offer public health protection in the form of reductions in direct outdoor exposure as well as reductions that arise due to indoor penetration. This composite effect, referred to as a concentration−response function, is what is typically measured in epidemiologic studies of pollution, even if the composition of this effect is rarely emphasized.[39]

Knowledge of the outdoor−indoor penetration, however, enables us to better isolate the biological effect, or exposure−response function, from the concentration−response function, which is a composite measure that reflects the ways in which ambient pollution levels are influenced by environmental and behavioral circumstances to translate into health impacts.[40,41] This distinction is important because the biological effect is likely to generalize across populations (with perhaps some important underlying heterogeneity), whereas the composite effect may not because exposure is mitigated by local circumstances.[42] For example, time spent indoors and the degree of outdoor−indoor penetration, driven by factors such as building stock, ventilation services, and climatic conditions, may vary by location such that estimates of the relationship between ambient pollution concentrations and health differ despite similar biological effects.

Second, our findings also have important implications for policies related to pollution avoidance. In many parts of the world, air quality alerts are designed to encourage individuals to spend more time indoors on polluted days.[43,44] For some pollutants, such as ozone, this is clearly an effective approach, but that protection is limited for PM given how rapidly it penetrates indoors. Air Quality Indices are a composite measure of air pollution that are constructed by taking the maximum over a piecewise-linear transformation of daily readings for a suite of pollutants. The value of the index is "triggered" by the pollutant with the highest transformed value.[45] At a minimum, our results suggest that air quality reporting systems should indicate the "triggering" pollutant so that individuals can make more informed decisions regarding their avoidance behavior.

Finally, explicit attention to indoor penetration of outdoor pollution also helps to broaden the regulatory conversation. Current regulatory approaches focus on limiting emission sources through mitigation technologies, such as scrubbers and catalytic converters. Since most human exposure to ambient pollution occurs indoors, policies to reduce this source of exposure through devices such as personal air filters may also have an important role to play. This is most clearly the case when emissions cannot be controlled by government policy, as is typically the case for nonanthropogenic emission sources (e.g 46.,) and when pollution crosses national boundaries (e.g., ref 47). One important caveat is that, in contrast to source-based policies that confer benefits on everyone, the purchase of air filters depends on income and the price to purchase and maintain those devices; as such, receptor-based policies may exacerbate inequalities absent policy interventions to counteract their potential to be regressive. Evidence from China indicates that lower-income families are less likely to purchase such filters.[48] In the end, receptor-based approaches may prove most useful for policy targeting, where they can be used as a supplemental approach to protect the most vulnerable,[49] such as pregnant women, infants, and those with respiratory ailments.[50−53]

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.est.0c08469.

> Cartoon schematic of two particle distributions (Figure S1); size bin distributions of indoor and outdoor particles in the study sample (Figure S2); distribution of coefficients from monitor-specific penetration rates (Figure S3); main coefficients from OLS regressions (Tables S1−S3 and S5−S8); and comparison of the daily coefficients to different sums of the hourly coefficients (Table S4) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Benjamin Krebs − *Faculty of Economics and Management, University of Lucerne, CH-6002 Luzern, Switzerland;* orcid.org/0000-0001-6870-1514; Email: benjamin.krebs@unilu.ch

Authors

Jennifer Burney − *School of Global Policy and Strategy, University of California, San Diego, La Jolla, California 92093, United States;* orcid.org/0000-0003-3532-2934

Joshua Graff Zivin − *School of Global Policy and Strategy, University of California, San Diego, La Jolla, California 92093, United States*

Matthew Neidell − *Mailman School of Public Health, Columbia University, New York 10032, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.est.0c08469

Notes

The authors declare no competing financial interest.

https://doi.org/10.1021/acs.est.0c08469
Environ. Sci. Technol. 2021, 55, 6107−6115

■ ACKNOWLEDGMENTS

For research assistance, the authors thank Peter Wallmüller. J.B. was supported by the Robert Wood Johnson Foundation (grant #76555). B.K. acknowledges funding by the Swiss National Science Foundation (fellowship P1LUP1 187787).

■ REFERENCES

(1) Burnett, R.; Chen, H.; Szyszkowicz, M.; Fann, N.; Hubbell, B.; Pope, C. A.; Apte, J. S.; Brauer, M.; Cohen, A.; Weichenthal, S.; et al. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, 9592−9597.

