Simultaneous measurement of health damage from transboundary and domestic air pollution in mixture*

Jaecheol Lee[§,1,2], Andrew J. Wilson[†,§,1,3], and Solomon Hsiang[1,4,5]

[1]Global Policy Laboratory, Doerr School of Sustainability, Stanford University
[2]Climate Impact Lab, University of Chicago
[3]Center on Food Security and the Environment, Stanford University
[4]Doerr School of Sustainability, Stanford University
[5]National Bureau of Economic Research

November 11, 2024

Click here for the most recent version

*We acknowledge valuable feedback from participants in the Seminar on Planetary Management, the Environment and Resource Economics and GPL Doctoral Fellows seminars at UC Berkeley, the Onteora Library speaker series, the Sustainable Development Seminar and Colloquium at Columbia University, the Coase Project at University of Chicago, the Occasional Workshop at UC Santa Barbara, TWEEDS, as well as thoughtful comments from Sebastien Annan-Phan, Shuyang Fang, Alejandro Favela Nava, Jenya Kahn-Lang, Minwoo Hyun, Jieun Kim, Junho Lee, Joseph Shapiro, and Reed Walker. The Global Policy Laboratory at UC Berkeley generously funded charges to access the dataset used in this analysis.

Funding: JL is supported by the Center for Korean Studies, the Institute of East Asian Studies, and the Center for Chinese Studies, all at UC Berkeley.

This study analyzed secondary data on human subjects under Korea National Institute for Bioethics Policy IRB P01-201811-22-008.

Any opinions, findings, conclusions, or recommendations expressed in this material are our views alone and do not necessarily reflect the views of organizations that support us. We declare no conflicts of interest.

[§]Equal contribution
[†]Correspondence: ajw2206@columbia.edu

Abstract

Background Particulate matter (PM) is a major, clinically important air pollutant. A large portion of emitted PM crosses borders, damaging health outside of its originating jurisdiction, but due to technical obstacles these pollutant flows remain unregulated. Proposed approaches also assume that units of PM originating in different jurisdictions cause the same harm, despite a widespread understanding that differing chemical and physical features of PM could generate distinct health effects.

Methods We use an atmospheric model to decompose the origins of PM individuals are exposed to at each location in South Korea, the nexus of the world's most contentious transboundary air pollution disputes, every day during 2005–2016. We then link these data to universal healthcare records in an econometric analysis that simultaneously measures and accounts for harms from seven types of PM, each from a distinct origin.

Findings We discover that the health harm of a unit of transboundary PM is approximately 5 times (North Korea) and 2.6 times (China) greater than a unit of PM originating within South Korea, and that the health response to natural dust differs from anthropogenic sources. Because harms differ by origin, we compute that transboundary sources contribute only 43

Interpretation Our results suggest that PM should be treated in research, WHO standards, and international treaties as a mixture of distinct pollutants, each with a unique measurable impact on human health.

Funding Jaecheol Lee was supported by the Center for Korean Studies, the Institute of East Asian Studies, and the Center for Chinese Studies, all at UC Berkeley.

Introduction

Particulate matter (PM) pollution has significant negative effects on health and well-being, altering risks of many outcomes, including cardiovascular and respiratory illnesses, dementia, suicide, and even crime [1–5]. Thus, to inform the public, support research, and provide a basis for regulation, PM is now widely measured worldwide. These measurements typically classify PM by particle diameter and are reported as a size-specific mass density for all particulates in a sample of air. However, this aggregation obscures differences in PM constituents, which include disparate species like oxidized volatile organic compounds from industrial complexes and mineral dust from desert erosion. Prior research has shown that differences in physical and chemical characteristics can affect the toxicity of PM in a laboratory setting [6–11]. Further, observational studies have shown that health outcomes are correlated with the average physical and chemical attributes of PM that groups are exposed to [12–17], although it is unknown if those associations are causal. Other papers have assessed the causal effects of pulses of PM from significant natural sources, including wildfire [18, 19] and dust storms [20, 21], but have not conclusively distinguished these effects from those of background PM. Nonetheless, taken together, this collection of facts has led researchers to hypothesize that the same measured quantities of PM from different jurisdictions may have distinct population-level health consequences [10, 14–17, 22]. We provide the first direct and unconfounded test of this hypothesis.

Distinguishing origin-specific health impacts can dramatically influence air quality management, since pollution is managed at its origin but reflects the scale of impacts in downstream locations [23, 24]. Air quality management maximizes social welfare when regulators impose shadow costs on air pollution emissions that equate the marginal benefits and marginal costs of regulation. In the case of sources emitting a uniform air pollutant like carbon monoxide with heterogeneity only in exposure, the efficiency gains from regulation that imposes differentiated (as opposed to uniform) shadow costs on emitters can be large [25, 26] but depend on the variance and uncertainty of marginal damages and abatement costs across emitters [27]. In the case of a heterogeneous class of air pollutants like particulate matter, the variance and uncertainty of emitter-specific marginal damages may be higher, which would suggest a broader range of potential gains from differentiation.

When evaluating marginal damages for transboundary air pollution, regulators must draw a causal link between emissions in one jurisdiction and realized damages in another. State-of-the-art long-range air pollution studies approach this challenge by linking emitters to recipient locations using atmospheric models [28–40] and then assuming PM has uniform toxicity; this leads to an allocation of responsibility for harms that reflects only the *quantity* of exposure to PM from each origin. However, if toxicity varies by jurisdiction, this approach will misallocate responsibility for damages—possibly to a large degree. In an impossible and unethical "ideal experiment," a researcher would distinguish the health effects of PM from different emitters by experimentally exposing a fixed set of subjects to different, known mixtures of PM from multiple origins and observing the resulting health outcomes. Our quasi-experimental approach approximates this "ideal experiment" using observational data.

Specifically, we simultaneously measure the health impacts of PM from multiple origins on the population of South Korea, which is at the nexus of the world's largest and most contentious transboundary air pollution disputes [34, 41–45]. We decompose the daily mixture of total PM_{10} (particulates less than 10 micrometers in diameter, which includes $PM_{2.5}$; henceforth "PM") observed at a location into contributions from prominent emitters and measure how changes in PM from each origin independently influence daily health costs. Crucially, by measuring the impact of each type of PM simultaneously, health costs are not "double counted" (i.e., a new hospital visit cannot be attributed to domestic pollution and then again to transboundary pollution), and each estimate of origin-specific toxicity accounts for the impacts related to doses of incident PM from all other origins. To do this, we first use an atmospheric transport model to decompose the provenance of local PM every 3 hours into seven sources: nonanthropogenic sources (including sea salt, nonanthropogenic aeolian dust [which we refer to as "dust"], and forest wildfire), anthropogenic activity (which we group by originating jurisdiction: South Korea, China, and North Korea), and "other sources." Collectively, we call this set of nonanthropogenic sources and anthropogenic jurisdictions "origins." We then combine these probabilistic estimates with high-resolution data on medical spending from South Korea's universal healthcare system, which tracks the public and private medical spending of 97% of South Korea's 52 million residents. The data allow us to de-

convolve daily fluctuations of respiratory health costs within each locality into contributions from random variation in origin-specific PM doses over time. To the best of our knowledge, this study is the first to directly and separately measure the simultaneous health effects of both domestic and transboundary PM.

Analysis

Partitioning observed PM by origin We create a mapping that assigns PM monitor readings on individual days and locations (i.e., one of South Korea's 147 districts with PM monitoring; see Supplementary Materials (SM) Figure S1) to a probability distribution over seven origins—China, South Korea, North Korea, wildfire, mineral dust, sea salt, and "other sources." To do this, we first isolate the influence of dust and sea salt on PM levels using the Copernicus Atmosphere Monitoring Service Global Reanalysis 4 (see SM section B for additional details) [46]. To apportion the remaining PM, we then generate a 240-hour backward trajectory starting every three hours at eight altitudes from each PM monitor ($N = 264$) on each day during 2005–2016 (4383 days) using an atmospheric transport model [47, 48]. This results in 67.9 million air parcel trajectories, each defined by 241 points (10 days of hourly observations; see Figure 1A for an example month at a single PM monitor). These trajectories enable us to estimate the contribution of various locations to the air arriving at a South Korean district on a given day. We combine these trajectories with information on the spatiotemporal distribution of emissions from anthropogenic sources [49] and fires [50–52] (applying an adjustment for chemical scavenging and deposition over time) to estimate the probability that a unit of PM arriving at location "i" originated at location "j." Figures 1B–C depict examples of the distribution of these probabilities across space (all j's) for the district of Busan (i) on two days in April 2016. We integrate these probability distributions over hours, altitudes, and originating jurisdictions to construct district-specific time series of PM contributions, which are then re-scaled to partition station-measured PM into contributions from all seven origins (e.g., Figure 1D). The resulting origin-specific PM time series (e.g., Figure 1E) thus encapsulate what is known about the emissions, transport, and ground measurement of all PM arriving in South

Korea during this period. See SM section B for more details.

Measuring impacts of exposure on health outcomes We jointly resolve the effects of all seven origins of PM on spending for respiratory illnesses for the entire population of South Korea. These district-specific data are from a representative sample of individuals covered by South Korea's universal healthcare system (N = 4.87 million) and include expenses for outpatient and emergency visits due to respiratory illness at all healthcare facilities in South Korea (e.g., see Figure 1F; for more details on this data, see "Data Collection" section in SM). We empirically estimate the health effects observed for a given district's population when it is exposed to different mixtures of PM (see Methods). Our approach accounts for complex, nonlinear time trends in medical expenditures (i.e., seasonal patterns, secular trends, variations across days of the week, and the impacts of holidays), differences in time-invariant subnational characteristics that may affect health spending (e.g., levels of wealth, the quality of local public health provision, or differences in data quality), non-PM pollutants (nitrogen dioxide, carbon monoxide, ozone, and sulfur dioxide), and changes in weather previously shown to impact human health [53]. We account for cumulative effects over a period sufficiently long (4 weeks) that we capture delayed changes in health expenditure associated with PM exposure (see Methods). In this context, delays can result from many factors: for example, respiratory symptoms may unfold over several days or weeks after pollution exposure, or South Korea's medical appointment referral system may lead to a gap between the initial and final visit for a given health issue. Finally, our empirical specification allows us to flexibly account for the possibility of avoidance behavior that individuals undertake autonomously to limit their total PM exposure (e.g., staying indoors or using masks) while still preserving our ability to distinguish between the effects of subcomponents of overall PM (see SM section C). Finally, we assess the robustness of our results to different methods for calculating uncertainty, alternative functional forms for how we model avoidance and health responses, and whether we account for air quality alerts; we also assess the consistency of our results over space, time, and seasons (for a detailed explanation of these tests, see Figure S4).

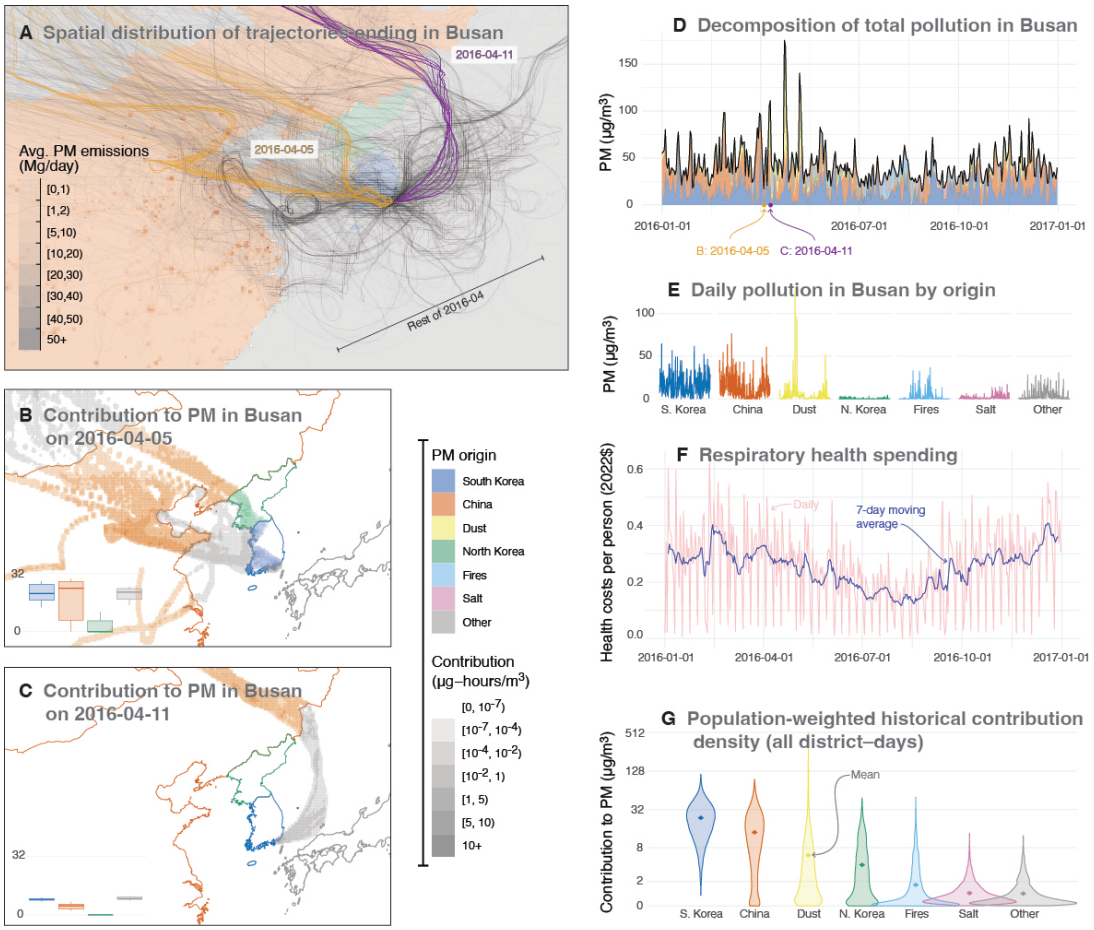


Figure 1: **Example decomposition of PM by origin.** (A) Example set of backtrajectories from a single district in Busan, South Korea, during April 2016. The full analysis includes all districts shown in Figure S1. Trajectories for April 5 and April 11 are shown in orange and purple, respectively. Anthropogenic PM emissions during this month are shown as changes in the opacity of the background map, with colors differing by originating jurisdiction. (B) The estimated contribution of each location's emissions to anthropogenic PM exposure experienced in Busan on April 5, 2016 and the distribution of estimates contributions of different origins across the set of backtrajectories (lower left box and whisker plot). (C) Same as (B), but for April 11, 2016. (D) Time series of total PM (black line) measured in Busan throughout 2016, with the two dates in (B)–(C) labeled. Stacked color areas indicate the decomposition of this measured PM by origin based on backtrajectory calculations illustrated in (A)–(C). (E) Time series for each component of total PM by origin for 2016. (F) Daily (pink) and seven-day moving average (blue) time series for respiratory health spending per capita in Busan throughout 2016. (G) Distributions of nationwide (not just Busan) population-weighted PM exposure over each of the seven PM origins we investigate. Distributions include all districts and days in our sample, with the mean exposure for each PM origin shown with a diamond shape (note nonlinear y-axis).

Results

Exposure by PM origin

During the period of our analysis, we estimate that three anthropogenic source regions—China, South Korea, and North Korea—account for roughly 80.9% of total population-weighted PM exposure in South Korea; mineral dust accounts for an additional 11.7%; and wildfire, sea salt, and all other sources together account for the remaining <8% of total exposure (Figure 1G). We also estimate that seasonal and spatial differences in exposure across South Korea are large. For example, while sea salt is on average a small component of PM (about 3%), it is a large portion of total PM in less industrial, coastal districts during summer (we estimate that sea salt accounts for over 50% of PM on 0.2% of district–day observations). Likewise, PM from China is highest across all districts in the winter and the spring, when prevailing winds are westerly. During this period, some regions occasionally experience "Yellow Dust" events; when concentrations of PM in South Korea from soil erosion in the deserts of China, Mongolia, and Kazakhstan can rise above 1000 $\mu g/m^3$. These events account for a large portion of total dust exposure, as contributions of dust are generally <5 $\mu g/m^3$. Because most PM (and most of the resulting health damage) originates from South Korea, China, and North Korea—and because these emissions can be modified by human activities—we focus our analysis primarily on anthropogenic PM from these three origins, but note that all analyses account for PM from all seven origins.

