The Effect of Air Pollution on Mortality: Evidence from Wildfire Smoke in Mexico

Maximilian Nohr*

November 5, 2024
(Click here for latest version)

Abstract

Wildfires produce large amounts of air pollution via smoke, which can travel far beyond the original fire location. This paper studies the causal effect of wildfire smoke exposure on mortality in Mexico. I merge satellite image data on wildfire smoke plumes with administrative death records and leverage high-frequency variation in smoke exposure within municipalities over time. Using data from air pollution monitors, I show that wildfire smoke over a municipality increases PM2.5 air pollution by 11%. At the same time, mortality increases by 1.87 deaths per million on the day of smoke, and by an additional 1.69 deaths per million over the next three days. The mortality effects are concentrated among individuals over 60 years old, with the largest effects for those over 80 and no effects for those below 60. The main effect on short-term mortality in Mexico is high compared to prior studies in developed countries. Within Mexico, the effects are also larger for individuals in poorer municipalities. Overall, this paper provides new evidence on short-term mortality effects of wildfire smoke across all age groups in Mexico, and suggests key heterogeneity in the harms of air pollution by income.

*PhD Candidate, Department of Economics, University of Texas at Austin, m.t.nohr@gmail.com.

Air pollution is considered to be one of the largest environmental risk factors to human health (McDuffie et al., 2021), and wildfires are the largest single source of fine particulate air pollution. The frequency of extreme wildfire events has increased dramatically in recent years, and with the effects of climate change on drought conditions, wildfires are projected to become even more frequent and destructive. The negative effects of this increase and the associated effects on air pollution are unlikely to be distributed equally. While pollution affects all countries and regions of the world, people in lower income countries are typically more affected. Lower income countries suffer from higher levels of baseline pollution, and they simultaneously lack the resources to combat this high pollution or address its negative effects on health. Despite these concerns, many studies on air pollution are conducted in high income countries like the United States, raising concerns about the external validity of their results to less developed countries.

Small particulate air pollution is thought to be the most harmful to human health, and wildfire smoke is a major and growing source of this pollutant. In North America, for example, the average share of fine particulate matter coming from wildfire smoke increased from 9% in 2002-2011 to 18% in 2012-2021.[1] While the average share of fine particulate matter coming from wildfire smoke was 9% from 2002-2011, it increased to an average 18% of all fine particulate air pollution over the next ten years. This example is illustrative of global trends in increasing wildfire smoke pollution, and Mexico is also directly affected, as smoke from North American wildfires travels hundreds of kilometers and frequently leads to increased air pollution in Mexico.

This paper estimates the causal effect of air pollution on short-term mortality, using plausibly exogenous variation in air pollution from wildfire smoke. My analysis combines nationwide administrative mortality records with satellite derived smoke data from the U.S. National Oceanic and Atmospheric Administration (NOAA) Hazard Mapping System. By overlaying smoke plumes with municipality outlines, I construct a municipality-level smoke exposure variable, that categorizes days as either smoke days or non-smoke days.

[1] Based on U.S. data from the EPA:`https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data`

Studying air pollution from wildfire smoke in this setting has several advantages. First, the large, high-frequency dataset I construct allows me to estimate short-term mortality. Spikes in air pollution from wildfire smoke are plausibly exogenous, driven by the combination of fire location and prevailing wind patterns, allowing credibly causal estimates of the effect on mortality. Second, mortality records provide reliable, daily observations that cover all of Mexico, and are available over a long period (2007-2019), resulting in enough power to estimate mortality effects, which can be difficult as death is such a rare event. Records are available for all ages, and all municipalities in Mexico, allowing me to study mortality effects for different age groups and heterogeneity by municipality-level characteristics. Additionally, the data contain cause of death identifiers, which I use to study potential mechanisms.

I first show that smoke plume coverage increases ground-level air pollution. To do so, I combine the NOAA data on smoke plumes with data from Mexico's network of pollution monitoring stations. I find that smoke plume coverage leads to a large, transient shock in ground-level air pollution. Smoke coverage increases small particulate air pollution ($PM_{2.5}$) by $2.27\mu g/m^3$, a 10.8% increase over non-smoke days (.21 standard deviation (SD)). Larger particulate air pollution (PM_{10}) also increases by an average of $3.72\mu g/m^3$, which is a 7.9% increase over non-smoke day PM_{10}, or an increase of .15 SD. I also find a small increase in ozone levels on days of smoke coverage by 2.7% over non-smoke days, a .05 SD increase, and no statistically significant effect on SO_2.

For my first main result, I study the effects of wildfire smoke on mortality. The analysis shows that smoke increases mortality significantly, and mortality effects are concentrated in the elderly population. One day of wildfire smoke increases mortality by 1.93 deaths per million on the same day, for those who are over 60 years old. Death rates stay elevated on the first few days after a smoke day, resulting in total excess mortality of 3.54 deaths per million over 3 days, from one day of smoke coverage. This is an increase of 4.3% over the average daily mortality rate. I find no increase in mortality for age groups under 60 years old, but among those over 60, I find the largest mortality effects in the older population groups. Among ages 60-69, three day excess mortality is 1.76 deaths per million, an increase of 5%

over the daily average. For those 70-79 years old, wildfire smoke leads to an excess mortality of 2.40 deaths per million, an increase of 3% over daily average. For those over 80 years old excess mortality is 12.08 deaths per million (5.0% increase). I find no evidence of offsetting decreases in mortality over the month following a smoke event, suggesting that the mortality effect is not driven by harvesting.[2]

Next, I study increases in mortality by cause of death for those over 60 years old. I find that the largest increase in mortality is among individuals with circulatory causes listed as their cause of death.[3] I don't find a statistically significant increase in respiratory causes, which is consistent with the interpretation that smoke days affect people with pre-existing conditions. Under this explanation, wildfire smoke is a stressor that affects those individuals most who are already weak from another illness, and causes their death. Mortality from external causes, such as car accidents, does not increase. I also find an increase in a broad summarizing category of "all other internal causes.[4]" These results are consistent with prior research conducted in the United States by Deryugina et al. (2019), who find increases in mortality from air pollution for circulatory causes, cancer, and internal causes.

Comparing my results to mortality estimates of smoke induced air pollution from the United States, I find that the mortality increase from air pollution in Mexico is much larger. Miller et al. (2024) find an increase of 1.33 deaths per million over 65 over 3 days, from one day of light smoke coverage. My estimate for Mexico of 3.54 deaths per million over 60 is over twice as large, highlighting the importance of studying the impacts of wildfire smoke in different settings.[5] There are different possible explanations why people in less developed countries could be

[2]Harvesting is when weak individuals die slightly earlier than they would have without the short-term shock in air pollution, which can lead to positive estimates for the short-term mortality effect from smoke, which would be offset by a equally sized decrease in mortality later, when those individuals would have died otherwise.

[3]The most common cause of death in this category is a heart attack, the category also includes strokes and other circulatory system related causes.

[4]Other internal causes includes, for example, different types of diabetes and liver cirrhosis.

[5]The comparison is not perfect, because we run different regression. Miller et al. (2024) use only light smoke, whereas I use all smoke, which slightly increases my estimates (still, most smoke coverage is classified as light coverage), and I study mortality over 60, instead of over 65, which slightly decreases my estimates, since death rates and mortality effects are lower among those 60-65.

more affected by air pollution, including differences in baseline health, differences in exposure due to time spent outside, and the accessibility of healthcare during acute need.

