THREE ESSAYS ON EXTREME HEAT, WILDFIRES, AND AIR POLLUTION IN THE UNITED STATES

By

Zeying Huang

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Agricultural, Food and Resource Economics – Doctor of Philosophy
Economics – Dual Major

2022

ABSTRACT

Over the past several decades, more frequent and intense extreme heat events have been an increasing threat to human health and economic performance. These extreme heat events result in more fatalities than all other types of extreme weather events, as well as a series of clinical syndromes and chronic diseases, which may expose those with underlying health problems to higher mortality risks. At the same time, extreme heat increases wildfire risks. In recent decades, the United States (U.S.) has experienced upward trends in total acreages burned by wildfires and the average size of wildfires. These trends are expected to grow with the changing climate and as more households move to the wildland-urban interface (WUI). Wildfires can lead to direct injuries and fatalities as well as direct damage to properties and infrastructures. Also, wildfires can lead to environmental changes in many ecosystems. The burning of biomass and soil-based organic matter can generate a large amount of haze and smoke composed primarily of fine particulate matter ($PM_{2.5}$), which is another threat in and of itself but also because it exacerbates the impacts of extreme heat, influences economic development patterns, and interferes with the enjoyment of environmental goods and services directly or indirectly.

Based on the evidence of the inter-relatedness of heat and air pollution and the health risks of heat and air pollution, the first chapter provides nationally representative, robust, and precise estimates of the joint impacts of heat and $PM_{2.5}$ on mortality in the U.S. The chapter employs a county-year balanced panel dataset covering 2,992 U.S. counties from 2001 through 2011 and applies a Fixed-effect Poisson model. I correct the endogeneity of $PM_{2.5}$ by applying the control function approach and exploring transboundary externalities of all-source and wildfire-caused $PM_{2.5}$. I find that the heat index and $PM_{2.5}$ are positively and significantly associated with all mortalities. $PM_{2.5}$ is a positive confounder of heat and vice versa. Failure to consider the endogeneity of $PM_{2.5}$ leads to a substantial underestimation of $PM_{2.5}$ risk. The overestimation bias caused by ignoring the potential confounding effect between heat and $PM_{2.5}$ is magnified once the endogeneity of $PM_{2.5}$ is further addressed. My evaluation also offers evidence of the spillover effects of both $PM_{2.5}$ and wildfires.

Wildfires affect human health directly and indirectly via the environmental (dis)amenities induced by wildfires. The second chapter employs the same dataset and further explores the mechanisms by which wildfires affect human health by examining the extent of the direct and indirect health impacts by applying a mediation analysis. In particular, it focuses on the air

pollution ($PM_{2.5}$) channel. It finds that complementary mediations exist for all-cause, respiratory system disease, and circulatory system disease mortality, and the indirect impacts of wildfires through $PM_{2.5}$ account for 58%, 47%, and 21% of the total effects of wildfires, respectively. I do not find evidence of a mediation effect through $PM_{2.5}$ for suicide, but the result suggests a potential delayed direct impact of wildfires on suicide. In addition, the analysis suggests that the spillover effect of wildfires is substantially larger than the local wildfire effect. Although most previous studies assume that wildfires are exogenous, this study finds that failing to consider the causes of wildfires will lead to upwardly biased estimates of health impacts.

The last chapter conducts a non-market valuation of the impact of wildfires and wildfire-induced $PM_{2.5}$ on the housing market by applying a hedonic price model and mediation analysis approach. This study also explores the potential reasons why more people choose to move in or near the wildland-urban interface (WUI). In particular, I examine the degree to which people may underestimate wildfire risks and the tradeoff between the enjoyment of natural resources and increased wildfire risks. I employ a nationwide repeat-sale dataset between 2010 and 2018, which covers 3,945,340 transaction records of 1,886,684 houses. I find that wildfires, especially distant wildfires, have a statistically significant detrimental impact on house prices via emitting $PM_{2.5}$. There are also significant price disparities between houses located upwind and downwind locations of the wildfires, which may be explained by the substitution effect, externality, and the existence of other channels other than air pollution by which wildfires affect house prices. Moreover, the longer the property's adjacent areas remain free of wildfires, and the farther the nearest recent wildfire, the higher the property's sale price. While I find that households place a higher value on homes in locations with more greenery, they are also aware of the dangers of living near a wildland-urban interface.

This thesis is dedicated to my parents and grandparents.
Thank you for always being there for me.

ACKNOWLEDGEMENTS

Foremost, I would like to express my deepest appreciation to my advisor and chair of my committee, Professor Mark Skidmore, for all the invaluable advice, continuous support, and encouragement through all the stages of writing my dissertation. I also would like to thank the rest of my committee: Professor Scott Loveridge, Professor Frank Lupi, and Professor Ajin Lee, who generously provide insightful comments and suggestions. My sincere thanks also go to Professor Felica Wu for offering the research assistant opportunities and guiding my work on exciting research projects.