Effects of origin on properties of PM mixtures

We analyze the observed chemical and physical properties of PM to validate that our decomposition of PM origin identifies distinct components of the PM mixture incident on South Korean locations. We use the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) atmospheric transport model [47] because it was designed for backtrajectory calculation, and its computational efficiency is crucial for handling the large number of trajectories we analyze. This model has been widely validated for many applications [54–59], although it is possible, in principle, that the absence of active chemistry calculations [60–64] or other aspects of the model could impact our estimates.

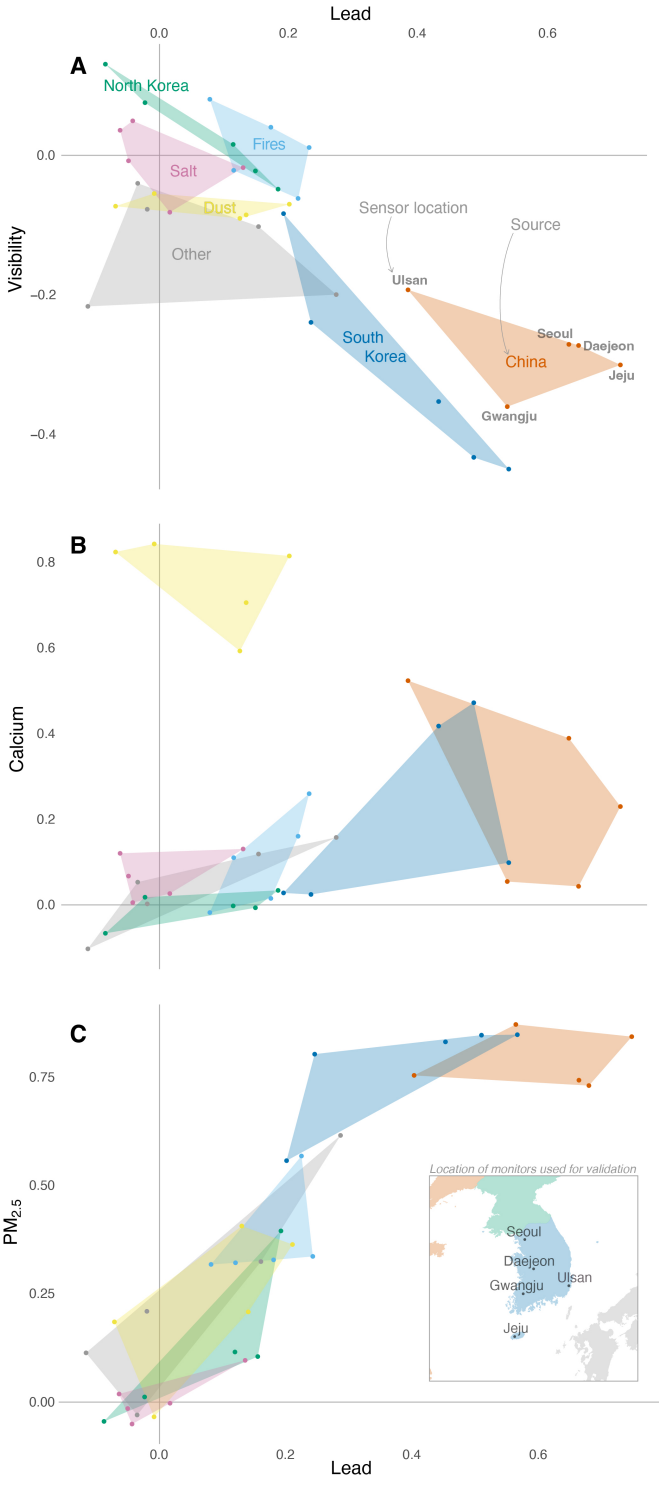


Figure 2: **Estimating the chemical and physical signature of PM from each origin at five locations in South Korea.** The partial correlations between the quantity of PM from different origins and concentrations of either lead, calcium, fine particulate matter ($PM_{2.5}$), or visibility across five monitors (see inset). Monitor locations are labeled in Panel A. In each panel, marker positions depict partial correlations between aggregate properties (for both x- and y-axes) measured at a given monitor and PM from each origin. Correlations for all origins are estimated simultaneously for each property at each sensor, but colored to group measurements across sensors based on PM origin. Colored regions depict the convex hull of these partial correlations for a single origin across the set of five monitors, indicating the chemical and physical "fingerprint" for each origin. Further explanation of this figure is shown in Figure S2.

We cannot directly test that hourly backtrajectories are correctly modeled; however, we validate that PM contributions from a specific origin appear (i) chemically and physically consistent across South Korea and (ii) chemically and physically distinct from PM attributed to other origins. We do this by identifying the chemical and physical "fingerprint" of PM from each origin, comparing it across sensors, and contrasting it with the "fingerprint" from other origins.

Aggregate chemical characteristics (lead and calcium levels) and physical properties (visibility and $PM_{2.5}$ concentration) of PM are monitored daily at five locations (see Figure 2 inset) and for limited periods of time (see SM section A). This quantity of data is insufficient for population-scale epidemiological analysis and does not quantify many important physical and chemical aspects of PM, but it can nonetheless validate our calculation of origin-specific contributions. If PM from different origins has distinct chemical and physical characteristics and we have correctly decomposed PM contributions by origin, then the observed aggregate properties of the PM mixture should be an approximately linear combination of properties for components in the mixture—with weights that reflect the fractional contribution from each origin (see SM section C). For example, on days when we compute that the PM mixture over Seoul is 90% from China, based on backtrajectories, then we would expect the chemical properties of PM on those days to be dominated by the properties of emissions from China. Based on this idea, we estimate how an influx of PM from each origin alters the observed aggregate chemical and physical properties of the PM mixture, accounting for the estimated contribution and properties of all other origins simultaneously. Specifically, we use multiple regression to decompose how PM from all origins simultaneously impacts aggregate chemical and physical properties at each measurement site (see SM section C).

We find that our estimates of PM origin are associated with consistent and distinct changes in lead, calcium, visibility, and the concentration of $PM_{2.5}$ (Figure 2). Three features of this result are notable. First, patterns across monitors are relatively consistent for each PM origin. For example, the chemical/physical signature of domestic-origin or Chinese PM is broadly consistent across all sensors in South Korea (high lead, low visibility, moderate calcium, and high $PM_{2.5}$). Second, the four-dimensional chemical/physical signature of each origin is well-separated from the signatures of other origins (note that overlap in Figure 2 primarily results from display in only two dimensions at

a time), implying that our approach to isolating PM by origin identifies collections of PM exposures that are physically and chemically distinct from one another.

Third, the patterns we recover are consistent with what is previously known about PM source characteristics. For example, we find that PM contributions from China and South Korea are associated with higher concentrations of lead (partial correlations of 0.39 to 0.73 for China and 0.25 to 0.69 for South Korea), which originates from industrial processes [65–68], and that PM attributed to mineral dust is strongly associated with the level of airborne calcium, consistent with prior analysis [69–71]. We observe similarly consistent and distinct patterns for visibility, which both is affected by ambient PM particle characteristics [72–74] and affects avoidance behavior [75]. We also find that PM attributable to our two main anthropogenic sources, China and South Korea, is more strongly associated with smaller particle sizes, compared to PM originating from dust and sea salt, also consistent with prior research [76–79]. We hypothesize that North Korean PM may appear different from PM from China and South Korea because of its distinct industrial structure and energy system [80–84]. Taken together, these results indicate that our decomposition of PM by origin consistently identifies distinct sources of pollution that are normally indistinguishable.

Distinguishing health impacts by PM origin

Dose–response by PM origin We simultaneously estimate the effect of PM from all seven origins on health outcomes in each district over time, discovering that comparably sized exposures to PM of different origins cause significantly different changes in health. To demonstrate these differences, we first replicate the standard approach of pooling all PM, estimating an average health response that is undifferentiated by PM origin (Figure 3A). Ignoring origins, we find that a 1 $\mu g/m^3$ increase in overall PM is associated with a nearly linear \$0.002 ($\pm$0.0005) per day increase in respiratory medical cost per person (noting that the model allows for non-linear relationships; see SM section C). However, this undifferentiated model masks variation in health responses across origins. Figure 3B presents the response of health costs from exposure to PM from different origins. PM from North Korea (green) is the most harmful per unit at all common doses (\$0.0048 per $\mu g/m^3$ per person per day), with an average effect per unit at the mean dose around $4\times$ larger than the per-

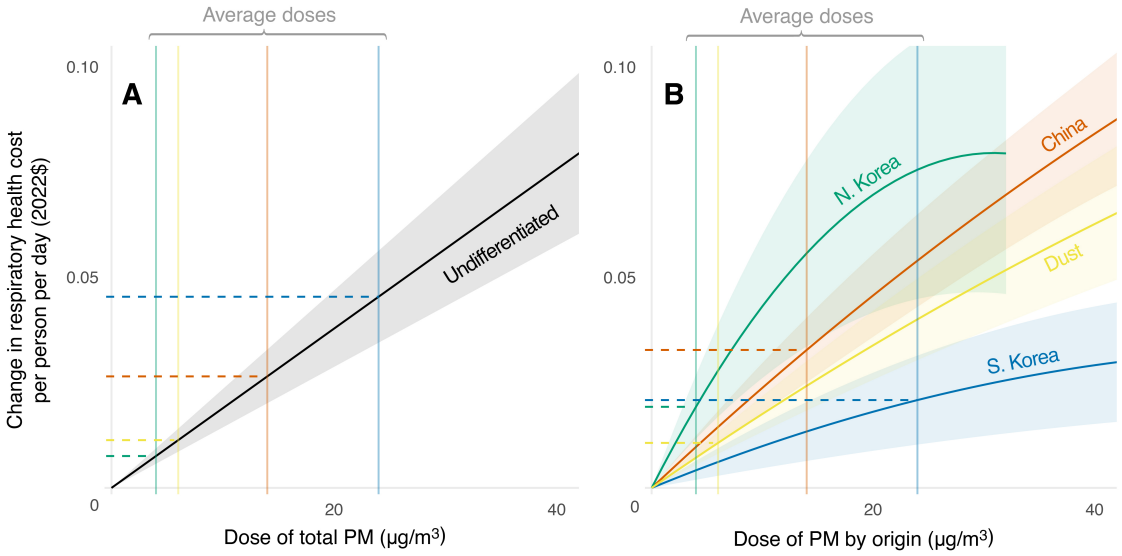


Figure 3: **Health-response to PM by origin.** (A) The cumulative respiratory health cost for the South Korean population from a single day of exposure to undifferentiated PM, aggregated across all origins. The relationship appears linear but is estimated allowing for a potentially nonlinear relationship. (B) Estimated effects of PM from four major origins: South Korea, North Korea, China, and dust (effects from seven sources are estimated simultaneously following the specification in SM section C). Each curve is plotted assuming the levels of other sources are zero. Panel A is shown only for comparison and is not used in our calculations for the attribution of respiratory health costs to PM origins. Vertical lines indicate the population-weighted mean dose for each PM origin across all district–days in the sample; horizontal lines indicate the predicted change in respiratory health costs per person. Dose–response relationships are shown only to the 99^{th} percentile dose for each PM origin if that value falls in the plotted range.

unit effect of the mean dose of PM from South Korea ("domestic-origin" PM, whose dose–response function is shown in blue; \$0.0009 per μg/m^3 per person per day). The average per-unit effects of PM from China (red-orange) and dust (yellow) are around 2.6× and 2× larger, respectively, than domestic-origin PM (\$0.0024 and \$0.0018 per μg/m^3 per person per day).

We note that the high harm per unit of PM originating in North Korea is not explained by the chemical and physical properties that we were able to analyze above (recall Figure 2). On dimensions we can observe, including particle size and lead concentrations, PM from North Korea appears less threatening to human health. This indicates that the qualities of North Korean PM that we do not observe are likely mediating this relationship, a topic that we believe merits further

study.

Total harm from average PM by origin Health costs (horizontal dashed lines in Figure 3A–B) result from the combination of dose–response relationships and dose from each origin (solid vertical lines). If the health burden from PM is computed using the undifferentiated dose-response, then the average dose of total PM (52 $\mu g/m^3$) would be estimated to generate average costs of \$0.10 ($\pm$0.026) per person per day (roughly \$5 million per day for the entire country), resulting from origin-specific harms that are proportional to the PM load attributed to each origin (Figure 3A). However, the estimated contribution of harm from each origin changes dramatically if origin-specific health responses are considered (Panel B). For example, we estimate that the small average dose of transboundary PM from North Korea (4.3 $\mu g/m^3$) is so harmful per unit that its impact is comparable to the impact of domestic-origin PM from South Korea, which has a mean dose almost six times higher (23.6 $\mu g/m^3$). Similarly, transboundary PM from China exhibits an intermediate mean dose (14 $\mu g/m^3$) that we estimate would generate health costs almost twice as large as domestic-origin PM. We estimate that average doses of PM from dust are small (6 $\mu g/m^3$), generating half the cost of domestic-origin PM, although Yellow Dust events, which lead to doses above 1000 $\mu g/m^3$, can generate substantial harm.

Temporal structure of health impacts by PM origin We find that cumulative respiratory medical expenditures emerge similarly across PM origins, rising gradually until leveling out after roughly three weeks, regardless of whether our model accounts for PM origin (Figure S3). The cumulative health response stabilizes after 21 days. Further, we find no evidence that indicates substantial temporal displacement of health costs ("harvesting," see SM section C).

Computing damages by PM origin in mixtures

Analyzing the health effects of PM by each origin in isolation provides a clear measure of relative impacts but is an incomplete picture of total PM impacts because PM is experienced as a mixture, and the presence of PM from one origin can affect the health impact of PM from other origins. In

particular, prior studies[85, 86] and our results suggest that individuals in South Korea engage in avoidance behavior (e.g., staying indoors) to protect themselves from the health effects of the entire PM mixture incident on their community. This implies that a unit of PM from origin j may have a health impact that depends on whether PM from origin k is high or low since higher PM from k may induce greater avoidance, mitigating the effect of PM from j (see SM section C). Stated another way, because individuals engaged in avoidance of PM from all origins, PM from each origin mediates the health impact of PM from all other origins. Figure 3 depicts only the *partial* effect of exposure to PM of each origin, holding PM from all origins at zero, thereby abstracting away from these interactions. However, computing the actual health impact of PM from any single origin requires accounting for the entire mixture of PM when exposure occurs.

We compute the total health impact of different PM mixtures, accounting for empirically estimated interactions between PM from different origins (see SM section C). Figure 4 presents a surface that describes the expected excess nationwide respiratory health spending that would result from exposing every person in South Korea to different mixtures of PM from South Korea, China, and North Korea for a single day (more complex combinations are possible to compute, but difficult to display). "Iso-damage" curves trace out mixtures of PM from the pairs of jurisdictions that generate the same health cost. Here, a slope of -1 would indicate a mixture where a one-unit increase in PM from either of the two originating jurisdictions would have the same incremental health impact (i.e., there would be no change in total spending for a one-to-one exchange of the two pollutants). However, we do not observe this for any observed mixtures (shading indicates historical frequency). Instead, iso-damage curves in both panels are steeper than -1, implying that incremental damages of anthropogenic PM from transboundary sources are always greater than those of domestic PM in our setting.

The mixture-damage surface in Figure 4 can be differentiated to compute the incremental harm caused by a unit of PM from a single origin that is contained within a mixture. Based on the average PM mixture from China and South Korea in our sample, assuming zero exposure to PM from other origins, we estimate that the incremental nationwide health costs of exposure to $+1$ $\mu g/m^3$ PM from China and South Korea—relative to historical levels–are \$108,808 ($\pm$\$10,896) and \$32,131

(±$7,943), respectively. For North Korea and South Korea, these values are $182,358 (±$45,842) and $26,665 (±$8,575), respectively. In terms of health costs, one additional µg/m^3 of PM from North Korea is equivalent to an increase of 6.84 µg/m^3 of PM from South Korea, and one µg/m^3 of PM from China is equivalent to 3.39 µg/m^3 of South Korean PM.