To examine potential reasons for the large difference in mortality effects, I study heterogeneity within Mexico. One possible explanation is the difference in baseline air pollution, which is twice as large in Mexico compared to the U.S. However, I study heterogeneity by baseline air pollution across municipalities within Mexico, and I do not find significant differences in mortality effects. Another possible explanation is differences in baseline health. Mexico ranks among the countries with the highest rates of diabetes in the world.[6] When studying mortality rates by differences in municipality-level diabetes rates, there is some indication that municipalities with higher rates are more affected, but the difference is not statistically significant. The same is true for municipality obesity rate and rate of people suffering hypertension. I also study heterogeneity by a municipality-level development index[7] A lower municipality development index is associated with a slightly higher estimated mortality effect, but again the difference is not statistically significant. Last, I study heterogeneity by income. I find that lower income municipalities have higher mortality from wildfire smoke compared to high income municipalities, and this difference is statistically significant. This suggests that the combination of different individual factors of health and development, which are correlated with income, can together explain some of the heterogeneity in mortality effects.

My paper contributes to the academic literature in several ways. First, it adds to a growing, quasi-experimental economics literature on the causal effect of air pollution on mortality. Much of the prior work in this literature has been conducted in the United States (Chay et al., 2003; Chay and Greenstone, 2003; Currie and Neidell, 2005; Currie et al., 2009; Moretti and Neidell, 2011; Deschênes and Greenstone, 2011; Deschênes et al., 2017; Schlenker and Walker, 2015; Knittel et al., 2016; Deryugina et al., 2019). Studies investigating mortality from air pollution in less developed countries mostly focus on infant mortality (Gutierrez, 2015; Arceo

[6]Based on World Bank data: `https://data.worldbank.org/indicator/SH.STA.DIAB.ZS`.

[7]Values from 2014, calculated by the Human Development Research Office (OIDH) in Mexico using United Nations Development Programme methodology.

et al., 2016; Adhvaryu et al., 2024). Both of the studies conducted in Mexico study infant health, and they find mixed results. Gutierrez (2015) finds that there is some suggestive evidence that infant mortality from air pollution is higher in lower socio-economic status municipalities, while Arceo et al. (2016) compare their estimates of infant mortality in Mexico City with U.S. estimates and find that infant mortality is similar or even lower in Mexico City. My paper uses high-frequency, nationwide data over 13 years to provide causal estimates for all age mortality from air pollution. These estimates inform the ongoing policy discussion about pollution in developing countries. They also underline the magnitude of the air pollution problem. Even when studying a rare outcome such as mortality, I find significant effects from air pollution, suggesting a large negative impact on acute health overall, both in terms of other negative short-term health outcomes from transient pollution, which I am unable to measure directly, as well as long-term impacts, which are generally harder to estimate causally.

Second, my paper contributes to a recently developing literature on (wildfire) smoke (e.g. Sastry (2002); Jayachandran (2009); Rangel and Vogl (2019); Arenberg and Neller (2023); Miller et al. (2024)). Of these, Sastry (2002) studies adult mortality from one extreme wildfire event in the context of a developing country, although they don't have a direct measure of smoke and instead use PM_{10} concentration. They are also limited to study one extreme event, rather than typical exposure to wildfire smoke throughout the year, which is arguably better suited to study the typical mortality effects of smoke. The other papers in this category study infant mortality (Jayachandran, 2009; Rangel and Vogl, 2019) or are again focused on the United States (Arenberg and Neller, 2023; Miller et al., 2024). Air pollution from wildfires is a growing concern, as it represents a large share of small particulate air pollution. This paper provides new, important estimates of the mortality effects of wildfire smoke for different age groups and in regions with different development.

The remainder of this paper is structured as follows. Section 1 describes the data. Section 2 explains the empirical strategy. Section 3 presents the main results, as well as robustness checks, and section 4 presents heterogeneity results. In section 5, I conduct a cost analysis. Section 6 concludes.

1 Data

I use three main sources of data to estimate the effect of wildfire smoke on mortality. First, I measure daily smoke exposure using high-frequency satellite images from the U.S. National Oceanic and Atmospheric Administration (NOAA). Second, I supplement this data with air quality measures from local pollution monitors collected by the Mexican National Institute of Ecology and Climate Change (SINAICA). Finally, I combine this with administrative mortality records from Mexican National Institute of Statistics and Geography (INEGI).

1.1 Smoke data

The smoke data are derived from daily satellite images, analyzed by NOAA. This data product was first published in 2003 to provide real time data on wildfires and wildfire smoke over all of North America, including Mexico. I use data from September 2006 through the end of 2019. Typically, two satellite images are taken during daylight hours and are combined into one daily smoke coverage file. Analysts can use the consecutive images to help distinguish smoke from clouds or snow coverage, but other sources of smoke, e.g. agricultural burns, are not distinguished in the data, a fact I address in the robustness section. The HMS then publishes one file per day, containing the location of all detected smoke plumes. I take the union of all smoke outlines to be the coverage area on that day, and calculate a daily measure of smoke cover at the municipal level by measuring the area of the municipality that is covered by smoke, and classifying a municipality as treated if at least 95% of its area are covered by smoke on that day.[8] I choose 95% as a cutoff to allow for some numerical noise in the geographical overlap calculation, and I show in Figure 1 that the cutoff of 95% is not decisive. Conditional on some part of the municipality being covered by smoke, in almost 90% of observations then the entire municipality is covered, which is due to the size difference of municipalities and average smoke plumes. The HMS also has daily fire date, which I use to create

[8]While the data include a density variable for each smoke plume starting from 2009 (low, medium high density), in this analysis I treat any smoke coverage the same for the purposes of my analysis.

an indicator variable that is equal to one if a fire was detected inside municipality boundaries on that day.

1.2 Air Pollution data

I supplement the smoke indicator with air quality monitoring data collected by the Mexican National Institute of Ecology and Climate Change (SINAICA). I collect hourly $PM_{2.5}, PM_{10}, O_3$ and CO concentration levels and take the average of all available measurements on a day to get a daily measure. To calculate a pollution measure at the municipal level, I average over all monitoring stations within municipality boundaries. The availability of monitoring data varies between the different types of pollutants, from 119,006 municipality-day obsevations for $PM_{2.5}$ and 224,281 observations for ozone (O_3). I also collect model-derived $PM_{2.5}$ measurements for the entire study period from 2005 through 2019 from the American National Aeronautics and Space Administration's MERRA2 reanalysis project, which incorporates satellite measurements of air pollution to estimate pollution at specific grid points. For this I match each municipality to the closest grid point value, and I use this data for baseline pollution classification in the heterogeneity section.

1.3 Mortality data

My mortality data come from the Mexican National Institute of Statistics and Geography (INEGI). Every year, INEGI publishes mortality files at the individual level. Working with administrative records is a great advantage, as Mexican records are considered to be of high quality (Hernández et al., 2011), and they contain the date of death, the municipality where the death occurred, date of birth and a cause of death categorization, as well as some demographic characteristics of the deceased. I aggregate daily deaths at the municipality level. To calculate my main outcome variables, mortality per million, I combine the mortality records with population data from the 2000, 2010 and 2020 censuses.[9] When studying cause of

[9]For years in between censuses, I linearly interpolate municipality-level population counts.

death, I aggregate daily mortality within each of the cause of death categories I study.

1.4 Additional data

I collect additional municipality-level data on development, education and health to study heterogeneity in mortality effects, these data come from INEGI. For weather controls, I use data from the NARR, a reanalysis model that incorporates large quantities of observational data into a regional climate model, which then provides temperatures, precipitation and additional climate variables on a grid.