My completion of this dissertation could not have been accomplished without the support and help from my friends at MSU: Yanan Jia, Ziwei Ye, Yixuan Gao, Pin Lu, and Xuche Gong. I am also grateful to my old friends in China: Qingping Xiong, Yao Xiong, Huijing Luo, Ning Zou, and Zhihong Chen for always accompanying me over the last five years. My appreciation should also go to the faculties and staff at MSU who inspired me and offered kind help throughout my Ph.D. studies.

Last but not the least, I wish to thank my family for always being with me and supporting me throughout my life.

TABLE OF CONTENTS

CHAPTER 1: THE IMPACTS OF HEAT AND AIR POLLUTION ON MORTALITY IN THE UNITED STATES

1. Introduction

In recent decades, the United States (U.S.) has experienced more frequent and intense extreme heat events. There were statistically significant increases in heat wave frequency and season length for most of the 50 largest U.S. cities over the 1961 to 2018 period (U.S. Global Change Research Program Indicator Platform (USGCRP), n.d.). Importantly, exposure to extreme heat events is associated with higher health risks. Extreme heat results in a large number of direct fatalities, with an annual average of 131 direct deaths in the U.S. over the last twenty years (Lim and Skidmore, 2020). Heat can also trigger a series of clinical syndromes and diseases (such as cardiovascular and respiratory diseases) that can lead to premature mortality (Rainham and Smoyer-Tomic, 2003; Kovats and Hajat, 2008; Anderson and Bell, 2009; Knowlton et al., 2009; Lin et al., 2009). However, the mortality risks caused by heat are not clearly reflected in death records because the fatalities are attributed to different diseases but do not indicate that extreme heat may have induced mortality.

Air pollution, especially fine particulate matter ($PM_{2.5}$), is another important factor that can result in severe health problems. While ambient $PM_{2.5}$ concentrations and extreme $PM_{2.5}$ days have seen overall declining trends in the U.S. as a result of the Environmental Protection Agency's (EPA) stringent air quality regulations (Zhang et al., 2017; U.S. EPA, 2020), $PM_{2.5}$ emitted by some specific sources, such as wildfire-caused $PM_{2.5}$, is a growing threat to human health. Respiratory and cardiovascular complications are two of the main detrimental effects exasperated by $PM_{2.5}$.

There is also a potential inter-relatedness between heat and air pollution. Air pollutants, including Particulate Matter (PM), Ozone (O_3), and Greenhouse Gases (GHGs) can lead to temperature change. Meanwhile, the temperature may affect air quality by influencing emission generation, inventories, and dispersion patterns (De Sario et al., 2013; Zhang et al., 2017). Due to the evidence of adverse health outcomes caused by heat and air pollution and the association between them, it is necessary to simultaneously consider the health impacts of heat and air pollution.

Although heat-related and air pollution-related health risks have been widely studied, particularly in epidemiology and economics, literature searches uncovered no study investigating the joint impact of heat and air pollution with consideration of the endogeneity problem of air pollution using a U.S. national-wide county-year panel dataset. Therefore, this paper aims to provide more representative, robust, and precise estimates of the joint mortality impacts of heat and air pollution in the U.S. by combining the advantages of the existing studies from different fields.

This study employs a nationwide county-year balanced panel dataset covering 2,992 U.S. counties in the 48 U.S. contiguous states and Washington, D.C. from 2001 through 2011[1], representing about 98.5% of the U.S. population. I use the heat index and $PM_{2.5}$ to measure heat and air pollution, respectively, based on the literature and data availability. I focus on three important adverse health outcomes of heat and air pollution: all-cause mortality and two cause-specific mortality (mortality caused by respiratory system diseases and mortality caused by circulatory system diseases). The evaluation starts from baseline models in which I estimate the separate impacts of heat and air pollution using the Fixed-effect Poisson estimation model. I then estimate the confounding and interaction effects to show the necessity of considering confounders. Next, to further address the endogeneity concern of air pollution, I apply the control function approach using the imported all-source air pollution and air pollution emitted by lightning-caused wildfires from distant counties as instrumental variables (IVs), which also provides evidence of cross-boundary spillover effects from air pollution and wildfires. Finally, I examine the robustness using different sets of IVs, model specifications, and samples.

As a prelude to the complete analysis, I find that both the heat index and $PM_{2.5}$ have significant positive impacts on all mortality categories in all the model settings I explore. The evaluation confirms that it is necessary to jointly consider the health impacts of heat and air pollution together because $PM_{2.5}$ (heat index) is a positive confounder of heat index ($PM_{2.5}$). Failing to consider the confounding effect of heat and air pollution leads to overestimated heat-related and air pollution-related risks. I also present new evidence that this overestimation bias is more significant after addressing the endogeneity of air pollution. In addition, I show that ignoring the endogeneity of $PM_{2.5}$ will underestimate the health risk of air pollution. Last, there are significant spillover effects of air pollution and wildfires.