(2) Anenberg, S. C.; Achakulwisut, P.; Brauer, M.; Moran, D.; Apte, J. S.; Henze, D. K. Particulate matter-attributable mortality and relationships with carbon dioxide in 250 urban areas worldwide. *Sci. Rep.* **2019**, *9*, No. 11552.

(3) Li, A.; Ren, T.; Yang, C.; Lv, W.; Zhang, F. Study on particle penetration through straight, L, Z and wedge-shaped cracks in buildings. *Build. Environ.* **2017**, *114*, 333−343.

(4) Liu, C.; Yang, J.; Ji, S.; Lu, Y.; Wu, P.; Chen, C. Influence of natural ventilation rate on indoor PM2.5 deposition. *Build. Environ.* **2018**, *144*, 357−364.

(5) *EPA Report to Congress on Indoor Air Quality, Volume II: Assessment and Control of Indoor Air Pollution*, Technical Report EPA/400/1-89/001C; U.S. Environmental Protection Agency, 1989.

(6) Klepeis, N. E.; Nelson, W. C.; Ott, W. R.; Robinson, J. P.; Tsang, A. M.; Switzer, P.; Behar, J. V.; Hern, S. C.; Engelmann, W. H. The National Human Activity Pattern Survey (NHAPS): A resource for assessing exposure to environmental pollutants. *J. Exposure Sci. Environ. Epidemiol.* **2001**, *11*, 231−252.

(7) Chen, C.; Zhao, B. Review of relationship between indoor and outdoor particles: I/O ratio, infiltration factor and penetration factor. *Atmos. Environ.* **2011**, *45*, 275−288.

(8) Dockery, D. W.; Spengler, J. D. Indoor-outdoor relationships of respirable sulfates and particles. *Atmos. Environ. (1967)* **1981**, *15*, 335−343.

(9) Thatcher, T. L.; Layton, D. W. Deposition, resuspension, and penetration of particles within a residence. *Atmos. Environ.* **1995**, *29*, 1487−1497.

(10) Özkaynak, H.; Xue, J.; Spengler, J.; Wallace, L.; Pellizzari, E.; Jenkins, P. Personal exposure to airborne particles and metals: Results from the Particle TEAM study in Riverside, California. *J. Exposure Anal. Environ. Epidemiol.* **1996**, *6*, 57.

(11) Brickus, L. S.; Cardoso, J. N.; de Aquino Neto, F. R. Distributions of indoor and outdoor air pollutants in Rio de Janeiro, Brazil: Implications to indoor air quality in bayside offices. *Environ. Sci. Technol.* **1998**, *32*, 3485−3490.

(12) Abt, E.; Suh, H. H.; Catalano, P.; Koutrakis, P. Relative contribution of outdoor and indoor particle sources to indoor concentrations. *Environ. Sci. Technol.* **2000**, *34*, 3579−3587.

(13) Jones, N.; Thornton, C.; Mark, D.; Harrison, R. Indoor/outdoor relationships of particulate matter in domestic homes with roadside, urban and rural locations. *Atmos. Environ.* **2000**, *34*, 2603−2612.

(14) Götschi, T.; Oglesby, L.; Mathys, P.; Monn, C.; Manalis, N.; Koistinen, K.; Jantunen, M.; Hänninen, O.; Polanska, L.; Künzli, N. Comparison of black smoke and PM2.5 levels in indoor and outdoor environments of four European cities. *Environ. Sci. Technol.* **2002**, *36*, 1191−1197.

(15) Allen, R.; Larson, T.; Sheppard, L.; Wallace, L.; Liu, L.-J. S. Use of real-time light scattering data to estimate the contribution of infiltrated and indoor-generated particles to indoor air. *Environ. Sci. Technol.* **2003**, *37*, 3484−3492.

(16) Wallace, L. A.; Mitchell, H.; O'Connor, G. T.; Neas, L.; Lippmann, M.; Kattan, M.; Koenig, J.; Stout, J. W.; Vaughn, B. J.; Wallace, D.; Walter, M.; Adams, K.; Sally Liu, L.-J. Particle concentrations in inner-city homes of children with asthma: The effect of smoking, cooking, and outdoor pollution. *Environ. Health Perspect.* **2003**, *111*, 1265−1272.

(17) Long, C. M.; Sarnat, J. A. Indoor-outdoor relationships and infiltration behavior of elemental components of outdoor PM2.5 for Boston-area homes. *Aerosol Sci. Technol.* **2004**, *38*, 91−104.