2 Methodology

In my main specification, I use a regression model with fixed effects to estimate the reduced form effect of wildfire smoke exposure on mortality. I include municipality by week-of-year fixed effects so that the effect is identified from within municipality-season variation. Additionally, I include date fixed effects, which control for nation-wide daily shocks, such as national holidays, which can affect mortality. Weather controls are contained in X_{it} and include for temperature, precipitation and precipitation squared.

$$y_{it} = \alpha + \Sigma_{k=-15}^{15} \beta_k smoke_{i(t-k)} + X_{it} + \gamma_{iw} + \gamma_t + \varepsilon_{it} \tag{1}$$

The treatment variable of interest, $smoke_{it}$, is equal to one if at least 95% of municipality i's area was covered by smoke on date t and is 0 otherwise. I include fifteen leads and lags of smoke exposure in the regression due to concerns about serial correlation that could otherwise bias my point estimates, and to provide some evidence of the validity of the estimation strategy. If the specification is correct, then coefficients on the leads of smoke should be close to 0, since smoke in the future should not affect mortality today. In the results tables, I present β_0, the coefficient on $smoke_i t$, as well as the sum $\Sigma_{k=0}^{2} \beta_k$, which is the sum of the coefficients on smoke and two lags of smoke, and can be interpreted as the effect of smoke exposure on the day of smoke and the two following days. I provide plots of the lead and lag coefficients to examine the impact of smoke exposure over time.

The identifying assumption of my regression is that conditional on included fixed effects and leads and lags, $smoke_{it}$ is uncorrelated with the error term ε_{it}. The validity of this assumption relies on the idea that smoke exposure in a given municipality on a given day depends on prevailing wind patterns and fire locations on that day, which are plausibly exogenous to other causes of mortality in the municipality. In the robustness analysis I will probe the validity of this identifying assumption.

In regressions that have mortality as the dependent variable, I weigh observations by the relevant population in order to improve the precision of my estimates and avoid over-weighting small municipalities. In all regression I cluster standard errors at the municipality level, given that errors are likely to be correlated between daily observations from the same municipality.

3 Results

3.1 Air Pollution

Table 2 presents the coefficient on smoke exposure from equation 1. Columns 1 shows the effect of a day of smoke on the concentration of fine particulate matter, particles smaller than 2.5 μm in diameter. A day of smoke increases the concentration of $PM_{2.5}$ particles by 2.27 units, which is an increase of approximately 10% over the daily mean measured at the ground-based monitors in my sample. Column (2) shows the effect of smoke on PM_{10}, which are slightly larger particulate matter. Smoke increases PM_{10} concentration by 3.72 units, an increase of 8% over the daily mean.

In panel (a) of Figure 5, I plot the coefficients on the leads and lags of smoke from the main equation 1 with $PM_{2.5}$ as the dependent variable. The coefficients on the leads of smoke are centered around zero, lending credibility to the research design, since future smoke should not increase air pollution today. However, I do find a significant increase in air pollution on the day before the smoke day, which can be explained by the fact that satellite pictures are taken during daylight hours, while pollution monitors run 24 hours. Hence, Smoke clouds that start covering the municipality during the night will already increase air pollution, but will only show up in the smoke dataset on the next day. After the smoke day, air pollution remains elevated for two days before returning back to centering around zero. The statistically significant increases in ground-measured air pollution show that the satellite derived smoke measure is a good proxy for increased air pollution on the ground.

3.2 Mortality

Table 3 shows the effect of wildfire smoke on mortality for 10-year wide age groups The table presents the coefficient on smoke from equation 1, as well as the sum of the coefficients on smoke and two lags of smoke, which estimate the increase in mortality one day and two days after a smoke event. I find no short-term mortality effects from wildfire smoke for all age groups under 50 years old, this is true for the one day mortality represented by the coefficient on smoke β_0, as well as the 3 day summed mortality. For the group 50-59 in column (6), there is potentially a small increase in mortality, but this effect is not statistically significant, and mortality over 3 days does not increase for this group.

The three oldest age cohorts on the other hand all show statistically significant increases in mortality. In the age cohort 60-69, a day of wildfire smoke increases mortality by 1.25 deaths per million on the day of the smoke, and by 1.76 deahts per million over 3 days, an increase of 4.8% over the average daily mortality of that group. For those age 70-79, mortality on the smoke day increases by 1.68 deaths per million, and by 2.4 deaths per million over three days, which is an increase of 2.8% percent over average daily mortality. I find the largest mortality effect for the age group of those over 80 years old, where mortality increases by 4.8 deaths per million on the day of smoke and by 12.08 deaths per million when including the two days following the smoke. This is an increase of 4.9% over average daily mortality in that group. Previous studies in the U.S. such as Deryugina et al. (2019); Miller et al. (2024) were limited to study mortality and health effects in the Medicare population, people over 65, so the null result of this paper for younger age groups adds to our knowledge and shows that the U.S. papers are, despite their data limitations, able to study the groups most affected by smoke induced air pollution. These results are robust to different

A concern when studying the short-term mortality effects is that some deaths might have occurred within a couple of days, with or without the smoke exposure. This mechanism of inter-temporal displacement of mortality is called "harvesting," by which a weak individual, who would have died within a short amount of time without any external influences, could die earlier because of the additional stress caused in this case from the wildfire smoke. This is a concern, because if that is the case, then the mortality estimated in this paper would not represent a significant cost to society, because those individuals most affected would have died shortly anyway, with or without the exposure to smoke.

In the main specification of this paper, this effect would show up in the coefficients on the lags of wildfire smoke. If increased mortality on the smoke day was offset by an equally large decrease in mortality in the next week, then that would show up as a negative coefficient on the eighth lag of smoke. Panel (b) of Figure 5 plots the coefficients on smoke and 15 of its leads and lags. Mortality increases on the day of smoke, marked at 0 in the figure, and stays elevated for two days, before returning back to zero. As with the effects on air pollution, mortality is not affected by future smoke, another indicator of the validity of the estimation strategy. If these results were driven by harvesting (from within 15 days of the smoke event), then that would show up through negative coefficients on the lags of wildfire smoke. Those coefficients are all centered around zero and statistically insignificant, making short-term harvesting less likely to be the cause of the large mortality effects I find, at least within two weeks of the exposure to wildfire smoke.

3.3 Cause of Death

In this section I study the heterogeneity of treatment effects. First, I will show results by different listed cause of mortality on the death certificate. Then I will compare my results to results from a comparable study in the U.S. and explore different potential mechanisms to explain the differences in results. The biological mechanisms of the acute effects of air pollution on the human body is an open area of research, but studies have found that air pollution can cause stress to the body, including the cardio-vascular system, inflammation, and other changes that can increase the risk of acute adverse health impacts (Mannucci and Franchini, 2007). Table 5 presents the effect of wildfire smoke on mortality by different cause categories, for mortality over 60. The first column shows deaths listed as caused by circulatory causes. The most commonly listed cause in this category are heart attacks, it also includes other causes related to the circulatory system such as strokes or acute hypertension. Deaths in this category increase by 0.74 deaths per million on the day of the smoke event, and by 1.92 deaths per million over three days. Column (2) shows results for respiratory causes, the most common of which is chronic pulmonary disease. In this category I don't find a significant increase in mortality from smoke induced air pollution. Other internal causes, which includes all other internal causes of mortality, such as for example different forms of diabetes or liver cirrhosis, increase by 1.08 deaths per million on the day of smoke and 1.6 deaths per million over three days. External causes, such as car or work accidents, do not increase. The results from this analysis are in line with the view that air pollution can exacerbate existing conditions and thereby effect those individuals the most who are already weaker because of their underlying health. My results are consistent with prior literature, such as Deryugina et al. (2019), who study the effect of air pollution on mortality in the United States using wind direction as an instrument for pollution and who find that air pollution increases cardiovascular mortality rates, cancer related deaths and other internal causes, but does not increase deaths from external causes.