We also note that the curvature of this surface is concave, such that (i) overall harm increases with higher PM levels from any origin, but at a declining rate, and (ii) higher PM levels from each origin reduce the incremental harm from other origins (see SM section C).

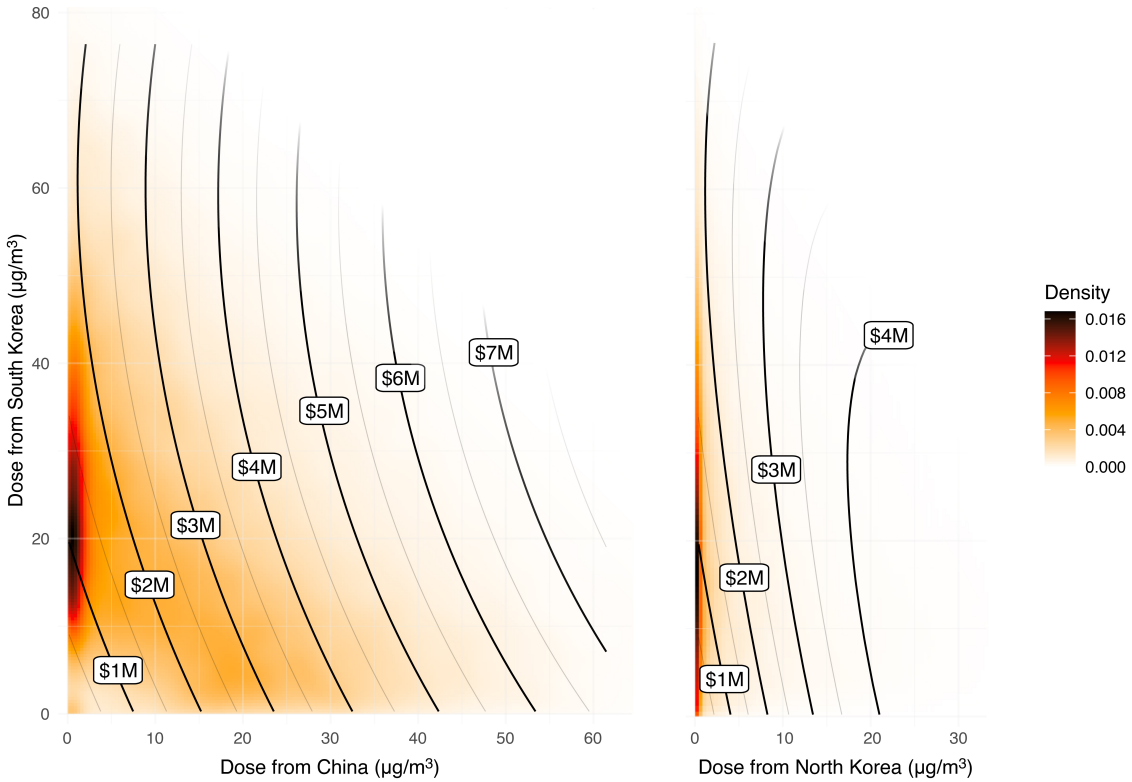


Figure 4: **Health costs from mixtures of domestic and transboundary PM.** Change in single-day total respiratory health spending (2022$) associated with nationwide exposure to mixtures of PM from combinations of domestic and transboundary PM from China (left) and North Korea (right). Estimates assume exposure to PM from all other origins are zero. Total spending is shown by black contour lines. Bivariate densities of exposure during our data period are shown as colored shading. Relationships are shown to the joint 99th-percentile of population-weighted historical exposures.

	Fraction of population exposure to anthropogenic PM in South Korea, % total	Estimated annual health cost, million 2022$ ± SE (% total, anthropogenic)	
		This study	Assuming undifferentiated dose–response
South Korea	56.6	305 ± 92 (27.6)	615 ± 70 (58.4)
North Korea	9.9	223 ± 66 (20.1)	97 ± 11 (9.2)
China	33.5	579 ± 66 (52.0)	341 ± 38 (32.3)

Table 1: Average annual contributions to anthropogenic PM exposure and associated change in respiratory health spending by origin.

Total damages from domestic and transboundary PM

We compute total nationwide damages traceable to each origin during our study period and find that accounting for differences in the per-unit harm by PM origin is critical for estimating the relative harms from domestic and transboundary PM (Table 1). We compute total harm from j by estimating the difference in health outcomes that would have been expected to occur in two different scenarios: one where PM emissions reflect actual historical emissions versus a scenario where j unilaterally reduces its emissions to zero (and the emissions of other countries are unchanged; see SM section C). Accounting for differences in harm per unit of PM, we estimate that PM originating in South Korea causes roughly $300 million in health costs in South Korea per year ($0.07 per person per day), while transboundary PM from China and North Korea generate roughly $580 and $220 million per year ($0.14 and $0.05 per person per day), respectively. (Note that health costs in South Korea reflect the low cost of medical care in a national healthcare system where costs for comparable treatments are roughly $6-10\times$ lower than in the United States [87].)

Prior state-of-the-art practice does not differentiate per-unit harms by origin [30, 33, 34, 36, 88]. Had we used the standard undifferentiated approach, we would have estimated that transboundary PM generated 41% of the costs from anthropogenic sources, rather than the 72% that we estimate here (Table 1). We estimate that 57% of anthropogenic PM that the South Korean population is exposed to originates domestically, but it generates only 28% of the damage from anthropogenic PM because it is relatively less harmful than transboundary PM.

Damages by origin over time We estimate that PM exposure and damages for major PM origins have exhibited different trends during our study period (see Figures 5A–C). We compute that exposure to PM from North Korea declined by roughly 12.5% per year, an effect that could be attributed to falling emissions, changes in meteorology, or other factors; over the same time, PM from South Korea remained essentially unchanged (+0.1% per year) and PM from China declined modestly (−1.6% per year). We estimate that the trends in estimated costs resulting from these exposures largely mirror these trends in exposure, although the overall baseline level of costs is relatively higher for PM from North Korea and China, reflecting the higher impact per unit of PM. Respiratory health costs attributable to PM from North Korea fell to about half of their 2005 level and costs related to PM from China declined slightly from 2013, reflecting emissions reductions associated with China's "war on pollution" [28, 89, 90]. In contrast, we estimate that health costs traceable to domestic PM emissions have increased steadily by 3.6% per year, an effect that can be explained by avoidance behavior: throughout our study period, domestic PM concentrations remained stable while transboundary PM concentrations declined, which on net increases the per-unit harms of domestic PM due to reductions in avoidance.

Damages by origin across space We find that the source of PM health damages varies substantially across space, with provinces nearer to emission sources generally experiencing relatively more damage from those sources (Figures 5D–F). In South Korea's southern and western provinces, we find that the largest portion of PM health damage results from emissions originating in China. PM from China accounts for as much as 80% of the costs in Jeju (the island province in the far southwest) and roughly 60% of the costs in other nearby provinces. PM from North Korea is responsible for relatively more damages in the north of South Korea—accounting for as much as 25% of the costs in the Seoul Metropolitan Area. Domestic emissions are responsible for around 28% of damages nationwide, but costs are relatively larger in South Korea's southeastern areas, rising to 40% of PM damage, where domestic heavy industry is concentrated.

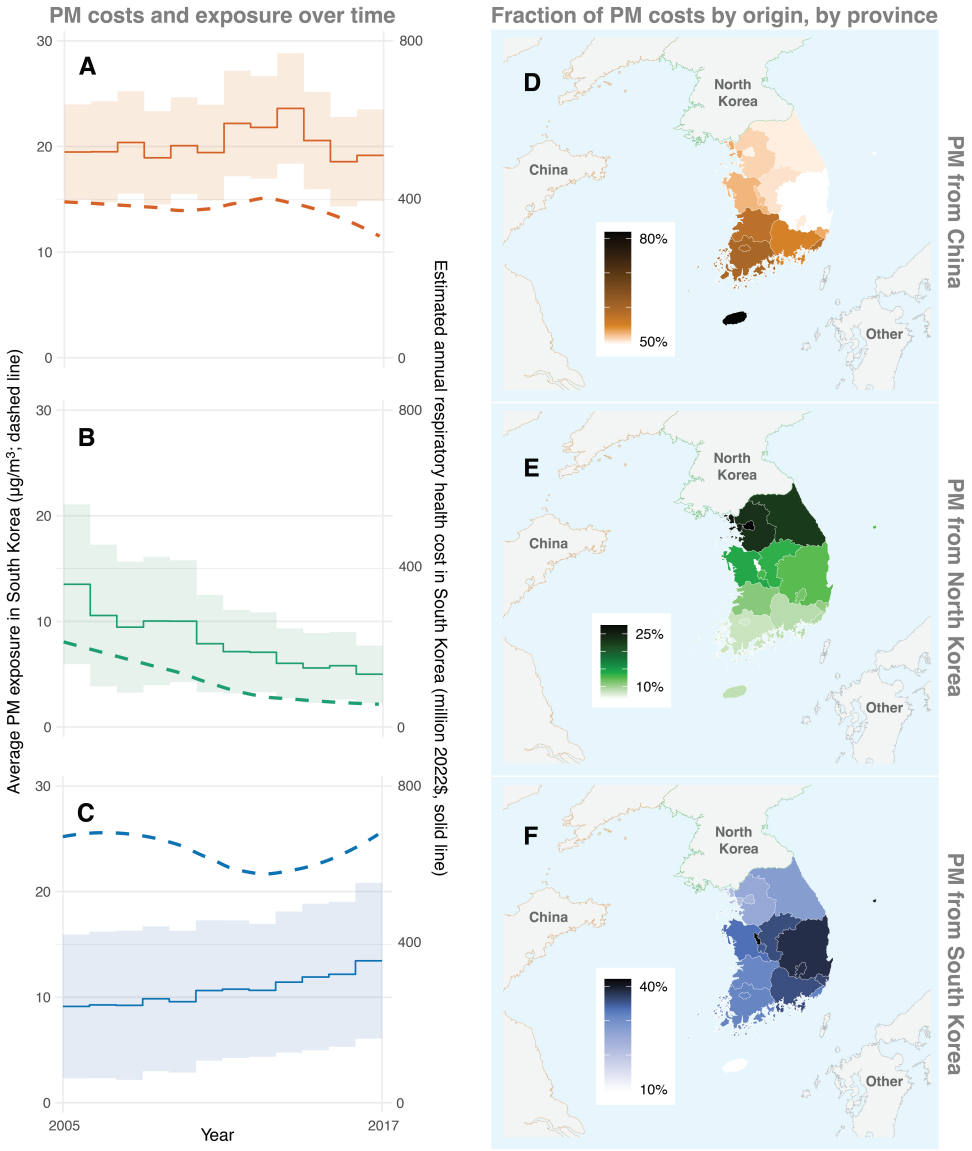


Figure 5: **Origin-specific damages from PM over time and across space.** (A) Trend of nationwide, population-weighted exposure in South Korea to PM from China (dashed lines, left y-axis) and annual respiratory health costs associated with PM exposures from China (solid lines, right y-axis, 95% confidence internal shown). (B)-(C) Same as (A), but for PM originating from (B) North Korea and (C) South Korea. (D) Province-level variation in the portion of respiratory health costs in South Korea attributable to anthropogenic PM from China. (E)-(F) same as (D), but for PM originating in (E) North Korea and (F) South Korea.

Discussion

Our findings show that distinguishing the health impacts of PM by its origin qualitatively changes our assessment of where health damage from PM originates. To the best of our knowledge, this study is the first to empirically demonstrate real-world differences in the health impacts of origin-specific contributions to a PM mixture incident on a single population. This is achieved by distinguishing pollutants from different origins and jointly estimating differences in their associated health impacts. Although we apply this approach to transboundary PM in Northeast Asia, it can be generalized to a broader class of pollution problems. For example, different water pollutants may similarly affect common metrics used to assess contamination, such as biochemical oxygen demand [91], but may have distinct per-unit impacts on a population or ecosystem; likewise, summary indices used for soil pollution [92] or ocean noise [93] may exhibit similar patterns. In settings like these, a key challenge for governance is identifying the linkages between multiple polluters and associated damages over large distances and long time horizons. The techniques we have developed here can support progress in these areas by providing critical information on the relative harms attributable to particular actors.

We have also demonstrated the quantitative implications of differentiating the harm caused by PM from different origins. Our findings suggest that when identifying sources of health damage, major biases can result from assuming equal per-unit harm, as is standard in the literature. For example, in our context, this assumption causes estimates of harm from multiple countries to be incorrect by a factor of two. We believe our results suggest that this assumption should be assessed in other contexts as well. In Southeast Asia, for example, transboundary haze is considered one of the region's more serious health concerns [94, 95]. In Sub-Saharan Africa, dust carried by winds from the Sahara is a major contributor to PM levels, though its effects may be distinct from the effects of PM from, for example, internal combustion engines [20, 21]. Even within a country, the diversity of regional pollutants can lead to divergent responses to subcomponents of PM. This issue may be especially relevant for larger countries, such as the U.S., China, Brazil, and India, where internal interstate air pollution flows are a growing concern [30, 33, 88, 96]. In all of these cases, the

approach we have developed here can be used to empirically test whether pollutants from different origins exhibit similar or distinct per-unit damages.

Two points are worth noting when interpreting our findings. First, data presented here should not be considered a complete cost-benefit analysis. The outcome of interest—health costs associated with outpatient visits due to respiratory illnesses—is an incomplete measure of total changes in social welfare. For example, defensive investments—such as the purchase of air purifiers and face masks—are often paid privately and are not accounted for in our analysis [97]. We also do not consider mortality costs, other morbidity costs, or any potential benefits associated with transboundary PM emissions, such as economic benefits from PM-emitting industrial activity in China [98]. We hope future work addresses these limitations.

Second, the magnitude of PM-driven health costs we measure will differ in other contexts. One reason for this is that our results are specific to South Korea: we report health effects of PM by origin, but those effects are net of any chemical or physical changes that occur to PM plumes after they are emitted. For example, because PM emitted in China must travel for many hours in the atmosphere before reaching South Korea, larger particles may have been preferentially removed or components of the emitted PM plume may have been oxidized into less harmful substances, both of which would affect resulting health costs. In other words, while we find that PM from China is more harmful per unit to the South Korean population than PM originating domestically, this does not necessarily imply that PM originating in China is more harmful *in general*. We also expect that, overall, monetized damages from PM will be larger in other high-income countries for two reasons. First, healthcare costs in South Korea are some of the lowest among developed countries. For instance, prior analyses have estimated that in 2020, the average fee for an initial primary care visit was about $13 (₩16,140) in South Korea while it was about $109 in the U.S. [99]. Similarly, the medical fees for hospital childbirth and appendectomy are $1,040 and $2,166 in South Korea, respectively, while the corresponding fees in the U.S. are $11,200 and $13,020 [87]. These differences in cost contribute to total annual health expenditure per capita of only $2,600 in South Korea, compared to $10,945 in the U.S. [100]. Thus, the costs we report here would almost certainly be higher if our study population were under an alternative healthcare system [85]. Second,

the South Korean population is very sensitive to information about PM levels, such as air quality warnings, and has already invested heavily in defensive adaptations, such as air filters [85, 86]. Populations that are less responsive to air quality information or that possess fewer defensive assets would likely experience larger health damages from any given quantity of PM exposure.

Lastly, our results suggest that institutions responsible for the management of transboundary pollution may need to consider origin-specific health responses to improve their assignment of damages and/or management of emissions. Since the *Trail Smelter* transboundary air pollution dispute was settled in 1941 [101], countries have attempted to control transboundary pollution within the frameworks of international institutions. The settlement served as the foundation for Principle 2 of the Rio Declaration, unanimously accepted by all 179 countries present at inaugural 1992 UN Earth Summit, which states "the polluter should, in principle, bear the cost of pollution" [102, 103]. Today, despite the existence of numerous commissions and organizations—including the United Nations Economic Commission for Europe (UNECE) Convention on Long-Range Transboundary Air Pollution, the Malé Declaration on Control and Prevention of Air Pollution and Its Likely Transboundary Effects for South Asia, and the Association of Southeast Asian Nations' Agreement on Transboundary Haze Pollution—none, to our knowledge, has suggested a method for accounting for origin-specific pollution health responses. Likewise, while the World Health Organization acknowledges the possibility that PM from different origins may have different health impacts, it currently does not offer guidelines to account for origin-specific health responses [104]. Our findings suggest that updated guidance and policies that account for these differences can improve public health.