[1] I focus on the period between 2001 and 2011 because of the data availability. The heat index data is only available before 2011.

I contribute to the literature in the following ways. First, this study offers a more representative assessment of the joint effect of heat and $PM_{2.5}$ on all-cause and cause-specific mortality across space and over time. The data cover about 98.5% population in the U.S. over 11 years. Moreover, panel data has advantages such as controlling for unobserved time-constant factors and nationwide trends compared with the time-series data. Second, the empirical analysis demonstrates the necessity of simultaneously considering heat and air pollution. This study further documents overestimation bias resulting from the omission of the confounding effects between heat and air pollution after considering the endogeneity of $PM_{2.5}$. Third, the instrumental variable strategy takes account of the spatial dependence of $PM_{2.5}$ and the contribution of wildfires on $PM_{2.5}$, as well as incorporates exogenous sources including wind direction, geographic distance, and the lightning phenomenon. This instrumental variable strategy provides new evidence of cross-boundary spillover effects of air pollution and wildfires. Last, instead of using the traditional method to measure extreme heat weather, I use the heat index, which combines temperature and humidity, to better capture people's physical experience of heat.

The rest of this essay is organized as follows. In the next section, I review the literature and construct the conceptual framework. In section three, the details of the data and methodology are described. In section four, I present and discuss the results. The last section concludes the paper.

2. Conceptual Framework

2.1 Extreme Heat

According to the U.S. Global Change Research Program Indicator Platform, a heat wave is defined as "a period of abnormally hot weather lasting days to weeks and is an indicator that describes trends in multi-day extreme heat events in cities across the United States" (USGCRP, n.d.). One problem with this definition is how "hot" weather should be measured. When studying the impacts of weather, one problem is to what extent weather conditions affect the human body. However, there is no standard formula to quantify this effect. De Freitas et al. (2017) categorized and classified 162 human thermal bioclimatic indices that considered a variety of atmosphere-related variables, rationales, underlying body-atmosphere heat exchange theories, and designs for application. "Heat Index" or "Apparent Temperature (AT)", which reflects thermal discomfort, is one of the indicators widely used to measure heat wave intensity. By the definition of the U.S. National Oceanic and Atmospheric Administration (NOAA) National Weather Service, the heat index measures "how hot it really feels when relative humidity is factored in with the actual air temperature"[2]. The concept of apparent temperature was first put forward for windy and cold situations (temperatures < 0°C) by Steadman in 1971 and then was extended in his following papers for warm and humid situations (temperatures > 25°C) in 1979 and the universal scale in 1984 (Steadman, 1984).

The North America Land Data Assimilation System (NLDAS) data on Centers for Disease Control and Prevention (CDC) WONDER provides the daily air temperatures and heat index data from 1979 through 2011[3], in which the heat index is calculated based on Steadman's 1979 work. After simplification, only the ambient dry bulb temperature (°F) and relative humidity (integer percentage) were included in the heat index equation[4]. The air temperature is a straightforward measurement of "hot"; however, adjusting for humidity is crucial since the high level of humidity influences the sweating process of the human body and thus influences the cooling off of the body face (Barreca, 2012; USGCRP, n.d.). The combination of temperature and humidity can influence the heat balance of human organisms (Heal and Park, 2016). Therefore, another consistent definition of the heat wave is "a period of two or more consecutive days where the daily minimum apparent temperature (actual temperature adjusted for humidity) in a particular city exceeds the 85th percentile of historical July and August temperatures (1981–2010) for that city" (USGCRP, n.d.). The maps of the distributions of the heat index and temperature across U.S. regions in Lim and Skidmore (2020) also show that the temperature alone does not fully explain the risks of heat, but humid regions have relatively higher heat index values than dryer hot regions[5]. The NOAA's National Weather Service provides a heat index chart, which presents the relationship between the temperature and relative humidity and the heat index. When there is a higher level of temperature and relative humidity, the value of the heat index is higher, which means that there is a higher level of heat stress on the human body.