(18) Brunekreef, B.; Janssen, N. A. H.; de Hartog, J. J.; Oldenwening, M.; Meliefste, K.; Hoek, G.; Lanki, T.; Timonen, K. L.; Vallius, M.; Pekkanen, J.; Van Grieken, R. *Personal, Indoor, and Outdoor Exposures to PM2.5 and Its Components for Groups of Cardiovascular Patients in Amsterdam and Helsinki*, Research Report; Health Effects Institute, 2005; pp 1−70.

(19) Meng, Q. Y.; Spector, D.; Colome, S.; Turpin, B. Determinants of indoor and personal exposure to PM2.5 of indoor and outdoor origin during the RIOPA study. *Atmos. Environ.* **2009**, *43*, 5750−5758.

(20) Zuo, J.; Ji, W.; Ben, Y.; Hassan, M. A.; Fan, W.; Bates, L.; Dong, Z. Using big data from air quality monitors to evaluate indoor PM2.5 exposure in buildings: Case study in Beijing. *Environ. Pollut.* **2018**, *240*, 839−847.

(21) Air Quality Sensor Performance Evaluation Center, Field Evaluation: Purple Air (PM-II) PM Sensor. http://www.aqmd.gov/docs/default-source/aq-spec/field-evaluations/346purple-air-pa-ii-field-evaluation.pdf?sfvrsn=4 (accessed August 22, 2020).

(22) Levy Zamora, M.; Xiong, F.; Gentner, D.; Kerkez, B.; Kohrman-Glaser, J.; Koehler, K. Field and laboratory evaluations of the low-cost Plantower particulate matter sensor. *Environ. Sci. Technol.* **2019**, *53*, 838−849.

(23) Yong, Z. Digital Universal Particle Concentration Sensor. http://www.aqmd.gov/docs/default-source/aq-spec/resources-page/plantower-pms5003-manual_v2-3.pdf (accessed August 22, 2016).

(24) Yamamoto, N.; Shendell, D.; Winer, A.; Zhang, J. Residential air exchange rates in three major US metropolitan areas: Results from the Relationship Among Indoor, Outdoor, and Personal Air Study 1999-2001. *Indoor Air* **2010**, *20*, 85−90.

(25) Johnson, T.; Myers, J.; Kelly, T.; Wisbith, A.; Ollison, W. A pilot study using scripted ventilation conditions to identify key factors affecting indoor pollutant concentration and air exchange rate in a residence. *J. Exposure Sci. Environ. Epidemiol.* **2004**, *14*, 1−22.

(26) Schwartz, J.; Dockery, D. W.; Neas, L. M. Is daily mortality associated specifically with fine particles? *J. Air Waste Manage. Assoc.* **1996**, *46*, 927−939.

(27) Klemm, R. J.; Mason, R. M., Jr.; Heilig, C. M.; Neas, L. M.; Dockery, D. W. Is daily mortality associated specifically with fine particles? Data reconstruction and replication of analyses. *J. Air Waste Manage. Assoc.* **2000**, *50*, 1215−1222.

(28) Lippmann, M.; Chen, L. C.; Gordon, T.; Ito, K.; Thurston, G. D. *National Particle Component Toxicity (NPACT) Initiative: Integrated Epidemiologic and Toxicologic Studies of the Health Effects of Particulate Matter Components*, Research Report; Health Effects Institute, 2013; p 5.

(29) Long, C. M.; Suh, H. H.; Catalano, P. J.; Koutrakis, P. Using time- and size-resolved particulate data to quantify indoor penetration and deposition behavior. *Environ. Sci. Technol.* **2001**, *35*, 2089−2099.

(30) Spengler, J. D.; Sexton, K. Indoor air pollution: A public health perspective. *Science* **1983**, *221*, 9−17.

(31) Zhang, J.; Smith, K. R. Indoor air pollution: A global health concern. *Br. Med. Bull.* **2003**, *68*, 209−225.

(32) Delp, W. W.; Singer, B. C. Performance assessment of US residential cooking exhaust hoods. *Environ. Sci. Technol.* **2012**, *46*, 6167−6173.

(33) Salthammer, T.; Schripp, T.; Wientzek, S.; Wensing, M. Impact of operating wood-burning fireplace ovens on indoor air quality. *Chemosphere* **2014**, *103*, 205−211.

(34) Chen, C.; Zhao, Y.; Zhao, B. Emission rates of multiple air pollutants generated from Chinese residential cooking. *Environ. Sci. Technol.* **2018**, *52*, 1081−1087.