References

1. Organization, W. H. *Review of evidence on health aspects of air pollution: REVIHAAP project: technical report* tech. rep. (World Health Organization. Regional Office for Europe, 2021).

2. Peters, R., Ee, N., Peters, J., *et al.* Air pollution and dementia: a systematic review. *Journal of Alzheimer's Disease* **70,** S145–S163 (2019).

3. Shehab, M. A. & Pope, F. D. Effects of short-term exposure to particulate matter air pollution on cognitive performance. *Scientific Reports* **9,** 1–10 (2019).

4. Braithwaite, I., Zhang, S., Kirkbride, J. B., Osborn, D. P. & Hayes, J. F. Air pollution (particulate matter) exposure and associations with depression, anxiety, bipolar, psychosis and suicide risk: a systematic review and meta-analysis. *Environmental Health Perspectives* **127,** 126002 (2019).

5. Burkhardt, J., Bayham, J., Wilson, A., *et al.* The effect of pollution on crime: Evidence from data on particulate matter and ozone. *Journal of Environmental Economics and Management* **98,** 102267 (2019).

6. Harrison, R. M. & Yin, J. Particulate matter in the atmosphere: which particle properties are important for its effects on health? *Science of the Total Environment* **249,** 85–101 (2000).

7. Kelly, F. J. & Fussell, J. C. Size, source and chemical composition as determinants of toxicity attributable to ambient particulate matter. *Atmospheric Environment* **60,** 504–526 (2012).

8. Strak, M., Janssen, N. A., Godri, K. J., *et al.* Respiratory health effects of airborne particulate matter: the role of particle size, composition, and oxidative potential—the RAPTES project. *Environmental Health Perspectives* **120,** 1183–1189 (2012).

9. Schmid, O. & Stoeger, T. Surface area is the biologically most effective dose metric for acute nanoparticle toxicity in the lung. *Journal of Aerosol Science* **99.** Inhaled Particle Dosimetry, 133–143. ISSN: 0021-8502 (2016).

10. Li, J., Chen, H., Li, X., *et al.* Differing toxicity of ambient particulate matter (PM) in global cities. *Atmospheric Environment* **212,** 305–315 (2019).

11. Park, M., Joo, H. S., Lee, K., *et al.* Differential toxicities of fine particulate matters from various sources. *Scientific Reports* **8,** 17007. ISSN: 2045-2322. `https://doi.org/10.1038/s41598-018-35398-0` (Nov. 2018).

12. Achilleos, S., Kioumourtzoglou, M.-A., Wu, C.-D., *et al.* Acute effects of fine particulate matter constituents on mortality: A systematic review and meta-regression analysis. *Environment International* **109,** 89–100 (2017).

13. Chen, J., Hoek, G., De Hoogh, K., *et al.* Long-Term Exposure to Source-Specific Fine Particles and Mortality— A Pooled Analysis of 14 European Cohorts within the ELAPSE Project. *Environmental Science and Technology* **56,** 9277–9290 (2022).

14. Seagrave, J., McDonald, J. D., Bedrick, E., *et al.* Lung toxicity of ambient particulate matter from southeastern U.S. sites with different contributing sources: relationships between composition and effects. *Environmental Health Perspectives* **114,** 1387–1393 (2006).

15. Strak, M., Steenhof, M., Godri, K. J., *et al.* Variation in characteristics of ambient particulate matter at eight locations in the Netherlands–The RAPTES project. *Atmospheric Environment* **45,** 4442–4453 (2011).

16. Wu, Y., Jin, T., He, W., *et al.* Associations of fine particulate matter and constituents with pediatric emergency room visits for respiratory diseases in Shanghai, China. *International Journal of Hygiene and Environmental Health* **236,** 113805 (2021).

17. Thurston, G. D., Chen, L. C. & Campen, M. Particle toxicity's role in air pollution. *Science* **375,** 506–506 (2022).

18. Aguilera, R., Corringham, T., Gershunov, A. & Benmarhnia, T. Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California. *Nature communications* **12,** 1493 (2021).

19. Xue, T., Geng, G., Han, Y., *et al.* Open fire exposure increases the risk of pregnancy loss in South Asia. *Nature Communications* **12,** 3205 (2021).

20. Heft-Neal, S., Burney, J., Bendavid, E. & Burke, M. Robust relationship between air quality and infant mortality in Africa. *Nature* **559,** 254–258 (2018).

21. Heft-Neal, S., Burney, J., Bendavid, E., Voss, K. K. & Burke, M. Dust pollution from the Sahara and African infant mortality. *Nature Sustainability* **3,** 863–871 (2020).

22. Gilmour, M. I., McGee, J., Duvall, R. M., *et al.* Comparative toxicity of size-fractionated airborne particulate matter obtained from different cities in the United States. *Inhalation Toxicology* **19,** 7–16 (2007).

23. Mendelsohn, R. Regulating heterogeneous emissions. en. *Journal of Environmental Economics and Management* **13,** 301–312. ISSN: 00950696. https://linkinghub.elsevier.com/retrieve/pii/009506968690001X (Dec. 1986).

24. Kolstad, C. D. Uniformity versus differentiation in regulating externalities. en. *Journal of Environmental Economics and Management* **14,** 386–399. ISSN: 00950696. https://linkinghub.elsevier.com/retrieve/pii/0095069687900295 (Dec. 1987).

25. Muller, N. Z. & Mendelsohn, R. Efficient Pollution Regulation: Getting the Prices Right. en. *American Economic Review* **99,** 1714–1739. ISSN: 0002-8282. https://pubs.aeaweb.org/doi/10.1257/aer.99.5.1714 (Dec. 2009).

26. Jacobsen, M. R., Knittel, C. R., Sallee, J. M. & Van Benthem, A. A. The Use of Regression Statistics to Analyze Imperfect Pricing Policies. en. *Journal of Political Economy* **128,** 1826–1876. ISSN: 0022-3808, 1537-534X. https://www.journals.uchicago.edu/doi/10.1086/705553 (May 2020).

27. Fowlie, M. & Muller, N. Market-Based Emissions Regulation When Damages Vary across Sources: What Are the Gains from Differentiation? en. *Journal of the Association of Environmental and Resource Economists* **6,** 593–632. ISSN: 2333-5955, 2333-5963. https://www.journals.uchicago.edu/doi/10.1086/702852 (May 2019).

28. Heo, S. W., Ito, K. & Kotamarthi, R. *International Spillover Effects of Air Pollution: Evidence from Mortality and Health Data* tech. rep. (National Bureau of Economic Research, 2023).

29. Chen, L., Lin, J., Martin, R., *et al.* Inequality in historical transboundary anthropogenic $PM_{2.5}$ health impacts. *Science Bulletin* **67,** 437–444 (2022).

30. Chen, S., Li, Y., Shi, G. & Zhu, Z. Gone with the wind? Emissions of neighboring coal-fired power plants and local public health in China. *China Economic Review* **69,** 101660 (2021).

31. Diao, B., Ding, L., Cheng, J. & Fang, X. Impact of transboundary $PM_{2.5}$ pollution on health risks and economic compensation in China. *Journal of Cleaner Production* **326,** 129312 (2021).

32. Han, S.-B., Song, S.-K., Shon, Z.-H., *et al.* Comprehensive study of a long-lasting severe haze in Seoul megacity and its impacts on fine particulate matter and health. *Chemosphere* **268,** 129369 (2021).

33. Dedoussi, I. C., Eastham, S. D., Monier, E. & Barrett, S. R. Premature mortality related to United States cross-state air pollution. *Nature* **578,** 261–265 (2020).

34. Lim, Y.-H., Kim, S., Han, C., *et al.* Source country-specific burden on health due to high concentrations of PM2.5. *Environmental Research* **182,** 109085 (2020).

35. Liu, S., Xing, J., Wang, S., *et al.* Revealing the impacts of transboundary pollution on $PM_{2.5}$-related deaths in China. *Environment International* **134,** 105323 (2020).

36. Sergi, B., Azevedo, I., Davis, S. J. & Muller, N. Z. Regional and county flows of particulate matter damage in the US. *Environmental Research Letters* **15,** 104073 (2020).

37. Gu, Y. & Yim, S. H. L. The air quality and health impacts of domestic trans-boundary pollution in various regions of China. *Environment International* **97,** 117–124 (2016).

38. Penn, S. L., Arunachalam, S., Woody, M., *et al.* Estimating State-Specific Contributions to PM $_{2.5}$ - and O $_3$ -Related Health Burden from Residential Combustion and Electricity Generating Unit Emissions in the United States. en. *Environmental Health Perspectives* **125,**

324–332. ISSN: 0091-6765, 1552-9924. https://ehp.niehs.nih.gov/doi/10.1289/EHP550 (Mar. 2017).

39. Fann, N., Fulcher, C. M. & Hubbell, B. J. The influence of location, source, and emission type in estimates of the human health benefits of reducing a ton of air pollution. en. *Air Quality, Atmosphere & Health* **2,** 169–176. ISSN: 1873-9318, 1873-9326. https://link.springer.com/10.1007/s11869-009-0044-0 (Sept. 2009).

40. Levy, J. I., Baxter, L. K. & Schwartz, J. Uncertainty and Variability in Health-Related Damages from Coal-Fired Power Plants in the United States. en. *Risk Analysis* **29,** 1000–1014. ISSN: 0272-4332, 1539-6924. https://onlinelibrary.wiley.com/doi/10.1111/j.1539-6924.2009.01227.x (July 2009).

41. Jia, R. & Ku, H. Is China's pollution the culprit for the choking of South Korea? Evidence from the Asian dust. *The Economic Journal* **129,** 3154–3188 (2019).

42. Jung, J., Choi, A. & Yoon, S. Transboundary air pollution and health: evidence from East Asia. *Environment and Development Economics* **27,** 120–144 (2022).

43. Crawford, J. H., Ahn, J.-Y., Al-Saadi, J., *et al.* The Korea–United States Air Quality (KORUS-AQ) field study. *Elementa: Science of the Anthropocene* **9,** 00163 (2021).

44. Jordan, C. E., Crawford, J. H., Beyersdorf, A. J., *et al.* Investigation of factors controlling $PM_{2.5}$ variability across the South Korean peninsula during KORUS-AQ. *Elementa: Science of the Anthropocene* **8** (2020).

45. Choi, J., Park, R. J., Lee, H.-M., *et al.* Impacts of local vs. trans-boundary emissions from different sectors on $PM_{2.5}$ exposure in South Korea during the KORUS-AQ campaign. *Atmospheric Environment* **203,** 196–205 (2019).

46. Inness, A., Ades, M., Agustí-Panareda, A., *et al. CAMS global reanalysis (EAC4)* https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-reanalysis-eac4?tab=overview (Copernicus Atmosphere Monitoring Service (CAMS) Atmosphere Data Store (ADS), 2019).

47. Draxler, R. R. & Hess, G. An overview of the HYSPLIT_4 modelling system for trajectories. *Australian Meteorological Magazine* **47,** 295–308 (1998).

48. Stunder, B. *Global Data Assimilation System (GDAS) 1° product* 2004. `http://ready.arl.%20noaa.gov/gdas1.php`.

49. European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL). *Emission Database for Global Atmospheric Research (EDGAR), release version 5.0* tech. rep. `http://edgar.jrc.ec.europa.eu` (European Commission, Joint Research Centre (JRC)/Netherlands Environmental Assessment Agency (PBL), 2019).

50. Van der Werf, G. R., Randerson, J. T., Giglio, L., *et al.* Global fire emissions estimates during 1997–2016. *Earth System Science Data* **9,** 697–720. `https://essd.copernicus.org/articles/9/697/2017/` (2017).

51. Randerson, J. T., Chen, Y., van der Werf, G. R., Rogers, B. M. & Morton, D. C. Global burned area and biomass burning emissions from small fires. *Journal of Geophysical Research: Biogeosciences* **117** (2012).

52. Giglio, L., Randerson, J. T. & van der Werf, G. R. Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4). *Journal of Geophysical Research: Biogeosciences* **118,** 317–328 (2013).

53. Carleton, T., Jina, A., Delgado, M., *et al.* Valuing the global mortality consequences of climate change accounting for adaptation costs and benefits. *The Quarterly Journal of Economics* **137,** 2037–2105 (2022).

54. Du, X., Jin, X., Zucker, N., Kennedy, R. & Urpelainen, J. Transboundary air pollution from coal-fired power generation. *Journal of Environmental Management* **270,** 110862 (2020).

55. Cheng, S., Wang, F., Li, J., *et al.* Application of trajectory clustering and source apportionment methods for investigating trans-boundary atmospheric PM_{10} pollution. *Aerosol and Air Quality Research* **13,** 333–342 (2013).

56. Sapkota, A., Symons, J. M., Kleissl, J., *et al.* Impact of the 2002 Canadian forest fires on particulate matter air quality in Baltimore City. *Environmental Science and Technology* **39,** 24–32 (2005).

57. Lee, S., Kim, M., Kim, S.-Y., *et al.* Assessment of long-range transboundary aerosols in Seoul, South Korea from Geostationary Ocean Color Imager (GOCI) and ground-based observations. *Environmental Pollution* **269,** 115924 (2021).

58. Escudero, M., Stein, A., Draxler, R., *et al.* Determination of the contribution of northern Africa dust source areas to PM_{10} concentrations over the Central Iberian Peninsula using the Hybrid Single-Particle Lagrangian Integrated Trajectory model (HYSPLIT) model. *Journal of Geophysical Research: Atmospheres* **111** (2006).

59. Alam, K., Qureshi, S. & Blaschke, T. Monitoring spatio-temporal aerosol patterns over Pakistan based on MODIS, TOMS and MISR satellite data and a HYSPLIT model. *Atmospheric Environment* **45,** 4641–4651 (2011).

60. Binkowski, F. S. & Roselle, S. J. Models-3 Community Multiscale Air Quality (CMAQ) model aerosol component 1. Model description. *Journal of Geophysical Research: Atmospheres* **108** (2003).

61. Wagstrom, K. M., Pandis, S. N., Yarwood, G., Wilson, G. M. & Morris, R. E. Development and application of a computationally efficient particulate matter apportionment algorithm in a three-dimensional chemical transport model. *Atmospheric Environment* **42,** 5650–5659 (2008).

62. Han, X., Cai, J., Zhang, M. & Wang, X. Numerical simulation of interannual variation in transboundary contributions from Chinese emissions to $PM_{2.5}$ mass burden in South Korea. *Atmospheric Environment* **256,** 118440 (2021).

63. Chen, D., Cheng, S., Liu, L., Chen, T. & Guo, X. An integrated MM5–CMAQ modeling approach for assessing trans-boundary PM_{10} contribution to the host city of 2008 Olympic summer games—Beijing, China. *Atmospheric Environment* **41,** 1237–1250 (2007).

64. Liu, J., Li, J. & Yao, F. Source-receptor relationship of transboundary particulate matter pollution between China, South Korea and Japan: Approaches, current understanding and limitations. *Critical Reviews in Environmental Science and Technology* **52,** 3896–3920 (2022).

65. Harrison, R. *Lead pollution: causes and control* (Springer Science & Business Media, 2012).

66. Del Rio-Salas, R., Ruiz, J., De la O-Villanueva, M., *et al.* Tracing geogenic and anthropogenic sources in urban dusts: Insights from lead isotopes. *Atmospheric Environment* **60,** 202–210 (2012).

67. Bellis, D., Cox, A., Staton, I., McLeod, C. & Satake, K. Mapping airborne lead contamination near a metals smelter in Derbyshire, UK: spatial variation of Pb concentration and 'enrichment factor' for tree bark. *Journal of Environmental Monitoring* **3,** 512–514 (2001).