3.4 Robustness Checks

I run different variations of the main estimating equation to show that the results I am finding are not dependent on particular choices regarding the estimating strategy. In table 6 I show that different specifications of weather controls yield similar results to my main specification, which is reproduced in column (1). Column

(2) reports results when not controlling for any weather, and column (3) controls for bins of temperature rather than a linear control, the three day mortality effects estimated from these regression are close to the main estimate of 3.54 deaths per million, with 3.77 and 3.11 deaths per million respectively. In column (4) I address the concern that results could be driven by smoke from agricultural burns, instead of wildfire smoke. This could present a problem if agricultural burns are correlated with unobservables that also affect mortality rates. In this regression, I exclude the months of March through May from the sample, which are the three months with the highest activity of agricultural burn activity in Mexico. The estimate for the one day effect from wildfire smoke of 2.23 deaths per million is close to the main estimate of 1.93 deaths per million, and estimated three day mortality is 5.67 deaths per million, slightly larger than the main estimate, suggesting that my results are not driven by agricultural smoke that may be correlated with economic activity.

To show that my results are not driven by the choice of the number of leads and lags of smoke, I run the main regression from 1 with no leads or lags, 7, 15 (main specification) and 20 leads and lags and present the results in appendix table 7. The estimate for the one day smoke effect is larger when no leads and lags are included, which is expected, since conditional on smoke today, the likelihood of the presence of smoke yesterday is quite high, and we are now not controlling for yesterday's smoke. For 7, 15 and 20 leads and lags, the one day smoke effects are all very similar at 1.87, 1.93 and 1.96 deaths per million, respectively, as are three day smoke effects, which are estimated to be 3.41, 3.54 and 3.60 deaths per million, showing that the choice of how many leads and lags to include in the regression is not a main driver of the results. Note that the number of observations for each of these regressions is slightly different, since the number of leads and lags determines how many observations need to be dropped from the beginning and end of the sample period due to missing values.

4 Heterogeneity

4.1 Comparison to estimates from the U.S.

The mortality effects from wildfire smoke I find in Mexico are large compared to previous papers. Comparing directly with Miller et al. (2024), while they find a comparable increase in $PM_{2.5}$ air pollution caused by wildfire smoke in counties in the United States, they estimate that a day of light smoke causes an increase in 3-day mortality for those over 65 years old of 1.33 deaths per million (confidence interval: [0.76,1.90]). For medium and thick smoke days, they find increases in mortality of 1.63 and 1.42 deaths per million, respectively. My estimate for the 3-day mortality effect from smoke exposure[10] in Mexico is 3.54 deaths per million, at least two times the size of the effect estimated for the elderly population in the U.S., with confidence intervals that don't overlap.

One explanation could be the difference in ambient air pollution, which is approximately twice as high in Mexico compared to the U.S., according to different sources such as the World Bank. Ambient air pollution has been linked to negative health outcomes (Mannucci and Franchini, 2017), but whether higher levels of underlying air pollution positively or negatively affect smoke induced mortality is not well understood. If higher ambient air pollution is a stressor, then people living in higher air pollution environments might be more susceptible to negative effects from air pollution. To test this theory I can study the heterogeneity by ambient air pollution within Mexico. I split my sample into four quartiles of average baseline $PM_{2.5}$ air pollution on non-smoke days[11], and separately estimate the mortality effects in each sample. The results for 3-day mortality are presented in figure 6. Even though the point estimates look like they are slightly different between the quartiles, the difference between the first and last quartile is not statistically significant.

[10] I don't distinguish between different intensities of smoke, but light smoke is the most common.

[11] Because the sample of ground-based $PM_{2.5}$ observations is so small, I use satellite-derived pollution values from a NASA atmospheric model for this heterogeneity analysis.

4.2 Heterogeneity within Mexico

Mexico and the United States differ in more than just ambient air pollution. Mexico is considered a developing or middle-income country, average income is low compared to the United States. While income itself is unlikely to directly affect mortality from air pollution, it is correlated with a lot of factors that could.

To test whether people in lower income municipalities are differently affected than those in high income municipalities, I split the sample into four quartiles by municipality average household income, and run the main regression from 1 separately for each quartile. The results are presented in figure 7. Each dot represents the 3-day mortality estimate on the y-axis, and the quarterly income in Mexican pesos in that quartile on the x-axis. The point estimates suggest that there is a difference in effect sizes by income quartile, with a mortality effect approximately of 5 deaths per million from a day of smoke in the lowest income quartile and slightly below 2 deaths per million for those in the highest income quartile. In this case, the difference between the lowest and the highest income quartiles is statistically significant at the 5% level. To better understand this heterogeneity of effects by income, I repeat the split analysis for the sample of over 70 year olds and over 80 year olds, the results are presented in the top panel of figure 8 and show that for all age groups, mortality from smoke is significantly higher in lower average income municipalities.

To try and better understand what could cause this heterogeneity by income, I repeat the same for different municipality-level indicators of health and socioeconomic development, for each plotting the difference between the lowest and highest quartile and the 95% interval for this difference, shown in the bottom three panels of figure 8. For all of the regressions used for the bottom three panels, the outcome variable is deaths per million over 60. None of the environmental factors show significant heterogeneity between quartiles. The difference between high and low $PM_{2.5}$ municipalities, as discussed previously, is not significant, and neither are the differences for the hottest versus coolest municipalities, or those with higher and lower precipitation.

The health panel shows heterogeneity by quartiles of municipality-level rates of obesity, hypertension and diabetes. Plotted is the difference between the lowest quartile of obesity, hypertension and diabetes rates and the highest quartile, which for obesity and hypertension are negative, suggesting that effects are slightly higher in municipalities with higher rates of obesity and hypertension, but these differences are not statistically significant. The bottom panel of the figure shows some development related correlates of income, an education index, access to healthcare, development index, and extreme poverty rate. Here, the only significant difference between highest and lowest quartiles is found for the poverty rate, where municipalities with the highest rates of extreme have higher mortality effect by 3.91 deaths per million compared to municipalities with the lowest poverty rates, consistent with the results of heterogeneity by average income. While not statistically significant, the mortality effects from air pollution are estimated to be 1.93 deaths per million higher in those municipalities in which a higher rate of the population lacks access to healthcare (p-value: 0.19).

5 Cost Analysis

Policy makers have some control over wildfire frequency and severity through a variety of different mitigation strategies, such as prescribed burns, wildfire suppression, and firefighting.[12] These strategies are complex and cost intensive. The results form this paper can help policy makers better understand the costs of wildfire smoke, and thereby help policy makers make better decisions about the optimal level of wildfire mitigation. In the following section, I will use the point estimates on mortality to estimate the social cost of the short-term mortality caused by wildfire smoke in Mexico. Additional to providing policy makers with important information on the costs of wildfire smoke, this analysis can also provide context on the size of the mortality effects estimated in this paper and the changes in impacts over time.

To estimate the overall impact of wildfire smoke on short-term mortality in Mexico over the time period of study, 2007 through 2019, I take the 3-day estimates

[12] While prescribed burns still cause air pollution, there is some evidence to suggest that it is less than the pollution from avoided larger wildfires (Selimovic et al., 2020)

from table 3 for those three groups with statistically significant effects, 60-69, 70-79 and over 80 years old, which are 1.76, 2.40 and 12.08 deaths per million, respectively. For each municipality, I multiply the population in each age group with the estimated mortality effect and the number of smoke days in that year. Over the entire 13 year period, I estimate that in the age group 60-69 there are 1,683 excess deaths caused form wildfire smoke. For the group 70-79, I estimate 1,232 excess deaths, and finally for those over 80 I estimate 3,076 excess deaths from wildfire smoke. Note how the total number of excess deaths are higher for the population in their 60s compared to those in their 70s, despite the smaller increase in deaths per million, because the population in their 60s is a larger group. On average, this additional mortality adds up to 428 deaths per year from short-term mortality caused by wildfire smoke.