Extreme heat can lead to clinical syndromes including heat stroke, heat exhaustion, heat

[2] U.S. National Oceanic and Atmospheric Administration (NOAA) National Weather Service. Retrieved from https://www.weather.gov/safety/heat-index on November 22, 2019.
[3] The heat index defined by NOAA and heat index data provided by NLDAS are not available for temperature below 80°F (27°C).
[4] North America Land Data Assimilation System (NLDAS) Daily Air Temperatures and Heat Index, 1979 – 2011, on CDC WONDER. Retrieved from https://wonder.cdc.gov/wonder/help/Climate/ta_htindx.PDF on November 22, 2019.
[5] This chart is available at https://www.wrh.noaa.gov/psr/general/safety/heat/heatindex.png.

cramps, and heat syncope (fainting), and exacerbate health issues including cardiovascular and respiratory diseases, thus increasing hospitalizations, emergency department (ED) visits, and mortality (Rainham and Smoyer-Tomic, 2003; Kovats and Hajat, 2008; Anderson and Bell, 2009; Knowlton et al., 2009; Lin et al., 2009; Karlsson and Ziebarth, 2018). In the U.S., extreme heat events result in more fatalities than all other extreme weather events (Luber and McGeehin, 2008; Lim and Skidmore, 2020). Extreme heat events not only result in direct fatalities (Lim and Skidmore, 2020) but also induce all-cause mortality (Deschênes and Greenstone, 2011; Yu et al., 2019) and mortality caused by cardiovascular diseases (Basu and Ostro, 2008; Anderson and Bell, 2009; D'Ippoliti et al., 2010; Deschênes and Greenstone, 2011; Yu et al. 2019). The impacts of excessive heat on deaths caused by respiratory diseases are mixed. Basu and Ostro (2008) did not find a significant association between respiratory mortality (as a primary or a secondary cause of death) and high ambient temperature, while Anderson and Bell (2009) found that respiratory disease fatalities are associated with heat waves, but the estimates were unclear. However, clear evidence of the significant impact of extreme heat on respiratory mortality has been found by D'Ippoliti et al. (2010) and Huang et al. (2010). Deschênes and Greenstone (2011) and Yu et al. (2019) found a U-shaped temperature-mortality relationship using the county-year panel data of the U.S. and China, respectively. Barreca (2012) considered the impacts of temperature and humidity simultaneously and found that both the temperature-mortality and humidity-mortality relationships are U-shaped using monthly county-level data of the U.S.

The relative magnitudes of heat effects on respiratory and cardiovascular mortalities are not clear. D'Ippoliti et al. (2010) found that heat had a greater effect on respiratory mortality than on cardiovascular mortality, a result that is consistent with other studies in Italy and the Netherlands. However, other studies obtained higher estimates for cardiovascular mortality (rate) than respiratory mortality (rate) (Anderson and Bell, 2009; Deschênes and Moretti, 2009). Finally, Huang et al. (2010) did not find a significant difference between the rate ratios (RRs) of cardiovascular and respiratory mortality during a heat-wave period. Within the context of these mixed findings, the present study sheds light on the relative impacts of heat waves on cardiovascular and respiratory mortality.

Extreme heat events pose a significant health threat to people, especially vulnerable groups such as the elderly, the very young, economically disadvantaged groups, the socially isolated, people who are bedridden or are on certain medications, people with inadequate English language skills or lacking access to media, residents in high-crime areas, and people living in mobile homes or rental homes (McGeehin and Mirabelli, 2001; Deschênes and Moretti, 2007; Anderson and Bell, 2009; Huang et al., 2010; Gabriel and Endlicher, 2011; Deschênes and Greenstone, 2011; Chindapol, 2017; Yu et al., 2019; Lim and Skidmore, 2020). The disparities in mortality during heat waves are found based on gender, ethnicity, and race (McGeehin and Mirabelli, 2001; D'Ippoliti et al., 2010; Madrigano et al., 2015), and there is substantial geographical heterogeneity among cities and between the urban areas and the rural or suburban areas (McGeehin and Mirabelli, 2001; Kovats and Hajat, 2008; U.S. EPA, 2008; Deschênes and Moretti, 2009; Buyantuyev and Wu, 2010; D'Ippoliti et al., 2010; Gabriel and Endlicher, 2011; Deschênes and Greenstone, 2011). Other living situations, such as the central air conditioning and other indicators of house quality, and geographic regions are also important in the relationship between weather (such as temperature) and mortality (McGeehin and Mirabelli, 2001; Ren et al., 2008; Anderson and Bell, 2009).

2.2 Air Pollution

Particulate Matter (PM) is a complex mixture of anthropogenic, biogenic, and natural materials, suspended as aerosol particles (mainly consisting of sulfate, nitrate, ammonium, organic carbon, elemental carbon, sea salt, and dust) in the atmosphere, which can be emitted by both anthropogenic and natural sources, such as combustion, evaporation, agricultural activities, and natural processes (Dawson et al., 2014; U.S. EPA, 2019). The health and welfare effects of PM are usually linked with the size of the particles including $PM_{2.5}$, $PM_{10-2.5}$, PM_{10}, and UFPs (U.S. EPA, 2019). $PM_{2.5}$ and PM_{10} are two particulate matters widely discussed in previous papers. The fine particulate matter ($PM_{2.5}$) is particulate matter with a nominal mean aerodynamic diameter generally less than or equal to 2.5 μm and PM_{10} is comprised of both fine and coarse fractions with a nominal mean aerodynamic diameter less than or equal to 10 μm (U.S. EPA, 2019). In this paper, I focus on $PM_{2.5}$.