68. Roy, D., Seo, Y.-C., Kim, S. & Oh, J. Human health risks assessment for airborne PM_{10}-bound metals in Seoul, Korea. *Environmental Science and Pollution Research* **26,** 24247–24261 (2019).

69. Heim, E., Dibb, J., Scheuer, E., *et al.* Asian dust observed during KORUS-AQ facilitates the uptake and incorporation of soluble pollutants during transport to South Korea. *Atmospheric Environment* **224,** 117305 (2020).

70. Kim, W.-H., Song, J.-M., Ko, H.-J., *et al.* Comparison of Chemical Compositions of Size-segregated Atmospheric Aerosols between Asian Dust and Non-Asian Dust Periods at Background Area of Korea. *Bulletin of the Korean Chemical Society* **33,** 3651–3656 (2012).

71. Park, S., Song, C., Kim, M., Kwon, S. & Lee, K. Study on size distribution of total aerosol and water-soluble ions during an Asian dust storm event at Jeju Island, Korea. *Environmental Monitoring and Assessment* **93,** 157–183 (2004).

72. Yuan, C.-S., Lee, C.-G., Liu, S.-H., *et al.* Correlation of atmospheric visibility with chemical composition of Kaohsiung aerosols. *Atmospheric Research* **82,** 663–679 (2006).

73. Kim, K. W. *et al.* Optical properties of size-resolved aerosol chemistry and visibility variation observed in the urban site of Seoul, Korea. *Aerosol and Air Quality Research* **15,** 271–283 (2015).

74. Kim, Y. J., Kim, K. W., Kim, S. D., Lee, B. K. & Han, J. S. Fine particulate matter characteristics and its impact on visibility impairment at two urban sites in Korea: Seoul and Incheon. *Atmospheric Environment* **40,** 593–605 (2006).

75. He, X., Luo, Z. & Zhang, J. The impact of air pollution on movie theater admissions. *Journal of Environmental Economics and Management* **112,** 102626 (2022).

76. Kar, A. & Takeuchi, K. Yellow dust: an overview of research and felt needs. *Journal of Arid Environments* **59,** 167–187 (2004).

77. Perrino, C., Canepari, S., Catrambone, M., *et al.* Influence of natural events on the concentration and composition of atmospheric particulate matter. *Atmospheric Environment* **43,** 4766–4779 (2009).

78. Pérez, N., Pey, J., Querol, X., *et al.* Partitioning of major and trace components in PM_{10}–$PM_{2.5}$–PM_1 at an urban site in Southern Europe. *Atmospheric environment* **42,** 1677–1691 (2008).

79. Gobbi, G., Barnaba, F. & Ammannato, L. Estimating the impact of Saharan dust on the year 2001 PM_{10} record of Rome, Italy. *Atmospheric Environment* **41,** 261–275 (2007).

80. Koen, V. & Beom, J. *North Korea: The last transition economy?* tech. rep. 1607 (OECD Economics Department, 2020). `https://www.oecd-ilibrary.org/content/paper/82dee315-en`.

81. Kim, I. S., Lee, J. Y. L. & Kim, Y. P. Energy Usage and Emissions of Air Pollutants in North Korea. *Journal of Korean Society for Atmospheric Environment (in Korean)* **27,** 303–312 (2011).

82. Lee, Y. W., Kim, Y. P. & Yeo, M. J. Estimation of air pollutant emissions from heavy industry sector in North Korea. *Particle and Aerosol Research (in Korean)* **17,** 133–148 (2021).

83. Kim, I. S. & Kim, Y. P. Characteristics of Energy Usage and Emissions of Air Pollutants in North Korea. *Journal of Korean Society for Atmospheric Environment (in Korean)* **35,** 125–137 (2019).

84. Kim, N. K., Kim, Y. P., Morino, Y., Kurokawa, J.-i. & Ohara, T. Verification of NOx emission inventories over North Korea. *Environmental Pollution* **195,** 236–244 (2014).

85. Anderson, M. L., Hyun, M. & Lee, J. *Bounds, Benefits, and Bad Air: Welfare Impacts of Pollution Alerts* tech. rep. (National Bureau of Economic Research, 2022).

86. Hahm, Y. & Yoon, H. The impact of air pollution alert services on respiratory diseases: generalized additive modeling study in South Korea. *Environmental Research Letters* **16,** 064048 (2021).

87. RIHP. *제20대 대통령 선거 보건의료 분야 정책제안서 [The 20th presidential election - policy proposal for health care]* tech. rep. (Research Institute for Healthcare Policy (RHIP), 2021).

88. Du, X., Guo, H., Zhang, H., Peng, W. & Urpelainen, J. Cross-state air pollution transport calls for more centralization in India's environmental federalism. *Atmospheric Pollution Research* **11,** 1797–1804 (2020).

89. Greenstone, M., He, G., Li, S. & Zou, E. Y. China's war on pollution: Evidence from the first 5 years. *Review of Environmental Economics and Policy* **15,** 281–299 (2021).

90. Zheng, S. & Kahn, M. E. A new era of pollution progress in urban China? *Journal of Economic Perspectives* **31,** 71–92 (2017).

91. Jouanneau, S., Recoules, L., Durand, M., *et al.* Methods for assessing biochemical oxygen demand (BOD): A review. *Water Research* **49,** 62–82 (2014).

92. Kowalska, J. B., Mazurek, R., Gasiorek, M. & Zaleski, T. Pollution indices as useful tools for the comprehensive evaluation of the degree of soil contamination–A review. *Environmental Geochemistry and Health* **40,** 2395–2420 (2018).

93. McCarthy, E. *International regulation of underwater sound: establishing rules and standards to address ocean noise pollution* ISBN: 978-1-4020-8077-7 (Kluwer, Boston, 2004).

94. Othman, J., Sahani, M., Mahmud, M. & Ahmad, M. K. S. Transboundary smoke haze pollution in Malaysia: Inpatient health impacts and economic valuation. *Environmental Pollution* **189,** 194–201 (2014).

95. Lee, J. S. H., Jaafar, Z., Tan, A. K. J., *et al.* Toward clearer skies: Challenges in regulating transboundary haze in Southeast Asia. *Environmental Science and Policy* **55,** 87–95 (2016).

96. DeCicca, P. & Malak, N. When good fences aren't enough: The impact of neighboring air pollution on infant health. *Journal of Environmental Economics and Management* **102,** 102324 (2020).

97. Deschenes, O., Greenstone, M. & Shapiro, J. S. Defensive investments and the demand for air quality: Evidence from the NO_x budget program. *American Economic Review* **107,** 2958–2989 (2017).

98. Zhang, Q., Jiang, X., Tong, D., *et al.* Transboundary health impacts of transported global air pollution and international trade. *Nature* **543,** 705–709 (2017).

99. RIHP. *주요국 의원급 의료기관 진찰료 [Medical examination fees in major countries]* tech. rep. (Research Institute for Healthcare Policy (RIHP), 2020).

100. Chow, R. D., Bradley, E. H. & Gross, C. P. *Comparison of Cancer-Related Spending and Mortality Rates in the U.S. vs 21 High-Income Countries* in *JAMA Health Forum* **3,5** (2022), e221229–e221229.

101. Henry, L. S., Kim, J. & Lee, D. From smelter fumes to Silk Road winds: Exploring legal responses to transboundary air pollution over South Korea. *Washington University Global Studies Law Review* **11,** 565 (2012).

102. Bratspies, R. M. & Miller, R. A. *Transboundary harm in international law: lessons from the Trail Smelter arbitration* (Cambridge University Press, 2006).

103. *Rio Declaration on Environment and Development* Report of the United National Conference on Environment and Development. United Nations, June 1992. `https://www.un.org/en/`

development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf (2020).

104. World Health Organization. *WHO global air quality guidelines: particulate matter ($PM_{2.5}$ and PM_{10}), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide* (World Health Organization, 2021).

105. NIER. *대기환경연보 2017 [Annual report of air quality in Korea 2017]* tech. rep. (National Institute of Environmental Research (NIER), 2018).

106. Rémy, S., Kipling, Z., Flemming, J., *et al.* Description and evaluation of the tropospheric aerosol scheme in the European Centre for Medium-Range Weather Forecasts (ECMWF) Integrated Forecasting System (IFS-AER, cycle 45R1). *Geoscientific Model Development* **12,** 4627–4659 (2019).

107. Ryu, Y.-H. & Min, S.-K. Long-term evaluation of atmospheric composition reanalyses from CAMS, TCR-2, and MERRA-2 over South Korea: Insights into applications, implications, and limitations. *Atmospheric Environment* **246,** 118062 (2021).

108. Korea Meteorological Administration (KMA). *황사업무 - 관측업무 [Yellow Dust - Observation]* https://www.kma.go.kr/kma/biz/asiandust01.jsp. Accessed: July 7, 2023.

109. Korea Meteorological Administration (KMA). *황사업무 - 예보업무 [Yellow Dust - Forecast]* https://www.kma.go.kr/kma/biz/asiandust03.jsp. Accessed: July 7, 2023.

110. Burke, M., Driscoll, A., Heft-Neal, S., *et al.* The changing risk and burden of wildfire in the United States. *Proceedings of the National Academy of Sciences* **118,** e2011048118 (2021).

111. Crippa, M., Solazzo, E., Huang, G., *et al.* High resolution temporal profiles in the Emissions Database for Global Atmospheric Research. *Scientific Data* **7,** 121. ISSN: 2052-4463. https://doi.org/10.1038/s41597-020-0462-2 (Apr. 2020).

112. Smith, W., Pan, L., Honomichl, S., *et al.* Diagnostics of Convective Transport over the Tropical Western Pacific from Trajectory Analyses. *Journal of Geophysical Research: Atmospheres* **126,** e2020JD034341 (2021).

113. Air Quality Modeling Group, U.S. Environmental Protection Agency. *User's guide for the AMS/EPA regulatory model (AERMOD)* (2004).

114. Lab, F. C. & for International Earth Science Information Network – CIESIN – Columbia University, C. *High Resolution Settlement Layer (HRSL)* Source imagery for HRSL © 2016 DigitalGlobe. 2016.

115. Neidell, M. Information, avoidance behavior, and health the effect of ozone on asthma hospitalizations. *Journal of Human Resources* **44,** 450–478 (2009).

116. Aguilar-Gomez, S., Dwyer, H., Graff Zivin, J. & Neidell, M. This Is Air: The "Nonhealth" Effects of Air Pollution. *Annual Review of Resource Economics* **14,** 403–425. eprint: `https://doi.org/10.1146/annurev-resource-111820-021816`. `https://doi.org/10.1146/annurev-resource-111820-021816` (2022).

117. Hsiang, S. Climate econometrics. *Annual Review of Resource Economics* **8,** 43–75 (2016).

118. Chetty, R. Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. *Annual Review of Economics* **1,** 451–488 (2009).

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

120. Ostro, B. D. Air pollution and morbidity revisited: a specification test. *Journal of Environmental Economics and Management* **14,** 87–98 (1987).

121. Schwartz, J. & Dockery, D. W. Particulate air pollution and daily mortality in Steubenville, Ohio. *American Journal of Epidemiology* **135,** 12–19 (1992).

122. Chay, K. Y. & Greenstone, M. The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession. *The Quarterly Journal of Economics* **118,** 1121–1167 (2003).

123. Ostro, B., Sanchez, J. M., Aranda, C. & Eskeland, G. S. Air pollution and mortality: results from a study of Santiago, Chile. *Journal of Exposure Analysis and Environmental Epidemiology* **6,** 97–114 (1996).

124. Greene, W. H. *Econometric Analysis* Upper Saddle River, NJ (Prentice Hall, 2003).

125. Angrist, J. D. & Pischke, J. *Mostly Harmless Econometrics: An Empiricist's Companion* (Princeton University Press, 2008).

126. Gasparrini, A., Guo, Y., Hashizume, M., *et al.* Mortality risk attributable to high and low ambient temperature: a multicountry observational study. *The Lancet* **386,** 369–375 (2015).

127. Newey, W. K. & West, K. D. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* **55,** 703 (May 1987).

128. Conley, T. G. GMM estimation with cross sectional dependence. *Journal of Econometrics* **92,** 1–45 (1999).

129. Hsiang, S. M. Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. *Proceedings of the National Academy of Sciences* **107,** 15367–15372 (2010).

Supplementary Materials

A Data Collection

Air pollutants

We obtained hourly data on ambient levels of PM_{10}, $PM_{2.5}$ (this is available only starting in 2015), ozone, sulfur dioxide, nitrogen dioxide, and carbon monoxide from the Air Korea portal (https://airkorea.or.kr/) administered by the Korea Environment Corporation (KECO), a South Korean government agency that manages air quality monitors (for a map of monitor locations and districts included in our sample see Figure S1). Holding the set of monitors constant, we estimated exposure for all missing monitor–hours over our sample period by calculating the empirical CDF for the monitor with missing data and then selecting the quantile of this CDF equal to the inverse squared distance-weighted quantile of observations for the five nearest non-missing monitors. Daily values for each monitor are the mean of hourly values.

Observations of airborne heavy metals

KECO also manages the observations of airborne heavy metals used in our analysis of PM chemistry. We requested daily observations of these measures using the official information disclosure process through the South Korean information request portal (https://open.go.kr/); use of the data was granted on the condition that it not be shared. KECO measures airborne heavy metals (lead and calcium) recorded by monitors of Air Pollution Monitoring Supersites every two hours. We aggregated those observations to daily average levels. We omitted readings from the Incheon supersite, located on Baekryeong Island, as it is far from any urban PM monitor.

Weather and atmospheric visibility

We collected hourly monitor-level readings for temperature, humidity, and precipitation using the Korea Meteorological Administration (KMA) portal. We then performed the same procedure for filling missing observations and assigning exposure to districts as detailed for pollutants above.

We also obtained measures of visibility (reported as visible distance) for the small set of reporting monitors in the KMA's automated synoptic observation system. We matched these hourly observations to those for the nearest monitors reporting levels of airborne heavy metals, applying a ceiling to all observations of 10 km to account for differences across stations in their censoring practices.

Health spending

Our health outcome of interest is the morbidity spending associated with outpatient and emergency visits due to respiratory disease. We were granted access to this data through a restricted-use agreement with the National Health Insurance Service (NHIS) data center, which manages the transaction information of South Korea's Social Health Care System. Our analysis uses a 10% (N $\approx$ 5 million) sample of insured individuals. The sampling was carried out by stratifying individuals by each pair of district and age group, the latter of which was based on five-year age groups (0 to 4, 5 to 9, and so on). Individuals who died before the end of our analysis period were excluded. We then filtered for records associated with respiratory disease (J-category illnesses according to the International Classification of Diseases, 10^{th} revision [ICD-10]) that occurred during the period 2005 to 2016. We then aggregated this individual expenditure information by date and district. To determine expenditure per capita, we used denominators reported by NHIS; to calculate total expenditure across South Korea, we linearly interpolate estimates of the district-level population between census years.

Some aspects of this data set are of note. First, health expenditures per capita include outpatient and emergency visits covering all the different levels of facilities in the South Korean healthcare system (public health centers, doctor's offices, clinics, hospitals, general hospitals, and tertiary general hospitals). We exclude costs incurred from inpatient visits, as they require several layers of referrals and tend to involve more complex sequelae.

Second, our data set incorporates private copayments plus costs covered by the healthcare system, allowing us to gauge the impact of air pollution on total health costs.

Third, we only include districts with 'urban' air pollution monitors during the analysis period.

KECO manages several types of air quality monitors, including "urban" and "rural" locations. We used urban monitors because they better represent the degree of air quality experienced by the population. Indeed, due to better representation, local government agencies use urban monitors to determine whether environmental standards are met [105]. We also exclude districts where air quality monitors were installed too recently (2016 or later). This filtering process left us with 147 districts out of the 229 districts in South Korea, while still covering the entirety of 17 provinces and over 91% of the population (see Figure S1).

Fourth, our data set excludes individuals who moved into or out of the districts we analyze during our data period. Moving can lead to discontinuous changes in district-level health expenditure, complicating interpretation of our results.