Because mortality is concentrated in the elderly, basing a cost analysis on the total number of deaths might overstate the true social cost of smoke induced air pollution. To address this concern, I use a World Health Organization life table for Mexico to calculate conditional life expectancy at the center of each of the three age bins at 65 years, 75 years and 85 years old. At 65 years old, the conditional life expectancy is 17.95 years, at 75 it is 11.34 years, and at 85 it is 6.39 years. I can use these values together with the previously calculated excess mortality to estimate life years lost. For the period from 2007-2019, I estimate 30,207 life years lost in the age group 60-69, 13,971 life years lost for people 70-79 and 19,656 life years lost in the age group over 80. The total number of life years lost in the period of study adds up to 63,834 life years. The average life years lost per death estimated in this analysis is approximately 10 years. A weakness in this approach is that it assumes that those who died from exposure to wildfire smoke had the average life expectancy of their age group. Considering the results from table 5 by cause of death, it is likely that those who died were weaker and had a lower life expectancy. This is what prior from the U.S. by Deryugina et al. (2019) suggests is happening.

To calculate costs I use the value of statistical life, a tool often used in cost-benefit analysis that assigns a monetary value to life. The Mexican government has previously used USD 1.99 million[13] in cost-benefit analyses. To estimate the welfare costs, I could multiply this number my the excess mortality caused from wildfire smoke, which would yield annual costs of USD 851 million. This would arguably overstate the true social costs considering that mortality is concentrated in the elderly. Because of that, I use the value of USD 1.99 million to derive a dollar value for a statistical life year, which I can then multiply with the values estimated for life years lost in the previous paragraph. To do so, I assume that the value of statistical life can be understood as the value assigned to the median Mexican's life. The median Mexican in the period of analysis is 28 years old and has a conditional life expectancy of 48 years, hence a value of a statistical life year implicit in the USD 1.99 million value of a statistical life can be calculated to be approximately USD 41,562.[14]

The average annual costs calculated using value of statistical life year are then USD 189 million, and total costs from 2007-2019 are USD 2.65 billion. Figure 9 plots the social costs by year calculated using value of statistical life years. 2019 was the year with by far the most wildfire smoke, leading to estimated costs of USD 1 billion in that year alone. Recall the limitations of the previous section, individuals who died from smoke are likely to be weaker, and thus would have had a shorter life expectancy than assumed in this analysis. To be more cautious, one can divide all cost estimates by 10, essentially assuming that the average life expectancy of those who died was not 10, but rather just one year. In that case, the costs from wildfire smoke induced short-term mortality in Mexico in 2019 are still estimated to be USD 100 million.

6 Conclusion

In this paper I study the effect of air pollution on mortality, utilizing plausibly exogenous variation in air pollution caused by the wind direction and fire location and comprehensive administrative mortality records from Mexico. I find that smoke coverage leads to a significant short term increase in ground measured air

[13] 2021 US Dollars.

[14] $VLY = \frac{\$1.99 \text{ million}}{48} = \text{USD } 41,562.$

pollution and has a significant negative impact on mortality rates in age groups 60 years and older, while there is no detectable increase in mortality in younger age groups. Mortality effects is larger in municipalities with lower average household income, and effects are much larger compared to previous estimates from the U.S. There is some suggestion that this heterogeneity in mortality effects is caused by a combination of factors including access to healthcare facilities, development, and baseline health. My results show that there are large costs due to excess mortality from air pollution of approximately USD 189 million per year from increased short-term mortality alone, implying potentially much larger costs when taking into account less severe health outcomes caused my pollution.

My study provides causal evidence on the effect of air pollution and wildfire smoke using nationwide data, and the results provide valuable information to policy makers, who have to balance the costs of fire and pollution mitigation in a world where wildfires are becoming more common and more destructive, and thereby also a growing source of overall air pollution. These estimates are of particular interest to policy makers in developing countries who were previously faced with a large literature focused on the U.S. and a dearth of evidence from settings that are more comparable to their own.

Beyond studying air pollution in a less developed country, this paper also provides estimates of the effect of smoke exposure on mortality for all age groups. Previous studies were sometimes limited to data on elderly and infant mortality, while administrative mortality records allow me to study effects throughout all age groups, showing that mortality effects are mainly concentrated in the elderly. My paper adds to and extends the growing literature on negative health effects of wildfire smoke and highlights the importance of further study of the socioeconomic determinants of differences in mortality effects from air pollution and wildfire smoke.

References

Adhvaryu, A., Bharadwaj, P., Fenske, J., Nyshadham, A., Stanley, R. 2024. Dust and death: Evidence from the west african harmattan. The Economic Journal. 134 (659), 885–912.

Arceo, E., Hanna, R., Oliva, P. 2016. Does the effect of pollution on infant mortality differ between developing and developed countries? evidence from mexico city. The Economic Journal. 126 (591), 257–280.

Arenberg, S., Neller, S. 2023. Ashes to ashes: The lifelong consequences of early-life wildfire exposure.

Chay, K. Y., Greenstone, M. 2003. 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 (3), 1121–1167.

Chay, K., Dobkin, C., Greenstone, M. 2003. The clean air act of 1970 and adult mortality. Journal of Risk and Uncertainty. 27 (3), 279–300.

Currie, J., Neidell, M. 2005. Air pollution and infant health: What can we learn from california's recent experience? The Quarterly Journal of Economics. 120 (3), 1003–1030.

Currie, J., Neidell, M., Schmieder, J. F. 2009. Air pollution and infant health: Lessons from new jersey. Journal of Health Economics. 28 (3), 688–703.

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

Deschênes, O., Greenstone, M. 2011. Climate change, mortality, and adaptation: Evidence from annual fluctuations in weather in the us. American Economic Journal: Applied Economics. 3 (4), 152–185.

Deschênes, O., Greenstone, M., Shapiro, J. S. 2017. Defensive investments and the demand for air quality: Evidence from the nox budget program. American Economic Review. 107 (10), 2958–2989.

Gutierrez, E. 2015. Air quality and infant mortality in mexico: evidence from variation in pollution concentrations caused by the usage of small-scale power plants. Journal of Population Economics. 28 (4), 1181–1207.

Hernández, B., Ramírez-Villalobos, D., Romero, M., Gómez, S., Atkinson, C., Lozano, R. 2011. Assessing quality of medical death certification: Concordance between gold standard diagnosis and underlying cause of death in selected mexican hospitals. Population Health Metrics. 9 (1).

Jayachandran, S. 2009. Air quality and early-life mortality. Journal of Human Resources. 44 (4), 916–954.

Knittel, C. R., Miller, D. L., Sanders, N. J. 2016. Caution, Drivers! Children Present: Traffic, Pollution, and Infant Health. The Review of Economics and Statistics. 98 (2), 350–366.

Mannucci, P., Franchini, M. 2007. Short-term effects of air pollution on cardiovascular diseases: outcomes and mechanisms. Journal of Thrombosis and Haemostasis. 5 (11), 2169–2174.

Mannucci, P., Franchini, M. 2017. Health effects of ambient air pollution in developing countries. International Journal of Environmental Research and Public Health. 14 (9), 1048.

McDuffie, E. E., Martin, R. V., Spadaro, J. V. et al. 2021. Source sector and fuel contributions to ambient pm2.5 and attributable mortality across multiple spatial scales. Nature Communications. 12 (1).

Miller, N., Molitor, D., Zou, E. 2024. The nonlinear effects of air pollution on health: Evidence from wildfire smoke. Working Paper.

Moretti, E., Neidell, M. 2011. Pollution, health, and avoidance behavior: Evidence from the ports of los angeles. The Journal of Human Resources. 46 (1), 154–175.

Rangel, M. A., Vogl, T. S. 2019. Agricultural fires and health at birth. The Review of Economics and Statistics. 101 (4), 616–630.

Sastry, N. 2002. Forest fires, air pollution, and mortality in southeast asia. Demography. 39 (1), 1–23.