From 1990 to 2014, the direct emissions of $PM_{2.5}$ remained relatively unchanged but the ambient concentrations of $PM_{2.5}$ experienced a declining trend across much of the U.S. in recent years (U.S. EPA, 2020), and the number of extreme $PM_{2.5}$ days generally decreased from 2000 to 2009 with some fluctuations (Zhang et al., 2017). However, air pollution is still a major threat to human health. Previous studies have documented the adverse impacts of PM exposure on morbidity (such as cardiovascular and respiratory diseases) and premature mortality (Dawson et al., 2014; Khawand, 2015; Zhang et al., 2017; U.S. EPA, 2019). Khawand (2015) found that the additional premature deaths associated with the increased level of $PM_{2.5}$ primarily come from deaths caused by cardiovascular and respiratory diseases for individuals over age 65.

The U.S. EPA Integrated Science Assessment (ISA) for Particulate Matter Report (2019) concluded that short-term and long-term exposure to $PM_{2.5}$ is causally related to cardiovascular effects, whereas the causal relationships of respiratory effects were only likely to exist. The evidence for morbidity provides biological plausibility for cause-specific mortality (such as mortality caused by cardiovascular and respiratory diseases) and ultimately total mortality (U.S. EPA, 2019). Moreover, the impacts of $PM_{2.5}$ short-term exposure on respiratory mortality has a bigger magnitude and confidence interval than that on cardiovascular mortality, while $PM_{2.5}$-related cardiovascular mortality has a more solid biological plausibility according to the evidence of morbidity assessment than $PM_{2.5}$-related respiratory mortality (U.S. EPA, 2020).

The factors that contribute to the heterogeneity of air pollution-related mortality risk estimates include age, gender, race, pre-existing health condition, genetic factors, medication use, smoking status, socioeconomic status (such as educational attainment, income, and occupation), healthcare availability, residential location, diet, etc. (Katsouyanni et al., 2001; Pope et al., 2006; Hoek et al., 2013; Khawand, 2015; Lavaine, 2015; US EPA, 2019). In addition, risk preference can make a difference. For example, risk-averse people may reduce the risks of air pollution by installing air filtration at home, moving to rural or suburban areas, etc. These averting behaviors can partially offset the adverse impacts of air pollution. The evidence of increased health risk of $PM_{2.5}$ has been found in children, nonwhite populations, people with the pre-existing disease (cardiovascular disease, respiratory disease, and obesity), people with specific genetic variants in genes in the glutathione transferase pathway, people of low socioeconomic status (SES), and people who smoke or were former smokers (U.S. EPA, 2019).

2.3 Relationship between Extreme Heat and Air Pollution

Evidence suggests that the frequencies of heat waves and extremely high temperatures are increasing in the United States (Wuebbles et al., 2017). Many scientists think that one reason for the increase in extreme heat events is human activity (Stott et al., 2016; Wuebbles et al., 2017; Hayhoe et al., 2018) such as excess emissions of Greenhouse Gases (GHGs), Ozone and Particulate Matter (PM). The GHGs can trap heat radiating from Earth toward space in the atmosphere and make our planet warmer. Atmospheric PM can influence climate through the absorption and scattering of incoming solar radiation, alterations in terrestrial radiation, effects on the hydrological cycle, and changes in cloud properties (U.S. EPA, 2019; U.S. EPA, 2020). The impact of PM on the temperature is complex since the composition of PM is complex. The scattering PM (such as sulfate and nitrate) can cool the surface below, whereas the absorbing PM may warm or cool the underlying surface depending on the factors such as the altitude of the PM layer relative to cloud cover and the albedo of the surface (U.S. EPA, 2020). For example, certain forms of PM such as black carbon (BC), brown carbon (BrC), or dust can absorb incoming sunlight (Chen et al., 2006; U.S. EPA, 2020). On the other hand, absorbing PM such as organic carbon (OC) and elemental carbon (EC) can diminish the surface albedo of snow and ice and lead to regional and even global warming (Xu et al., 2006; U.S. EPA, 2020).

Temperature variation also leads to variation in air pollution such as ozone and particulate matter (PM), although the association between temperature and PM is complex and less clear than that between temperature and ozone. Heat waves (temperature) have effects on air quality by influencing emission generation, inventories, and dispersion patterns (De Sario et al., 2013; Zhang et al., 2017). Zhang et al. (2017) found that the $PM_{2.5}$ extreme days were highly positively correlated with daily maximum temperature and minimum relative humidity. That is, based on the research discussed above, if the warming trend continues, $PM_{2.5}$ extreme days would be more likely to occur.