B Decomposition of PM by source

The PM observations used in this analysis come from a network of monitors located throughout South Korea. We decompose the PM values reported by these monitors in two steps. First, we extract the portion of PM associated with dust and sea salt using a global reanalysis product that estimates surface concentrations. Second, we combine a probabilistic estimate of pollution transport with high-resolution estimates of emission rates to apportion the remaining PM among two sources: human activities (which we further subdivide into four originating jurisdictions: China, South Korea, North Korea, and "other") and wildfire. Finally, we assign these exposures to districts (second-level administrative units) in South Korea.

Dust and sea salt

In the first step in our source decomposition, we partition out dust and sea salt from total observed PM. We collect 3-hourly rasters of surface-level PM components from the European Centre for Medium-Range Weather Forecasts (ECMWF) Atmospheric Composition Reanalysis 4, or EAC4 [46], and calculate the inverse distance weighted mean of the nine nearest raster values of each EAC4 surface-level PM component for each monitor–day. Following the weighting scheme used in the ECMWF Integrated Forecasting System (IFS-AER) [106], we determine the portion of total

PM that is attributable to mineral dust, which we rescale to match total PM values reported by each station.

In addition, we adjust for "Yellow Dust" days, as EAC4 is known to systematically underestimate this particular source of PM [107]. Though these events are infrequent, they can produce daily average PM values in South Korea above 1000 $\mu g/m^3$ as large quantities of coarse dust are removed by aeolian erosion from the soil and rocks of the Mongolian plateau and are carried by strong winds across the Yellow Sea.[1] This adjustment is made only on days and in provinces when the Korea Meteorological Agency reported a Yellow Dust event, a determination made by visual inspection by trained meteorologists in conjunction with the output of a physical model of dust erosion [108, 109]. For these observations, if we estimate that PM that is not EAC4 dust is above its 90^{th} percentile for that monitor's weekday–month, all such PM above the monitor's weekday–month median value is attributed to Yellow Dust. Our main analysis treats EAC4 dust and this additional Yellow Dust component jointly. In addition, to account for misreporting, we also apply this adjustment to neighboring provinces or the same province on an adjacent day if its reported values are above its weekday-by-month 99^{th} percentile. This adjustment affects 2.3% of province–days. We note that this procedure follows from a similar approach developed to account for extreme $PM_{2.5}$ from wildfires in North America [110].

We perform a similar procedure for sea salt by determining the portion of non-dust PM that is attributable to sea salt from EAC4 and scale that to match the non-dust PM value for each station.

PM from human activity and wildfire

The second step of our decomposition determines the contributions of human activities and wildfire to the portion of observed PM that remains after removing dust and sea salt. We first generate a large number of air parcel backtrajectories using a physical model of atmospheric dynamics, the Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) model [47], and the GDAS1 global atmosphere reanalysis product [48]. These backtrajectories represent the paths taken by a hypothetical parcel of air traveling backwards in time from its observed destination, based on

[1] Importantly, the meteorological conditions on these days are unusual and may also carry higher- or lower-than-usual quantities of PM from other sources, though we have no way of directly testing this.

the timing and distribution of observed wind flows (e.g., Figure 1A). Note that wind flow in the atmosphere is well observed and tightly constrained in this mid-latitude region by its relationship to air temperature and pressure, both of which are also well observed.

We initialize backtrajectories at eight times of day (every 3 hours beginning at 2 a.m. local time) and eight heights (2, 4, 8, 16, 32, 64, 128, 256, and 512 meters) at each PM monitor on each day and track each parcel for 240 hours back in time (16.3 billion locations along 67.9 million paths). We define trajectories that pass twice the height of the planetary boundary layer as exiting the model and provide small vertical perturbations to particles that collide with the surface to keep them aloft. For each hour of each trajectory that arrives at a monitor on a specific day, this results in 16 parcel locations where emissions are potentially entrained in the air parcel before it arrives at its destination in South Korea. We rasterize these locations into a $0.1^\circ \times 0.1^\circ$ grid by counting the number of parcel instances that occur within each grid cell and smooth these estimated counts across space using a Gaussian kernel ($\sigma = 1$ pixel) to account for uncertainty in trajectories. Finally, we normalize the sum of each raster to one, creating a map that we interpret as the probability (denoted π) that each grid cell contributes material to that specific air parcel in that hour of its trajectory.

To estimate the quantity of emissions that are entrained in the air parcel from each location during a specific hour, we multiply these contribution-probability rasters by sector-specific PM emissions (denoted q) from the Emissions Database for Global Atmospheric Research (EDGAR) 5.0 [49]. We rescale EDGAR values to account for hourly, weekly, seasonal, and secular variation in emissions intensity using a set of scalars for high-resolution temporal profiles [111]. We follow the same procedure for 3-hourly Global Fire Emissions Database (GFED) 4.1s [50–52], where we use emissions rates based on estimated fuel type for burn fire location. HYSPLIT does not contain active chemistry, so we apply the simple approximation that non-methane volatile organic compounds (VOCs), sulfur dioxide, nitrogen oxides, and ammonia emissions estimated in EDGAR 5.0 are converted to an equivalent mass of PM.[2]

[2]Somewhat surprisingly, we find that our results are not sensitive to this assumption; this may be because the correlations between primary PM and NO_x, SO_2, NMVOC, and NH_3 emissions are very high. Assuming zero conversion into secondary aerosols changes our results by only a small margin. That said, seasonal variation in the conversion efficiency or rate may affect our results, but the direction of this adjustment is not known.

We model deposition, scavenging, and chemistry in a simplified framework that assumes net exponential decay. We empirically calibrate a separate decay rate (denoted r) for each quarter of the year—to account for seasonal differences in atmospheric moisture, temperature, and cloud physics—that maximizes the correlation between the observed values of non-dust and non-salt PM (in South Korea) and the values predicted by our trajectory model (exponential decay rates of 0.004, 0.008, 0.016, and 0.006, respectively, with maximum Pearson correlations for all seasons around 0.6). We note that the assumption of a constant decay rate/particle lifetime is a simplification made elsewhere, such as in a trajectory analysis over the western Pacific [112], and the decay rates we recover (implying a half-life of particles around 24 hours) are similar to those suggested elsewhere in the literature, such as in the AMS/EPA Regulatory Model [113]. We then apply these rates to each parcel throughout its trajectory.

We separately track PM for each non-dust, non-salt origin j, scaling the predicted relative contributions of all origins to match the residual level of PM that arrives, and is observed, at monitor l at time t. Our estimate for the relative quantity of PM from origin j arriving at monitor l at time t is

$$\hat{p}_{ljt} = \sum_{\tau=0}^{T} \sum_{v \in j} \underbrace{\pi_{lv(t-\tau)}}_{\substack{\text{Probability that} \\ \text{air mass from } v \text{ at } t-\tau \\ \text{arrives at } l \text{ at } t}} \cdot \underbrace{s_{v(t-\tau)}}_{\substack{\text{Emissions} \\ \text{from } v \text{ at } t-\tau}} \cdot \underbrace{(1-r)^{\tau}}_{\text{Decay factor}} \tag{1}$$

where $\pi_{lv(t-\tau)}$ is the probability, estimated from HYSPLIT, that an air mass has arrived at location l at time t from pixel v in origin j at time $t-\tau$, $s_{v(t-\tau)}$ is a measure of actual emissions at location v and time $t-\tau$ (from EDGAR and GFED), and $(1-r)^{\tau}$ is the fraction of emissions that remain at arrival following deposition/decay/scavenging. To compute $\hat{p}_{ljt}$, we sum what remains of the emissions $s_{v(t-\tau)}$ over all locations v in origin j (e.g., over all grid cells in China) and over all emission times (i.e., for emissions that have traveled to our exposed population for 0 hours [$\tau = 0$], 1 hour [$\tau = 1$], 2 hours [$\tau = 2$], and so on).

Partitioning observed PM

The estimate $\hat{p}_{ljt}$ represents only the *relative* contribution of origin j to non-dust and non-salt PM observed at monitor l at time t. To apportion this PM to an origin, we begin with $\ddot{p}_{lt}$, the PM observed at l at time t that remains after accounting for dust and sea salt using the method described above in "Dust and sea salt" (i.e., $\ddot{p}_{lt} \equiv p_{lt} - p_{lt,\text{dust}} - p_{lt,\text{sea salt}}$). This residual is apportioned proportionately as

$$p_{ljt} = \ddot{p}_{lt} \cdot \frac{\hat{p}_{ljt}}{\sum_j \hat{p}_{ljt}} \tag{2}$$

where p_{ljt} is the amount of observed PM at monitor l and time t from origin j, where j includes South Korea, North Korea, China, wildfires, and "other sources." The procedure above, in "Dust and sea salt," provides a similar value for $j \in \{\text{dust, sea salt}\}$. Figures 1D–E illustrate this partition. We then aggregate these seven estimates by district: for each district i, we interpolate based on values of p_{ljt}, calculating an inverse squared distance-weighted average to each district i's center of population (calculated using Meta's High Resolution Population Density Maps [114]). This procedure yields 600,119 observations, p_{ijt}, for all seven origins j.

C Econometric analysis

The main text contains two separate econometric analyses. The first analysis validates our decomposition of PM by origin by evaluating how the contribution of PM from different origins affects the chemical and physical properties of PM observed at a small number of South Korean monitors that collect these data. The second analysis estimates the impact of PM from each origin on health costs in districts across South Korea.

Decomposing properties of PM mixtures by origin

We estimate the extent to which PM from each origin is associated with differences in the chemical and physical characteristics of the PM mixture that arrives at each destination. To do this, we empirically decompose the chemical and physical properties of PM (concentrations of lead and

calcium, $PM_{2.5}$, and visibility) observed at a limited number of advanced monitors into contributions from each origin, based on the amount of PM that we attribute to origin j on each day t in districts containing these advanced monitors. The chemical or physical property c_{it} (e.g., lead concentration) observed in district i (treating the district's advanced monitor and the district as synonymous) on day t is an additive mixture of contributions from different origins

$$c_{it} = c_{i,\text{South Korea},t} + c_{i,\text{China},t} + \ldots + c_{i,\text{wildfire},t} + \epsilon_{it} \quad (3)$$

where c_{ijt} is the contribution from j that arrives at monitor i on day t and ϵ_{it} is unmodeled error. For these chemical or physical proprieties, the contribution from location j is the product of the total PM from origin j (p_{ijt}) and the change in c per additional unit of origin j PM (λ_{ij}; e.g., the change in ambient lead observed at location i for a given increase in PM from origin j). Specifically:

$$c_{it} = p_{i,\text{South Korea},t} \cdot \lambda_{i,\text{South Korea}} + p_{i,\text{China},t} \cdot \lambda_{i,\text{China}} + \ldots + p_{i,\text{wildfire},t} \cdot \lambda_{i,\text{wildfire}} + \epsilon_{it} \quad (4)$$

which we solve via multiple linear regression. Here, we allow λ to differ by location i precisely so we are able to test its consistency over space, but we assume it is stable over our sample period. To aid comparison, we present in Figure 2 the set of λ_{ij}'s as partial correlation coefficients. The specific construction of this figure is illustrated in Figure S2: (1) Multiple correlation coefficients for all origins j but only a single destination site i (and a pair of chemical/physical variables) are plotted (Figure S2A). (2) We repeat this process for all destinations i. Then, coefficients are regrouped based on their origin, pooling across destinations to define a convex hull for each origin (Figure S2B). This allows us to assess patterns in the chemical/physical features of air that arrives at a set of destinations from each origin (Figure S2C).

Estimating the health response to undifferentiated total PM

Our objective is to develop a statistical framework that can be used to directly measure the relationship between exposure to PM from a specific origin and changes in health costs. In prior analyses, researchers have focused on modeling the health effects of total PM without distinguishing between contributing origins. Thus, we first analyze the health response to undifferentiated PM as in the previous literature. Then, we extend this framework to include the impacts of PM from different origins.

Conceptually, our model for respiratory health costs, h_{it}, from undifferentiated PM occurring in district i on day t is:

$$h_{it} = \underbrace{\bar{D}(\mathbf{p}_{it},\, \bar{A}(\mathbf{p}_{it}))}_{\text{damage from PM}} + \underbrace{\bar{f}(\mathbf{Q}_{it},\, \mathbf{W}_{it},\, \bar{\boldsymbol{\zeta}}_{it},\, \bar{\theta}_t)}_{\text{other causes}} + \bar{e}_{it} \tag{5}$$

where $\bar{D}(\cdot)$ is the damage from PM ($\mathbf{p}$), accounting for avoidance ($\bar{A}$). Other patterns in health costs that can be netted out by covariates are captured in $\bar{f}(\cdot)$, including the effects of other pollutants ($\mathbf{Q}$), meteorological conditions ($\mathbf{W}$), and a rich set of non-parametric fixed effects ($\bar{\boldsymbol{\zeta}}$) and trends ($\bar{\theta}$). We discuss each of these model elements below. Unmodeled variation in health spending is captured by $\bar{e}$. We use a bar above estimated model elements to denote that they describe components of the model for *undifferentiated PM*, in contrast to an alternative model that differentiates contributions by origin, in which bars are omitted (presented below). We discuss each of the arguments of Equation 5 in turn and then describe how we estimate the model.

Total particulate matter The term $\mathbf{p}_{it}$ describes the levels of total PM daily from the time $t - K$ to t, where K is the number of daily lags over which impacts accrue. We set $K = 28$ days to ensure potential delayed health effects are accounted for. As discussed in the main text, delays sometimes occur because lower-level referrals are required to visit a specialty hospital in the South Korean healthcare system. Thus, $\mathbf{p}_{it} = (p_{it},\, \cdots,\, p_{i(t-28)})$.

Accounting for avoidance behavior $\bar{A}$ is a summary measure of avoidance behavior. As ambient PM increases, individuals may recognize higher levels of air pollution based on changes in atmospheric visibility or information provided by governments and take defensive action (e.g., wearing face masks) or avoid certain activities (e.g., by staying at home). These actions may then in turn reduce the average damages incurred per additional unit of PM [53, 85, 97, 115, 116]; thus, it is usually thought that $\frac{\partial \bar{A}}{\partial p} > 0$ and $\frac{\partial \bar{h}}{\partial \bar{A}} < 0$. We thus design a function for the net damages caused by PM ($\bar{D}$) capable of incorporating this potential mechanism. Conceptually, we assume health damages are generated by a process:

$$\bar{D}(\mathbf{p}_{it},\, \bar{A}(\mathbf{p}_{it})) = \sum_{k=0}^{28} \underbrace{\bar{g}_k(p_{i(t-k)})}_{\text{biological response}} \cdot \underbrace{(1 - \bar{A}_{t-k}(p_{i(t-k)}))}_{\text{attenuation from avoidance}} \tag{6}$$

where we have explicitly expanded the number of terms in the summation to account for the 28 lagged days of pollution exposure that might contribute to harm on day t. Here, $\bar{g}_k$ is the biological impact of $p_{i(t-k)}$ on the health response in the absence of avoidance, where k indexes the damage from sequential lags. $\bar{A}_{t-k}$ measures the level of avoidance triggered by the PM level at time $t-k$, and can take on a value between zero (no avoidance) and one (total avoidance), inclusive. Thus, the effects of $\bar{g}_k$ can be partially mitigated by avoidance behavior, captured by the multiplicative term $(1-\bar{A}_{t-k})$ which drives damages to zero as $\bar{A}_{t-k} \to 1$. This expression for $\bar{D}$ describes, conceptually, how the data are assumed to be generated, however we do not observe avoidance behavior directly and thus cannot rely on explicit measurements of avoidance in our estimation. To overcome this challenge, we draw on techniques developed in the study of climate change damages [53, 117] and note that estimation can be simplified because avoidance behavior is an explicit response to the level of pollution; that is, $\bar{A}_{t-k}$ can be written as a function of $p_{i(t-k)}$ (i.e., pollution levels are a "sufficient statistic" for avoidance [118]). Thus, under the simplifying assumption that both $\bar{g}_k$ and $\bar{A}_k$ are approximately linear in $p_{i(t-k)}$, the expression for damages in Eq. 6 reduces to a quadratic function with two parameters for each lag:

$$\bar{D}(\mathbf{p}_{it},\, \bar{A}(\mathbf{p}_{it})) = \sum_{k=0}^{28} \bar{\alpha}_k \cdot p_{i(t-k)} + \bar{\beta}_k \cdot p_{i(t-k)}^2 \tag{7}$$

which can be directly estimated from data. Consequently, the shape of our response function is non-linear in pollution, consistent with prior analyses [119], and this non-linearity embeds information about the degree of unobserved avoidance undertaken by individuals.