Schlenker, W., Walker, W. R. 2015. Airports, air pollution, and contemporaneous health. The Review of Economic Studies. 83 (2), 768–809.

Selimovic, V., Yokelson, R. J., McMeeking, G. R., Coefield, S. 2020. Aerosol mass and optical properties, smoke influence on o3, and high no3 production rates in a western u.s. city impacted by wildfires. Journal of Geophysical Research: Atmospheres. 125 (16).

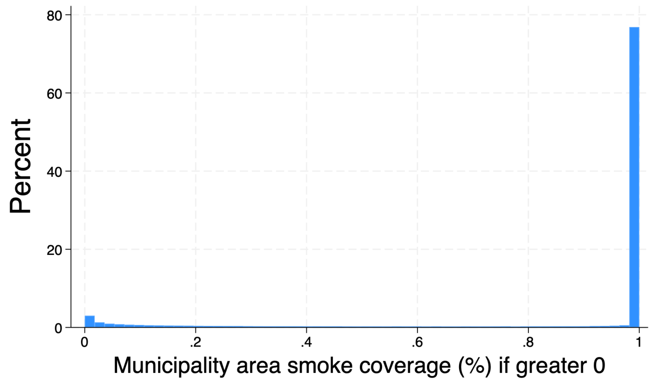


Figure 1

The figure shows the distribution of smoke coverage, conditional on municipality smoke coverage being greater than 0 on that day. Smoke plumes are usually large compared to municipalities, therefore conditional on being covered by any smoke, municipalities are most often entirely covered.

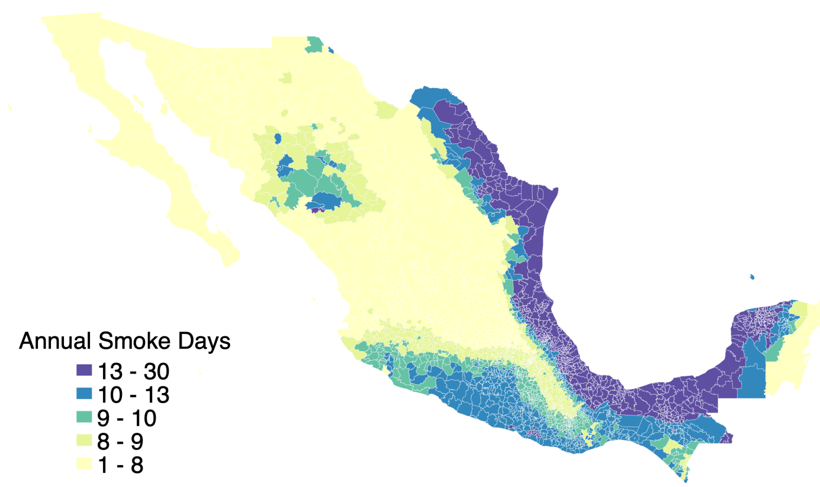


Figure 2: Average Number of Municipality Smoke Days

Figure plots average annual smoke days at the municipality level in the years 2007-2019. Population weighted average smoke coverage per year is 9.64 days.

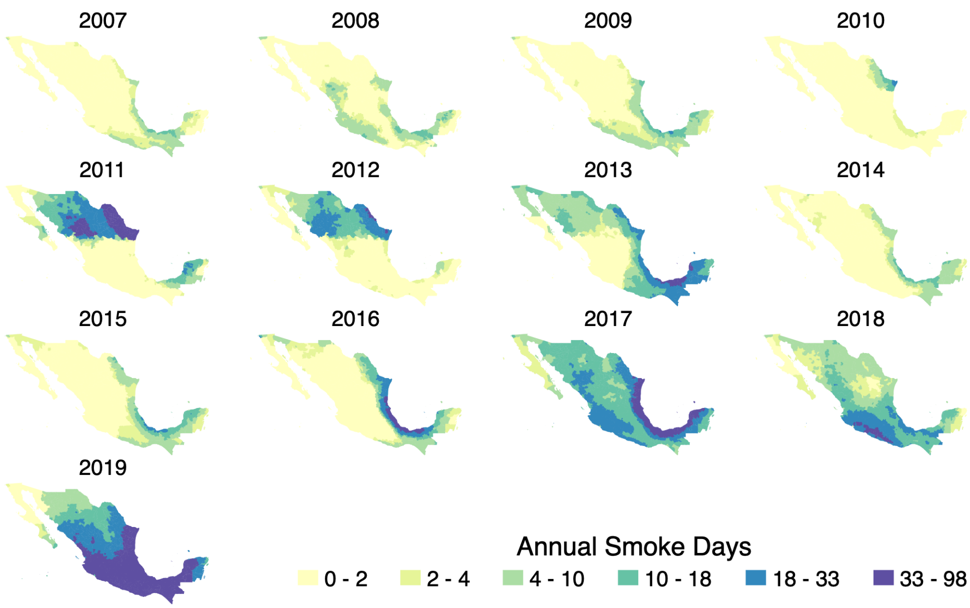


Figure 3: Number of Municipality Smoke Days by Year

Figure plots annual smoke days by year at the municipality level from 2007-2019. Population weighted average smoke coverage per year is 9.64 days.

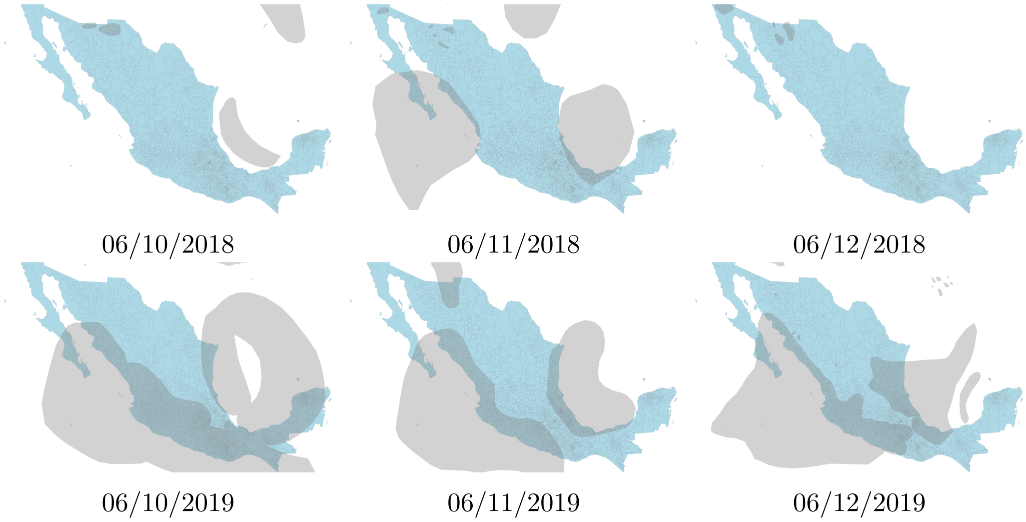
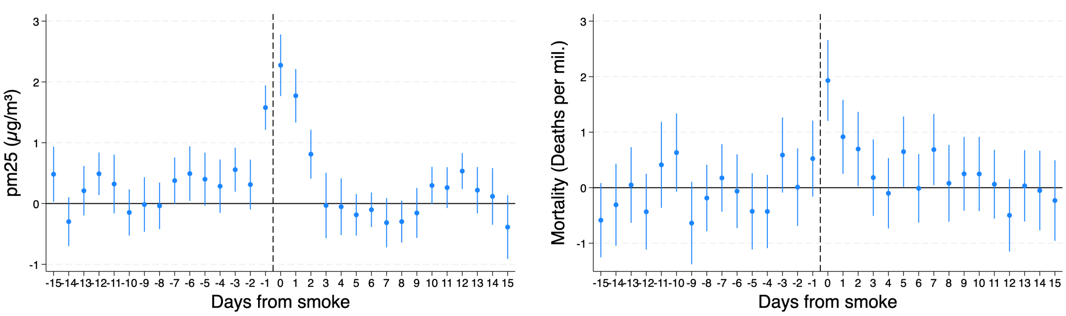


Figure 4: Smoke Cover on Three Consecutive Days in 2018 and 2019

The figure shows smoke plumes over Mexico on three consecutive days in June of 2018 (top panels) and the same three days in June 2019 (bottom panels).