The impact of temperature on PM is unclear and complicated. The possible reasons mentioned in previous studies include the diversity of PM components[6] and compensating effects[7] (Jacob and Winner, 2009; Tai et al., 2010; Dawson et al., 2014). However, there is still evidence that shows a significant positive association between PM and temperature. For example, one of the important sources of PM is wildfire, which is associated with warmer temperatures and frequency of drought (or precipitation) (Dawson et al., 2014; U.S. EPA, 2020). More frequent wildfires can significantly increase PM concentration (Jacob and Winner, 2009; Dawson et al., 2014). According to the National Emissions Inventory (NEI), wildfire smoke results in 10% ~ 20% of primary PM emissions in the U.S. per year (U.S. EPA, 2020). Khawand (2015) and Burke et al. (2021)

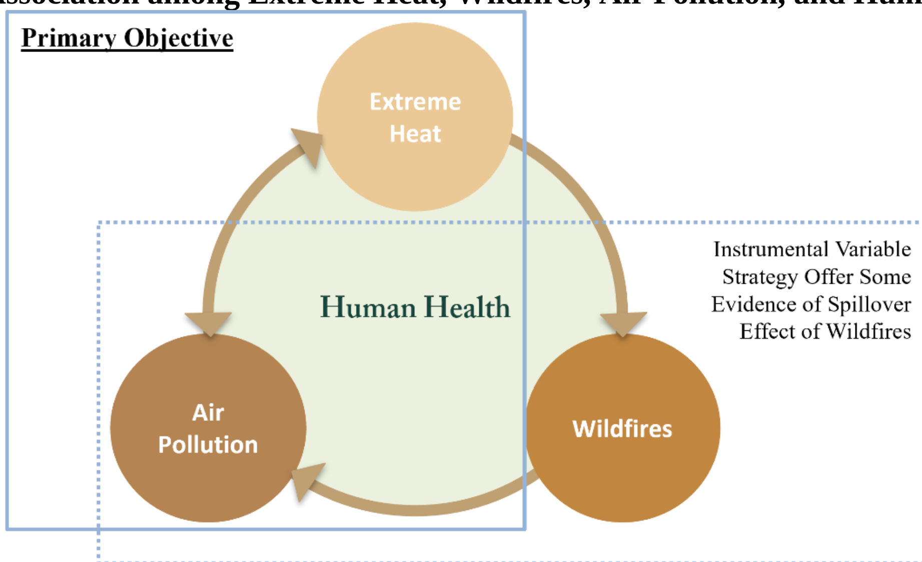
---

[6] The dependence of different $PM_{2.5}$ components on meteorological variables can be complex (Tai et al., 2010).
[7] "The changes in the most relevant meteorological factors for PM such as temperature, precipitation, and mixing will often have competing impacts and these impacts and interactions are difficult to diagnose by focusing on longer-term monthly, seasonal, and annual averages or by grouping various regions or PM species together" (Dawson et al., 2014).

simulated or assessed the air pollution emitted by wildfires, indicating that they contribute roughly 15% and up to 25% of $PM_{2.5}$ in the U.S., respectively. In contrast, Hernandez et al. (2017) found that temperature had a negative correlation with PM_{10} over a diurnal timescale in New Zealand and that the relative humidity generally had a positive correlation with PM_{10} up to a threshold value of 75% relative humidity.

The associations among extreme heat, wildfires, air pollution, and human health are illustrated in **Figure 1. 1**.

Figure 1. 1 Association among Extreme Heat, Wildfires, Air Pollution, and Human Health



2.4 Confounding and Interaction Effects between Heat and Air Pollution

Based on the atmospheric evidence of a potential association between heat and air pollution and the epidemiologic evidence of heat-related and PM-related risks on health, it is important for us to explore the potential confounding effects and interaction effects (effect modification) between heat and air pollution. Ignoring confounding effects can lead to a biased estimation of causal association, and positive (negative) confounding can lead to an overestimate (underestimate) of the effect of interest (Hajian Tilaki, 2012). The terms, interaction effect and effect modification, are often used interchangeably in practice, but their definitions are different (VanderWeele, 2009; Knol et al., 2012). The effect modification is defined as the causal effect of one exposure varying across strata of a second exposure, whereas the interaction effect is defined as the causal effects of two exposures together (VanderWeele, 2009; Knol et al., 2012). In epidemiological studies, the modification effect is more commonly used. In contrast to the confounding effect, failure to consider the modification effect will not result in a distortion of overall effects (Hajian Tilaki, 2012).