Covariates We also account for a matrix of potential confounding variables related to our health outcomes of interest and pollutant levels, which are included as covariates in our regression analysis. We account for the levels of other air pollutants (nitrogen dioxide, carbon monoxide, ozone, and sulfur dioxide), denoted by the matrix $\mathbf{Q}_{it} = \{\mathbf{q}_{nit}\} = \{\mathbf{q}_{1it}, \mathbf{q}_{2it}, \mathbf{q}_{3it}, \mathbf{q}_{4it}\}$, and meteorological conditions (temperature, humidity, and precipitation), denoted as $\mathbf{W}_{it} == \{\mathbf{q}_{mit}\} = \{\mathbf{w}_{1it}, \mathbf{w}_{2it}, \mathbf{w}_{3it}\}$. Similar to our notation describing PM, these matrices include the contemporaneous and lagged values of each variable, capturing effects within the 28-day window for exposure that we are studying.

Non-parametric controls and trends To account for a wide variety of time-dependent and time-invariant unobserved determinants of both pollution levels and health spending [120–123], we include district-by-year-by-month-by-weekday ($\bar{\boldsymbol{\zeta}}_{it}$) and holiday ($\bar{\theta}_t$) intercepts (also referred to as "fixed effects" [124]), which absorb these groups' mean values. Our identification of the effect of PM on health relies on the ability of this adjustment to account for potentially confounding variations in both pollution and health outcomes that are correlated with the day of the week (e.g., weekdays and weekends), season (e.g., flu season from December to February), year (e.g., the outbreak of swine flu pandemic in 2009), or district (e.g., differences in industrial infrastructure, health system, demographics, and socioeconomic conditions). With this approach, we assess the relationship between PM and health outcomes by comparing only observations within the same weekday, month, year, and district group; in other words, our estimates describe how health outcomes change as a function of pollution by comparing outcomes on the set of, for example, non-holiday Mondays within March 2012 within a single district of South Korea [125].

Estimation Combining Eqs. 5 and 7 with model components described above, we empirically estimate how health costs respond to undifferentiated PM by solving the panel regression:

$$h_{it} = \underbrace{\sum_{k=0}^{28} \left[\bar{\alpha}_k \cdot p_{i(t-k)} + \bar{\beta}_k \cdot p_{i(t-k)}^2 \right]}_{\bar{D}(\boldsymbol{p})} + \underbrace{\sum_{n=1}^{4} \bar{\boldsymbol{\gamma}}_n(\mathbf{q}_{nit}) + \sum_{m=1}^{3} \bar{\boldsymbol{\delta}}_m(\mathbf{w}_{mit}) + \bar{\boldsymbol{\zeta}}_{it} + \bar{\theta}_t}_{\bar{f}(\cdot)} + \bar{e}_{it} \tag{8}$$

where h_{it} is a health outcome of interest in district i in time t and $p_{i,t-k}$ is the level of the total undifferentiated PM in district i at time $t-k$. Other model variables are as described above. $\bar{\boldsymbol{\gamma}}_n$ and $\bar{\boldsymbol{\delta}}_m$ are flexible nonlinear functions of pollution and meteorological covariates, respectively, that account for the potentially nonlinear impact of these variables on health. For temperature, we use a cross-basis function that models the dose–response function as a three-knot natural cubic spline with knots placed at the 10^{th}, 50^{th}, and 90^{th} percentiles of the temperature distribution and with the lag–response function modeled as a piecewise zero-order spline with knots at lags 1, 2, 4, 7, and 14. For humidity and all non-PM pollutants, we adopt the same lag–response specification but choose an equally-spaced two-knot natural cubic spline for the dose–response function. For precipitation, we also adopt the same lag–response specification but adopt a zero-order spline dose–response function with knots spaced in logs throughout the range of historically observed of daily total precipitation (3.13, 11, 23, 38, 56, 78.5, 101.5, and 129 mm). This structure for the dose–lag–response specifications we adopt for our covariates are motivated by prior findings in the literature [53, 126] and efficient computation for a high-dimensional model. We also include a rich set of district-by-year-by-month-by-weekday fixed effects described above and indicated by $\bar{\boldsymbol{\zeta}}$; $\bar{\theta}$ is a holiday fixed effect. The error term, $\bar{e}_{it}$, accounts for variation in the outcome not explained by PM and the other covariates. We cluster standard errors at the province level, which accounts for arbitrary forms of temporal auto-correlation in health outcomes within a province between all days in our sample, as well as arbitrary spatial auto-correlation across all districts within a province [127–129].

Computing damages by origin assuming undifferentiated PM To compute partial damages for PM from specific origins using the undifferentiated PM model, we estimate Eq. 8 and isolate the estimated terms in the damage component $\bar{D}(\cdot)$. We then compute the damages that

would be experienced if the entire South Korean population were exposed to a constant specific PM level $\bar{\rho}$. We thus compute damages $\bar{D}(\boldsymbol{p} = \bar{\boldsymbol{\rho}})$ where the vector $\bar{\boldsymbol{\rho}}$ is defined such that all 29 elements (representing lags) are set to the same value—i.e., $\bar{\boldsymbol{\rho}} := (\bar{\rho}, ..., \bar{\rho})$. Fig. 3A illustrates the function $\bar{D}(\bar{\boldsymbol{\rho}})$ evaluated across different values of $\bar{\rho}$. We then compute the damages that would be attributed to a specific origin j by evaluating $\bar{D}(\cdot)$ at $\boldsymbol{p} = \bar{\boldsymbol{\rho}}_j$, the average level of PM originating from j that is incident on the South Korean population.

Estimating the health response to PM from different origins

Simultaneously estimating damages for PM from multiple origins (both transboundary and domestic) is the central contribution of this analysis. To do this, we extend the framework and approach described in Eqs. 5-8 to differentiate impacts of PM from different origins, rather than estimating a undifferentiated pooled effect. Conceptually, when accounting for multiple PM sources, we continue to consider health impact that originate from PM damage and other factors:

$$h_{it} = D(\mathbf{P}_{it},\, A(\mathbf{P}_{it})) + f(\mathbf{Q}_{it},\, \mathbf{W}_{it},\, \boldsymbol{\zeta}_{it},\, \theta_t) + e_{it} \tag{9}$$

where terms correspond to analogs in Eq. 5 but with the bar notation removed (to indicated that these are no longer from the undifferentiated model) and with one key substantive difference. In Eq. 9, PM is represented by the matrix $\mathbf{P}_{it}$, which contains PM values from different origins observed in location i and time from $t - 28$ to t, which contrasts with the vector of undifferentiated PM in Eq. 5. $\mathbf{P}_{it}$ consists of seven vectors describing PM from each origin: South Korea, China, North Korea, dust, wildfire, sea salt, and other sources. Each of the seven vectors is composed of the PM values from their respective origin for the period from $t - 28$ to t; hence, $\mathbf{P}_{it}$ is composed of 203 elements.

Accounting for avoidance with multiple PM sources Accounting for avoidance in a model with multiple sources of PM is more complex than in the undifferentiated PM model, since individuals may exhibit a different avoidance response to each type of PM and each type of avoidance can alter the health response to PM from each origin. Individuals likely respond differently to PM

from different origins because they have different chemical and physical properties (e.g., they may smell different or alter visibility in a distinct way). However, it is likely that the total level of PM that is incident on a population, from all sources, also affects their avoidance behavior.

Thus, we adjust the approach in Eq. 6 to allow for these possible different avoidance responses, one for PM from each origin and for total PM. Further, each type of avoidance can alter the health response for PM from each origin. Replacing the terms that include the undifferentiated PM with those that include the PM levels from heterogeneous origins and adding avoidance terms yields:

$$D(\mathbf{P}_{it},\ A(\mathbf{P}_{it})) = \sum_{k=0}^{28} \underbrace{\left(\sum_{j=1}^{7} g_{jk}(p_{ij(t-k)}) \right)}_{\text{biological response}} \cdot \left(1 - \underbrace{A_{0k}(p_{i(t-k)})}_{\text{total PM avoidance}} - \underbrace{\sum_{j=1}^{7} A_{jk}(p_{ij(t-k)})}_{\text{origin-specific avoidance}} \right) \tag{10}$$

In this differentiated framework, there are seven g_{jk} for each lag k, corresponding to the biological responses from the seven origins indexed by j. The second set of parentheses contains multiple avoidance terms. A_0 is similar to $\bar{A}$ in Eq. 6 in the sense that it captures the degree of avoidance behavior triggered by the total level of the undifferentiated PM. This is important because government agencies provide information on the total level of PM ($p_{i(t-k)} = \sum_{j=1}^{7} p_{ij(t-k)}$) and distribute advisory information based on it. In the case of South Korea, ambient air quality is classified by PM level into four categories: good (0–30 µg/m^3), moderate (31–80 µg/m^3), unhealthy (81–150 µg/m^3), and very unhealthy (151–600 µg/m^3), and air pollution advisories and alerts are issued based on the hourly PM level in extreme cases [85]. Avoidance via this channel is reflected in the term $A_{0k}(p_{i(t-k)})$.

Eq. 10 also accounts for the impact of any form of avoidance specifically associated with each origin, since the the extent of induced avoidance likely differs for PM from each origin. For example, as demonstrated in Fig. 2, the physical properites of PM differ across origins, resulting in different degrees of salience and coresponding defensive behaviors. For example, sea salt, wildfire smoke, and industrial-origin PM have different optical qualities and thus likely induce different responses. This potential heterogeneity is reflected by independent and separate functions A_{jk} for each PM origin

j. As a result, an increase of $p_{ij(t-k)}$ induces the avoidance behavior through both a component of avoidance that is origin-specific (A_{jk}) and also a component that reflects total PM (A_{0k}).

Expansion of the product in Eq. 10 yields numerous terms. Some terms are a function of PM from a single origin (e.g., $g_{jk}(p_{ij(t-k)})$ and $g_{jk}(p_{ij(t-k)}) \cdot A_j(p_{ij(t-k)})$). In addition, there are non-linear cross-terms that describe interactions between PM from differing origins (e.g., $g_{jk}(p_{ij(t-k)}) \cdot A_0(p_{i(t-k)})$ and $g_{jk}(p_{ij(t-k)}) \cdot A_{j'}(p_{ij'(t-k)})$, $j \neq j'$). Thus, under the parsimonious assumption that g_{jk} and A_j are each approximately linear (similar to the approach for undifferentiated PM) expansion of Eq. 10 yields a model that reduces to containing first and second-order terms for PM from each origin (similar to Eq. 7) as well as second-order terms that interact PM from different origins ($p_{ij(t-k)} \cdot p_{ij'(t-k)}$). Both sets of terms appear in the estimating equation below.

Estimation Combining Eqs. 9–10, expanding and simplifying the expression, and including the non-parametric trends and controls contained in $f(\cdot)$ from Eq. 8 yields our complete preferred model specification:

$$\begin{aligned} h_{it} = & \underbrace{\sum_{j=1}^{7}\sum_{k=0}^{28}\left(\alpha_{jk} \cdot p_{ij(t-k)} + \beta_{jk} \cdot p_{ij(t-k)}^2\right) + \sum_{j=1}^{7}\sum_{j' \neq j}\sum_{k=0}^{28} \mu_{jj'k} \cdot p_{ij(t-k)} \cdot p_{ij'(t-k)}}_{D(\mathbf{P}_{it})} \\ & + \underbrace{\sum_{n=1}^{4} \boldsymbol{\gamma}_n(\mathbf{q}_{nit}) + \sum_{m=1}^{3} \boldsymbol{\delta}_m(\mathbf{w}_{mit}) + \boldsymbol{\zeta}_{it} + \theta_t}_{f(\cdot)} + e_{it} \end{aligned} \tag{11}$$

where h_{it} is health costs and all terms correspond to their analogs in Eq. 8. Bar notation is removed to indicate where estimated values for parameters will differ from values in the undifferentiated model in Eq. 8 (since the model is estimated jointly, many parameters will be affected by allowing the effect of PM to be differentiated by origin, even if the specification for these terms does not change).

Analogous to the derivation of Eq. 7 for undifferentiated PM, $\alpha_{jk} \cdot p_{ij(t-k)}$ and $\beta_{jk} \cdot p_{ij(t-k)}^2$ terms in the first summation correspond to the estimated values that represent $g_{jk}(p_{ij(t-k)}) \cdot \left(1 - A_{0k}(p_{ij(t-k)}) - A_{jk}(p_{ij(t-k)})\right)$ in Eq. 10. However, there are now seven terms per lag, re-

flecting the seven origins (j) of differentiated PM. Consistent with the model for undifferentiated PM, there are 28 lags. The second summation of Eq. 11 contains many cross-terms that interact PM from different origins (j and $j' \neq j$). Each $\mu_{jj'k}$ corresponds to values for $g_{jk}(p_{ij(t-k)}) \cdot \left(A_{0k}(p_{ij'(t-k)}) + A_{j'k}(p_{ij'(t-k)})\right)$ terms in Eq. 10. Here, negative coefficient estimates would imply that the impact of PM from j on h_{it} is partially mitigated by avoidance behavior in response from simultaneously incident PM from j'. Other variables are the same as in Eq. 8.

Computing partial damages by origin To compute partial damage for PM from specific origins accounting for different simultaneous health impacts for each, we estimate parameters in Eq. 11 and isolate $D(\mathbf{P}_{it})$, the component describing heath damages from simultaneous incidence of PM from multiple different origins. We then estimate what health impacts would be if incident PM from a single origin were ρ and incidence from all other origins were zero. This approach is analogous to the calculations presented for the undifferentiated PM model, however in that case it was not necessary to be explicit about the level of PM from j' when calculating impacts from j; that said, the implicit assumption is identical, enabling a comparison between the two sets of results in Fig. 3.

Specifically, to compute partial damages from origin j, we compute $D(\mathbf{P}_{it} = \boldsymbol{\varrho}_j)$ where we define $\boldsymbol{\varrho}_j := (\mathbf{0}, \cdots, \mathbf{0}, \boldsymbol{\rho}_j, \mathbf{0}, \cdots, \mathbf{0})$ and $\boldsymbol{\rho}_j := (\rho_j, \cdots, \rho_j)'$. Thus, we set all 29 terms describing incidence from j to have the constant value ρ_j (i.e., $\mathbf{p}_{ijt} = \boldsymbol{\rho}_j$) and all elements in other vectors describing PM from non-j origins are set at zero (i.e., $\mathbf{p}_{ij't} = \mathbf{0}$, $\forall\ j' \neq j$). We then vary the value of ρ_j for each j independently to trace out a partial damage function for that origin. Fig. 3B presents $D(\boldsymbol{\varrho}_j)$ for the top four origins that contribute the most to the observed total PM in South Korea.

Damages from mixtures of PM To compute damages from mixtures of PM from different sources we use an approach that is similar to computing partial damages by origin, but we relax the assumption that PM from all non-j origins is zero. Instead, we continuously vary the quantity of PM incident from two different origins at a time and compute the resulting health damages, setting incidence from other origins to zero. Our calculation of the differentiated PM health damage function $D(\mathbf{P}_{it})$ enables us to compute damages from arbitrary mixtures of origin-specific PM,

however we restrict our presentation to two origins at a time for interpretability, since it is visually complex to present more dimensions simultaneously. Nonetheless, in our calculations of total health burden (described below), we compute impacts from observed mixtures of PM from all seven origins simultaneously.