(a) (b)

Figure 5: Effect of Wildfire Smoke on Air Pollution and Mortality

This figure plots coefficients on the leads and lags of smoke of equation 1. Dependent variable is average daily $PM_{2.5}$ concentration (panel (a)) and mortality over 60 per million (panel (b)). The estimate shown for Day 0 is the point estimate reported in tables for "Smoke". Regressions include municipality by week-of-year and date fixed effects and controls for temperature and precipitation. Estimates are weighted by the population over 60. Standard errors are clustered by municipality.

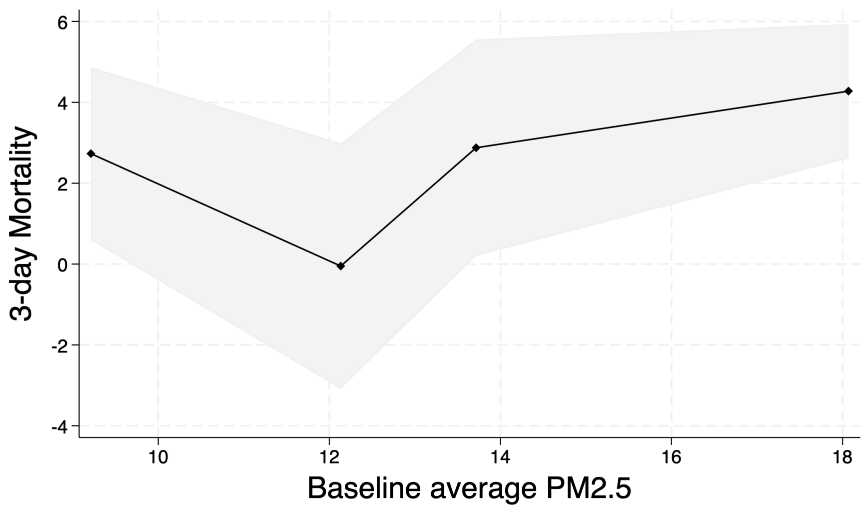


Figure 6: Treatment Effect by Baseline $PM_{2.5}$

Figure plots the 3-day mortality from wildfire smoke by four different quartiles of baseline $PM_{2.5}$ air pollution on days without smoke. Each point estimate is from a separate sub sample regression. Standard errors are shaded grey and clustered by municipality.

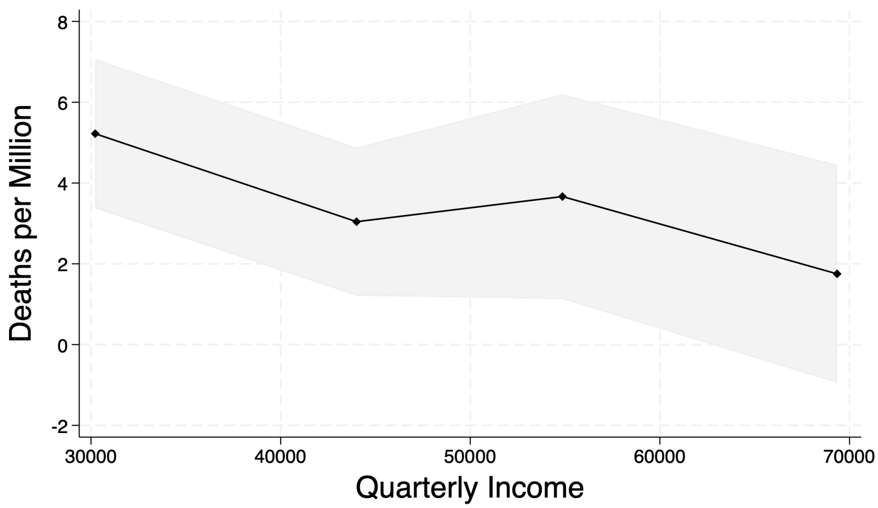


Figure 7: Treatment Effect by Quarterly Income

Figure plots the 3-day mortality from wildfire smoke by four different quartiles of quarterly income. Each point estimate is from a separate sub sample regression. Standard errors are shaded grey and clustered by municipality.

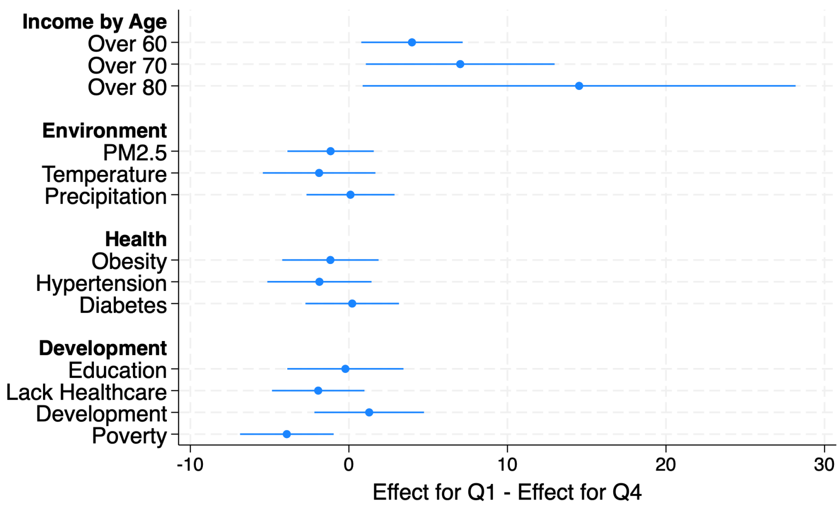


Figure 8: Treatment Heterogeneity by Socioeconomic Indicators

Figure plots the difference in 3-day mortality from wildfire smoke between the first quartile and the last quartile of the category listed on the left side of the figure. Standard errors are clustered by municipality.

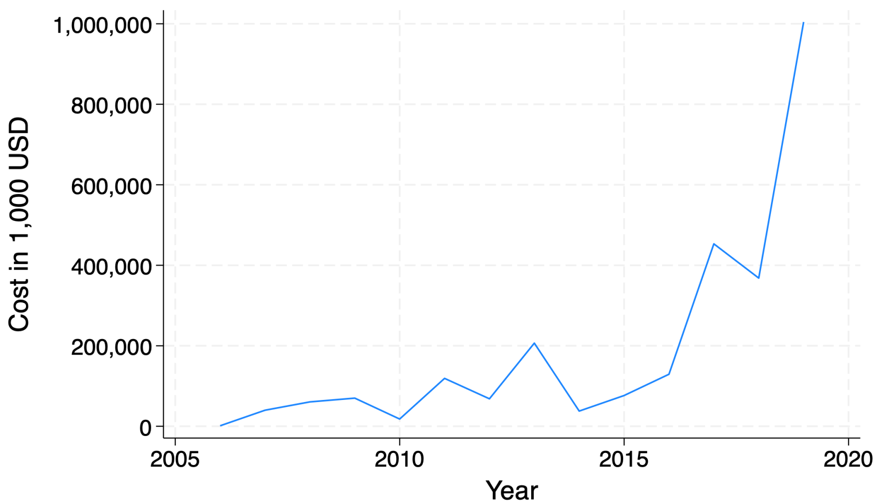


Figure 9: Welfare Cost by Year

The figure plots the annual welfare costs of mortality from wildfire smoke, estimated by using the value of statistical life year method.