Rainham and Smoyer-Tomic (2003) considered the confounding effect of air pollution when studying the heat-mortality effect, arguing that failing to consider the confounding effect of air pollution may slightly overestimate heat-related mortality risk (for all the groups examined except females). Analitis et al. (2014) also found the attenuation of the association between heat waves and mortality if including air pollutants in the model. On the other hand, previous studies also controlled for the potential confounding effect of meteorological variables (such as temperature and relative humidity) when studying the air pollution-related mortality effect (Katsouyanni et al. 2001; U.S. EPA, 2019; U.S. EPA, 2020). Omitting potential confounders was found to lead to overestimates or underestimates of the magnitude of the $PM_{2.5}$ effect (U.S. EPA, 2020). The U.S. EPA ISA concluded that there were consistent positive associations between short-term $PM_{2.5}$ exposures and total mortality as well as the cardiovascular-related emergency department (ED) visits, hospital admissions, and mortality and asthma and chronic obstructive pulmonary disease (COPD) ED visits and hospital admissions across studies that use different approaches to control for the potential confounding effects of weather (e.g., temperature) (U.S. EPA, 2019; U.S. EPA, 2020).

The interaction effect (or modification effect) is widely discussed in studies of air pollution, meteorology, and public health. The modification effect is commonly discussed and is supported by seasonal and geographic differences observed previously (Katsouyanni et al., 2001). Meteorological variables (such as temperature and relative humidity) are regarded as potential effect modifiers when estimating the impact of air pollution on mortality (Katsouyanni et al., 2001; D'Ippoliti et al., 2010). The temperature was found to be one of the best modifiers and played a much more important role than humidity, and the estimated effects of air pollution are larger in warmer arid countries (Katsouyanni et al., 2001). There is some evidence that the impacts of $PM_{2.5}$ on mortality may be modified by temperature, but studies conducted within the U.S. have not provided evidence of the modification effect of temperature, and whether temperature can modify the $PM_{2.5}$-mortality relationship is still unclear (U.S. EPA, 2019). There is inconsistent evidence of whether temperature modifies the associations between short-term exposure to $PM_{2.5}$ and respiratory mortality as well as the associations between short-term exposure to $PM_{2.5}$ and cardiovascular morbidity or mortality (U.S. EPA, 2019). Shaposhnikov et al. (2014) found that the interaction effects of high temperature and air pollution from wildfires in Russia added substantially to deaths. Analitis et al. (2014) also found that the impact of heat waves on mortality was larger when the level of ozone and PM_{10} were higher using the European data.

2.5 Assessments of the Impact of Heat and Air Pollution on Mortality

Although heat-related and air pollution-related health risks have been widely studied, particularly in epidemiology and economics, researchers in different fields explored this topic using different data structures, variables of interest, and analytical models. Many epidemiological studies analyze heat-related or air pollution-related health problems by employing daily time-series data on a specific area or focusing on a specific event (Deschênes, 2014; Chen et al., 2017; Yu et al., 2019; Yang et al., 2021). However, studies using daily time-series data on a specific area are sensitive to the sample population and cannot be generalized across different spatial contexts. Studies on particular events are sensitive to the selection of study periods and cannot be generalized across different temporal contexts. Recent economic studies employ panel datasets with monthly or annual observations covering larger geographical areas and longer timeframes, controlling for fixed effects to mitigate endogeneity problems (Deschênes and Greenstone, 2011; Barreca, 2012; Yu et al., 2019). Karlsson and Ziebarth (2018) found a larger heat-mortality association when using economic models relative to epidemiological models, and they attribute the difference to the control of county-fixed effects that partial out the unobserved time-constant heterogeneity.

The variables of interest and analytical methodologies are also different across fields. Economic studies typically use the mortality rate as the outcome variable and apply parametric OLS modeling with temporal and spatial fixed effects, whereas epidemiological studies typically use mortality counts as the outcome variable and apply log-linear Poisson models with seasonal effects as smooth spline functions (Karlsson and Ziebarth, 2018; Yu et al., 2019). Further, most discussions of the confounding and interaction effects between heat and air pollution exposures are included in these studies. Although many studies on air pollution control for the weather, few studies on the effects of weather consider the role of air pollution (Rainham and Smoyer-Tomic, 2003). Furthermore, epidemiological studies seldom consider the endogeneity and spillover effects of air pollution. Economic studies often take advantage of panel data structures and apply empirical strategies to mitigate endogeneity problems but may omit the important confounding effect between heat and air pollution, particularly in heat-risk studies. Also, although the transboundary externality of air pollution has been widely studied in developing countries, exploration of this issue in the context of the U.S. is limited.