(35) Tryner, J.; L'Orange, C.; Mehaffy, J.; Miller-Lionberg, D.; Hofstetter, J. C.; Wilson, A.; Volckens, J. Laboratory evaluation of low-cost PurpleAir PM monitors and in-field correction using co-located portable filter samplers. *Atmos. Environ.* **2020**, *220*, No. 117067.

https://doi.org/10.1021/acs.est.0c08469
Environ. Sci. Technol. 2021, 55, 6107−6115

(36) Sayahi, T.; Butterfield, A.; Kelly, K. Long-term field evaluation of the Plantower PMS low-cost particulate matter sensors. *Environ. Pollut.* **2019**, *245*, 932−940.
(37) Bi, J.; Wildani, A.; Chang, H. H.; Liu, Y. Incorporating low-cost sensor measurements into high-resolution PM2.5 modeling at a large spatial scale. *Environ. Sci. Technol.* **2020**, *54*, 2152−2162.
(38) Stavroulas, I.; Grivas, G.; Michalopoulos, P.; Liakakou, E.; Bougiatioti, A.; Kalkavouras, P.; Fameli, K. M.; Hatzianastassiou, N.; Mihalopoulos, N.; Gerasopoulos, E. Field evaluation of low-cost PM sensors (Purple Air PA-II) under variable urban air quality conditions, in Greece. *Atmosphere* **2020**, *11*, No. 926.
(39) Pope, C. A., III; Cropper, M.; Coggins, J.; Cohen, A. Health benefits of air pollution abatement policy: Role of the shape of the concentration-response function. *J. Air Waste Manage. Assoc.* **2015**, *65*, 516−522.
(40) Cox, L. A. T. Effects of exposure estimation errors on estimated exposure-response relations for PM2.5. *Environ. Res.* **2018**, *164*, 636−646.
(41) Daniels, M. J.; Dominici, F.; Samet, J. M.; Zeger, S. L. Estimating particulate matter-mortality dose-response curves and threshold levels: an analysis of daily time-series for the 20 largest US cities. *Am. J. Epidemiol.* **2000**, *152*, 397−406.
(42) Graff Zivin, J.; Neidell, M. J. Environment, health, and human capital. *J. Econ. Lit.* **2013**, *51*, 689−730.
(43) Saberian, S.; Heyes, A.; Rivers, N. Alerts work! Air quality warnings and cycling. *Resour. Energy Econ.* **2017**, *49*, 165−185.
(44) Neidell, M. J. Air pollution, health, and socio-economic status: The effect of outdoor air quality on childhood asthma. *J. Health Econ.* **2004**, *23*, 1209−1236.
(45) Mintz, D. *Technical Assistance Document for the Reporting of Daily Air Quality—The Air Quality Index (AQI)*; U.S. Environmental Protection Agency, Office of Air Quality Planning and Standards, 2012; pp 1−26.
(46) Evan, A. T.; Flamant, C.; Gaetani, M.; Guichard, F. The past, present and future of African dust. *Nature* **2016**, *531*, 493−495.
(47) Persad, G. G.; Caldeira, K. Divergent global-scale temperature effects from identical aerosols emitted in different regions. *Nat. Commun.* **2018**, *9*, No. 3289.
(48) Ito, K.; Zhang, S. Willingness to pay for clean air: Evidence from air purifier markets in China. *J. Political Econ.* **2020**, *128*, 1627−1672.
(49) Graff Zivin, J.; Zilberman, D. Optimal environmental health regulations with heterogeneous populations: Treatment versus "tagging". *J. Environ. Econ. Manage.* **2002**, *43*, 455−476.
(50) Chay, K. Y.; Greenstone, M. The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *Q. J. Econ.* **2003**, *118*, 1121−1167.
(51) Currie, J.; Neidell, M. Air pollution and infant health: What can we learn from California's recent experience. *Q. J. Econ.* **2005**, *120*, 1003−1030.
(52) Currie, J.; Neidell, M.; Schmieder, J. F. Air pollution and infant health: Lessons from New Jersey. *J. Health Econ.* **2009**, *28*, 688−703.
(53) Isen, A.; Rossin-Slater, M.; Walker, W. R. Every breath you take−every dollar you'll make: The long-term consequences of the Clean Air Act of 1970. *J. Political Econ.* **2017**, *125*, 848−902.

https://doi.org/10.1021/acs.est.0c08469
Environ. Sci. Technol. 2021, 55, 6107−6115