Table 1: Summary Statistics

	Mean	Std.	50%	90%
Smoke Days per Year	9.75	58.85	0.00	0.00
Length of smoke event (Days)	3.42	4.71	2.00	8.00
Quarterly Household Income (Pesos)	49,528	15,427	49,851	69,380
PM2.5	21.73	10.78	20.00	35.50
PM10	46.87	24.17	42.50	78.00
Municipality population Age 60-69	2,973.18	9,334.10	799.70	4,826.61
Municipality population Age 70-79	1,592.76	4,751.57	494.05	2,684.52
Municipality population over 80	789.48	2,217.81	273.60	1,361.49
Daily mortality per million over 60	79.83	117.36	62.66	171.34
Daily mortality per million over 70	136.38	218.39	97.79	306.11
Daily mortality per million over 80	244.39	498.87	96.92	593.73

The table presents summary statistics which are calculated at the municipality-day level. All averages are calculated weighting by municipality population over 60, except for the mean of population over 60.

Table 2: Effect of Wildfire Smoke on Air Pollution

	(1) $PM_{2.5}$	(2) PM_{10}	(3) O_3	(4) SO_2
Smoke	2.2745	3.7239	0.0006	0.0001
	(0.2542)	(0.2790)	(0.0002)	(0.0001)
Dep. Var. Mean	21.2995	47.3837	0.0216	0.0042
Municipalities	78	83	91	86
Observations	118,183	188,194	222,834	202,169

Table reports main estimates β_0 of equation 1 for $PM_{2.5}$ and PM_{10}. Standard errors are clustered by municipality and reported in parentheses.

Table 3: Effect of Wildfire Smoke on Mortality by Age

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	0-9	10-19	20-29	30-39	40-49	50-59	60-69	70-79	over 80
Smoke	-0.03	-0.03	-0.07	-0.10	-0.00	0.37	1.25	1.68	4.80
	(0.07)	(0.04)	(0.07)	(0.08)	(0.10)	(0.24)	(0.30)	(0.64)	(1.58)
3 Days: $\sum_{k=0}^{2} \beta_k$	-0.00	-0.03	-0.03	-0.06	0.13	-0.18	1.76	2.40	12.08
	(0.09)	(0.06)	(0.11)	(0.12)	(0.15)	(0.34)	(0.44)	(1.00)	(2.49)
Dep. Var. Mean	4.81	1.63	3.86	5.36	8.77	23.08	36.68	84.91	243.71
Clusters	2457	2457	2457	2457	2457	2457	2457	2457	2457

Table reports main estimates β_0 of equation 1 for age groups of ten years, indicated at the top of the column. Dependent variable is the age group mortality per million. Estimates are weighted by the respective population. Standard errors are clustered by municipality and reported in parentheses.

Table 4: Effect of Wildfire Smoke on Mortality by Income

	Lowest	Second	Third	Highest
Smoke	3.10	1.51	2.22	0.29
	(0.66)	(0.69)	(0.62)	(0.94)
3 Days: $\sum_{k=0}^{2} \beta_k$	5.22	3.04	3.67	1.75
	(0.94)	(0.93)	(1.28)	(1.35)
Dep. Var. Mean	72.08	76.27	84.73	92.53
Municipalities				
Observations	7,758,596	3,046,468	767,652	289,680

Table reports main estimates β_0 of equation 1, with the sample split into four population-weighted quartiles of municipality median income. Dependent variable is the mortality over 60 per million. Estimates are weighted by the population over 60. Standard errors are clustered by municipality and reported in parentheses.

Table 5: Effect of Wildfire Smoke on Mortality by Cause

	(1) Circulatory	(2) Respiratory	(3) Other internal	(4) External
Smoke	0.74	0.07	1.08	0.04
	(0.21)	(0.11)	(0.26)	(0.06)
3 Days: $\sum_{k=0}^{2} \beta_k$	1.92	0.08	1.60	-0.06
	(0.30)	(0.18)	(0.39)	(0.09)
Dep. Var. Mean	26.31	8.87	43.85	2.51
Municipalities	2,457	2,457	2,457	2,457
Observations	11,886,966	11,886,966	11,886,966	11,886,966

Table reports main estimates β_0 of equation 1 for different listed causes of mortality. Dependent variable is the mortality over 60 per million by the cause specified at the top of the column. Estimates are weighted by the population over 60. Standard errors are clustered by municipality and reported in parentheses.

Table 6: Robustness of the Effect of Wildfire Smoke on Mortality

	(1)	(2)	(3)	(4)
Smoke	1.93	2.06	1.72	2.23
	(0.37)	(0.37)	(0.37)	(0.65)
Temperature	0.14			-0.02
	(0.04)			(0.04)
Precipitation	-0.16		-0.15	-0.16
	(0.01)		(0.01)	(0.02)
3 Days: $\sum_{k=0}^{2} \beta_k$	3.54	3.77	3.11	5.67
	(0.56)	(0.56)	(0.56)	(1.01)
Dep. Var. Mean	81.54	81.54	81.54	82.26
Municipalities	2,457	2,457	2,457	2,457
Observations	11,886,966	11,886,966	11,886,966	8,948,394

Table reports estimates β_0 for different specifications. Column (1) is the main specification from 1. Column (2) omits weather controls, column (3) controls for temperature bins, and column (4) excludes the three months with the highest agricultural burn activity in Mexico. Standard errors are clustered by municipality and reported in parentheses.

Appendix

Table 7: Robustness of Results to More and Fewer Leads and Lags

	(1)	(2)	(3)	(4)
	0	7	15	20
Smoke	2.56	1.87	1.93	1.96
	(0.39)	(0.38)	(0.37)	(0.37)
3 Days: $\sum_{k=0}^{2} \beta_k$		3.41	3.54	3.60
		(0.59)	(0.56)	(0.56)
Dep. Var. Mean	81.49	81.53	81.54	81.55
Municipalities	2,457	2,457	2,457	2,457
Observations	11,960,676	11,926,278	11,886,966	11,862,396

Table reports main estimates β_0 of equations that are like 1, but with different numbers of leads and lags of smoke, indicated at the top of the column. Note that sample sizes are slightly different because observations from the beginning and end of the sample need to be excluded because of missing lead/lag observations. Dependent variable is the mortality over 60 per million. Estimates are weighted by the population over 60. Standard errors are clustered by municipality and reported in parentheses.

Table 8: Robustness of Results to Different Fixed Effect Specifications

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Smoke	2.26	2.28	2.16	1.93	2.26	1.79	1.81
	(0.34)	(0.31)	(0.41)	(0.37)	(0.32)	(0.34)	(0.40)
Municipality FE	No	No	No	No	Yes	No	No
Year FE	No	No	No	No	Yes	Yes	No
Week-of-year FE	No	No	No	No	Yes	No	No
Day-of-week FE	No	No	No	No	Yes	Yes	No
Date FE	No	No	Yes	Yes	No	No	Yes
Municipality x Week-of-year FE	No	Yes	No	Yes	No	No	No
Municipality x Day-of-year FE	No	No	No	No	No	Yes	Yes
3 Days: $\sum_{k=0}^{2} \beta_k$	4.95	5.02	4.20	3.54	4.75	3.80	3.45
	(0.58)	(0.42)	(0.74)	(0.56)	(0.43)	(0.43)	(0.59)
Dep. Var. Mean	81.54	81.54	81.54	81.54	81.55	81.55	81.55
Municipalities	2,457	2,457	2,457	2,457	2,457	2,457	2,457
Observations	11,886,966	11,886,966	11,886,966	11,886,966	11,862,396	11,862,396	11,862,396

Table reports main estimates β_0 of equation 1 for specifications with different fixed effects (Column (4) is the main specification). Dependent variable is the mortality over 60 per million. Estimates are weighted by the population over 60. Standard errors are clustered by municipality and reported in parentheses.