Specifically, we estimate Eq. 11 and compute damages $D(\mathbf{P}_{it} = \boldsymbol{\varrho}_{jj'})$ where we define $\boldsymbol{\varrho}_{jj'} := (\mathbf{0}, \cdots, \mathbf{0}, \boldsymbol{\rho}_j, \mathbf{0}, \cdots, \mathbf{0}, \boldsymbol{\rho}_{j'} \mathbf{0}, \cdots, \mathbf{0})$, where the j^{th} vector is $\boldsymbol{\rho}_j = (\rho_j, \cdots, \rho_j)$ (i.e., $\mathbf{p}_{ijt} = \boldsymbol{\rho}_j$) and the j'^{th} vector is $\boldsymbol{\rho}_{j'} = (\rho_{j'}, \cdots, \rho_{j'})$ (i.e., $\mathbf{p}_{ij't} = \boldsymbol{\rho}_{j'}$). All other vectors are set at zero (i.e., $\mathbf{p}_{ij''t} = \mathbf{0}$, $\forall j'' \notin \{j, j'\}$). We compute the value for $D(\cdot)$ while continuously varying values for ρ_j and $\rho_{j'}$ to generate surfaces that describe damages from mixtures of PM from these two different origins.

Fig. 4 illustrates the surface of health responses that result from different mixtures of PM from distinct origins. For example, in the left panel of Fig. 4, we demonstrate the healthcare expenditure response to the mixtures of PM from South Korea and that from China, the first and second largest contributing origins to PM observed in South Korea, while setting PM from the other origins at zero. Fig. 4 visualizes the curvature of the surface $D(\boldsymbol{\varrho}_{jj'})$ with contours ("iso-damage" curves) on the $\rho_j \times \rho_{j'}$ plane. Each contour traces the combinations of $(\rho_j, \rho_{j'})$ that results in a fixed quantity of health expenditure.

Substitutability of components in a PM mixture We evaluate the behavior of health damages as a PM mixture is incrementally altered by considering the "substitutability" of PM constituents in the mixture. The slope of each contour in Fig. 4 enables us to evaluate the degree of substitutability for PM from the two origins in the mixture, given a specific baseline level of PM incidence from each $(\rho_j, \rho_{j'})$. Substitutability is the degree to which PM from origin j would need to be reduced in order to offset the health damage resulting from an incremental increase in PM from origin j'. Note that due to the of the nonlinear structure of the damage surface $D(\cdot)$, the substitutability of PM from each origin is a function of the PM load.

Substitutability is equal to the slope of a contour associated with health expenditure G in Fig. 4 (or a similar surface). The slope $\frac{\mathrm{d}\rho_j}{\mathrm{d}\rho_{j'}}\Big|_{D(\boldsymbol{\varrho}_{jj'})=G} = -\left(\frac{\partial D(\boldsymbol{\varrho}_{jj'})}{\partial \rho_{j'}}\right) / \left(\frac{\partial D(\boldsymbol{\varrho}_{jj'})}{\partial \rho_j}\right)$ will be -1 if a one-unit

increase (decrease) in PM from origin j could be substituted with a one-unit decrease (increase) in PM from origin j' with no change in the health response (holding PM from other origins at zero). A steeper or more negative slope would imply that the health impact of one unit of PM from the origin on the x-axis (PM from origin j) is larger than the impact of one unit of PM on the y-axis (PM from origin j') since a one-unit increase (decrease) of PM from origin j necessitates a larger decrease (increase) of PM from origin j' to achieve the same health response. For example, in the left panel of Fig. 4, across the contour corresponding to one million USD, the slope is approximately -2.6, implying that exposure to one additional $\mu g/m^3$ of PM from China has the same incremental impact on health costs in South Korea as exposure to an additional 2.6 $\mu g/m^3$ of PM from South Korea.

Estimation of total PM-specific health damages by origin We use our empirical measurements of PM-specific damages to estimate the total health costs experienced by the South Korean population due to PM emissions from each origin. We do this by computing health costs due to actual incidence of PM by origin (estimated by partitioning observed PM, described above) relative to the health costs that would be expected if emissions from a single origin j were eliminated. This approach holds the emissions of all non-j at their observed level in both cases, so all that is changed is the PM from j. For example, we compute health costs in South Korea with all origins contributing PM and with contributions of PM from China shut off—with the difference representing the net health impact of PM from China. Specifying the emissions of all non-j origins in both scenarios is important because $D(\cdot)$ is nonlinear in PM from all origins. We believe this approach provides the most accurate estimate for the overall total impact that individual origins have on health in South Korea.

To estimate total harm by origin, we estimate Eq. 11 and compute damages $D(\cdot)$ for our study period in two scenarios: one with PM equal to actual incidence ($\mathbf{P}_{it}$) and a counterfactual scenario with emissions from origin j set to zero ($\mathbf{P}^0_{ijt}$). The difference in damage between these two is the

harm traceable to PM from origin j. Specifically, we compute

$$\widehat{\text{total damage}}_j = \sum_{i \in I} \sum_{t \in \mathbf{T}} \phi_{it} \cdot \left[D(\mathbf{P}_{it}) - D(\mathbf{P}^0_{ijt}) \right] \tag{12}$$

where $\mathbf{P}_{it}$ is the matrix of PM values that are actually incident on poulations (the same values used in estimation) and $\mathbf{P}^0_{ijt}$ is the same, except that the PM values corresponding to origin j are set at zero (i.e., $\mathbf{p}_{ijt} = (p_{ijt}, \cdots, p_{ij(t-28)}) = \mathbf{0}$). For example, for the origin with index $j = 1$, $\mathbf{P}^0_{i1t} = (\mathbf{0}, \mathbf{p}_{i2t}, \cdots, \mathbf{p}_{i7t})$. Therefore, the difference $D(\mathbf{P}_{it}) - D(\mathbf{P}^0_{ijt})$ is the average per capita damages in district i on day t attributable to origin j. This value is then scaled by ϕ_{it}, the population in district i at time t, and aggregated across days $t \in \mathbf{T}$ and regions $i \in \mathbf{I}$. The bounds of aggregation vary, with values in Fig. 5A–C aggregating over specific years (and all districts) and Fig. 5D–F aggregating across districts within a province (and all years). Table 1 aggregates across all districts and all years in the sample.

D Supplementary Figures

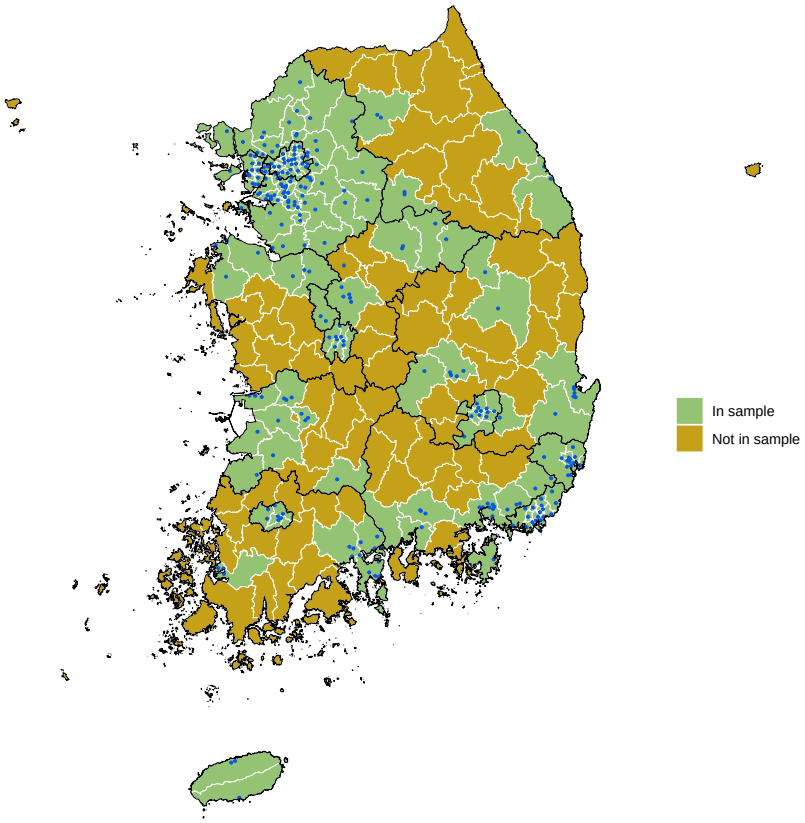


Figure S1: **Boundaries of districts and provinces; locations of air pollution monitors** Districts in our main model (those meeting minimum data length and monitor coverage requirements) are shown in green. Borders between districts within a province are shown in white, and borders between provinces are shown in black. Pollution monitors used in the main analysis are shown as blue points.

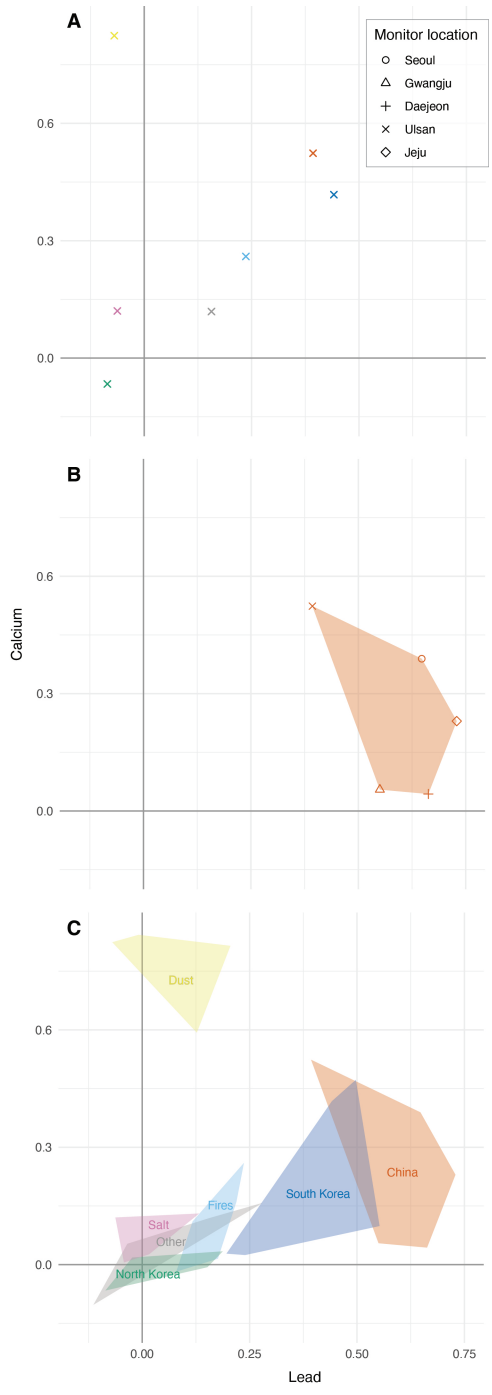


Figure S2: **Extended explanation of Figure 2.** (A) Partial correlations between PM of each origin and the concentrations of atmospheric lead (x-axis) and calcium (y-axis) using only data from the chemistry monitor in Ulsan. (B) Partial correlations between PM from a single origin (China) and atmospheric lead (x-axis) and calcium (y-axis) using data from all five chemistry monitors; a convex hull is drawn around these points to aid pattern recognition. (C) depicts these convex hulls for PM of each origin and is the form shown in Figure 2.

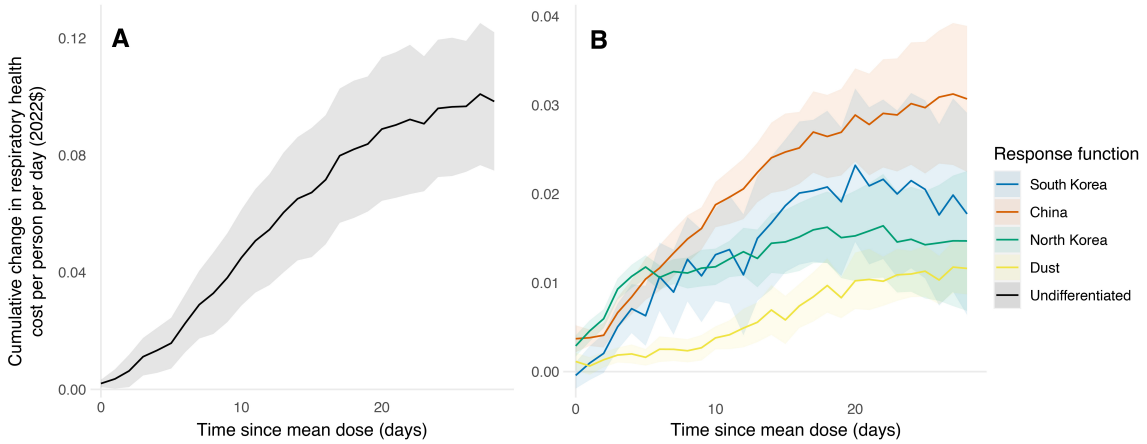


Figure S3: **Lag–response relationships.** (A) The undifferentiated lag–response relationship, calculated with levels of all sources at their mean values (is shown only for comparison and is not used in our calculations for the attribution of respiratory health spending to PM origins). For full model specification, see Methods. (B) The cumulative lag–response relationships for PM by origin.

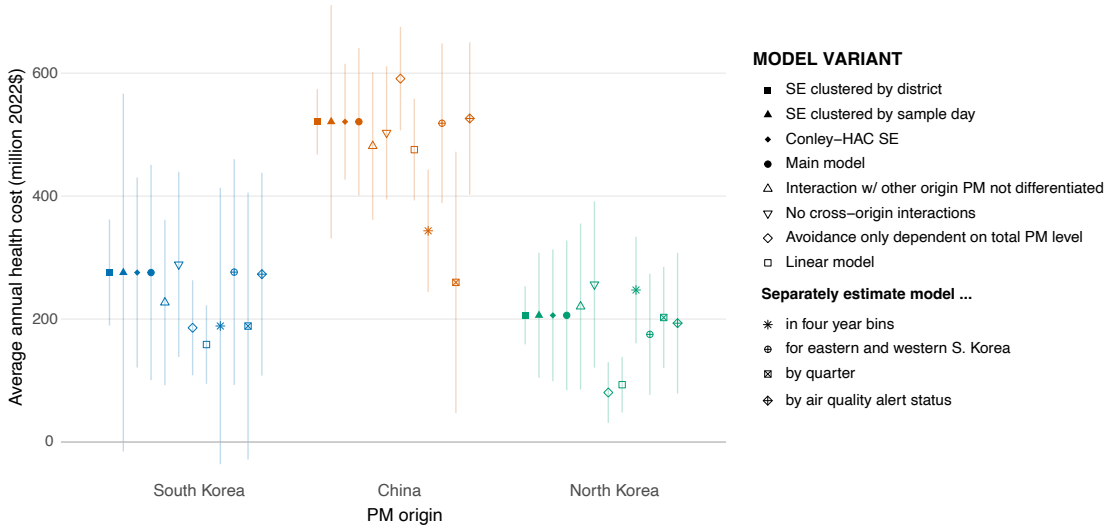


Figure S4: **Model robustness checks.** For ease of comparison, the model is estimated using alternate specifications and model results are combined with estimated exposures to determine differences in average annual attributed health costs by origin. Line ranges indicate 95% confidence intervals. Results for the main model, which follows Eq. 11 with standard errors clustered at the province level, are shown fourth from the left. The three estimates to the left of the main estimates show sensitivity to changing the method used for calculating standard errors: for the first set of estimates, standard errors are clustered at the district level; for the second set of estimates, standard errors are clustered at the day level; for the third set, standard errors are calculated using a spatial heteroskedasticity and autocorrelation robust (Conley-HAC) method [128]. Estimates 5 through 8 present variants of Eq. 11, but with a modified functional form that is progressively less flexible: for the fifth set of estimates, the model includes linear and quadratic terms for each origin of PM and interactions between each origin of PM and the total quantity of PM from every other origin (interactions with other PM are not differentiated); for the sixth set of estimates, the model includes only linear and quadratic terms for PM of each origin (no cross-origin interactions); for the seventh set, the model includes only a linear term for each origin of PM and an interaction of that term with the total level of PM (the curvature of each origin's dose–response function is dependent only on the total level of ambient PM); for the eighth set, the model includes only a linear term for each origin of PM. Estimates 9 through 12 separately estimate Eq. 11 by four year bins, for the eastern and western half of South Korea, for each quarter of the calendar year, and for days when air quality alerts were and were not issued, respectively.