3. Data and Methods

The evaluation begins with baseline models in which I estimate the separate impacts of heat and air pollution using the Fixed-effect Poisson estimation model. I then estimate the confounding and interaction effects to show the necessity of considering confounders. To further address the endogeneity concern of air pollution, I apply the control function approach using the imported all-source air pollution and air pollution emitted by lightning-caused wildfires from distant counties as instrumental variables, which provides evidence of cross-boundary spillover effects from air pollution and wildfires. Finally, I examine the robustness of the results using different sets of IVs, model specifications, and samples. Below, I introduce the data and present more detailed discussions of the models used in each step.

3.1 Baseline Models

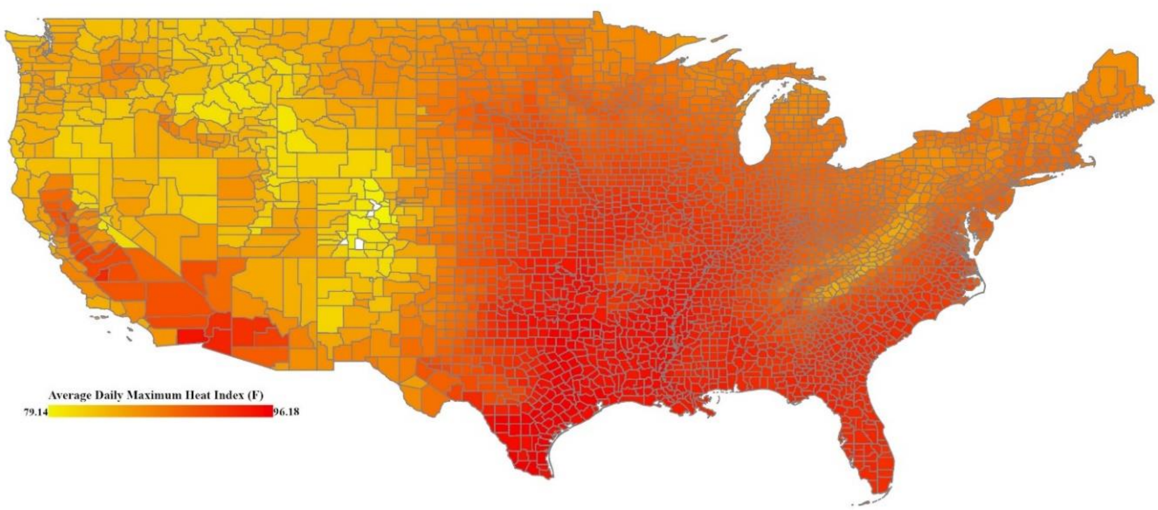
I first estimate the separate impacts of heat and air pollution. For the dependent variable, I focus on all-cause deaths, deaths caused by respiratory system disease (International Classification of Diseases, Revision 10 (ICD-10) code: J00-J98), and deaths caused by circulatory system disease (ICD-10 code: I00- I99). County-level data for annual mortality are collected from the Center for Disease Control and Prevention (CDC) WONDER online database. Since the outcome variables are measured as the number of annual deaths in each county, the priority is to employ count data models. Based on the distributions of death data, which have characteristics of non-negativity, discreteness, and left skewness, I employ a conditional Fixed-effect Poisson quasi-maximum likelihood model. This model is robust to arbitrary misspecified distribution and any serial correlation if the conditional mean is correctly specified (Wooldridge, 2010; Cameron and Trivedi, 2013). The regression equations for heat and air pollution impacts are shown below.

$$E(Deaths_{it}|HI_{it}, X_{it}, C_i, T_t) = C_i \cdot \exp(\alpha_B HI_{it} + \boldsymbol{\alpha} \cdot \boldsymbol{X_{it}} + T_t) \quad (1)$$

$$E(Deaths_{it}|PM_{it}, X_{it}, C_i, T_t) = C_i \cdot \exp(\beta_B PM_{it} + \boldsymbol{\beta} \cdot \boldsymbol{X_{it}} + T_t) \quad (2)$$

The heat index incorporates temperature and humidity, which can better characterize the physiological experience of heat on the human body (Kovats and Hajat, 2008; Lim and Skidmore, 2020). Therefore, this paper uses the average of the daily maximum heat index, HI_{it}, to measure heat for the baseline analysis. The county-level data for the heat index are available at the CDC WONDER online database and are initially from the North America Land Data Assimilation System (NLDAS)[8]. **Figure 1. 2** provides the distribution of the average daily maximum heat index at the county level across the U.S. from 2001 to 2011. The heat indices are generally higher in southern areas and present some geographical variation among counties located at similar latitudes. To examine robustness, I use the number of days with a daily maximum air temperature of 90˚F or more ($Temp90_{it}$) and the average daily precipitation ($Precip_{it}$) to substitute for the heat index. The temperature and precipitation data are also obtained from the CDC WONDER system and are initially from the NLDAS.

Figure 1. 2 Average Daily Maximum Heat Index (F) (2001-2011)



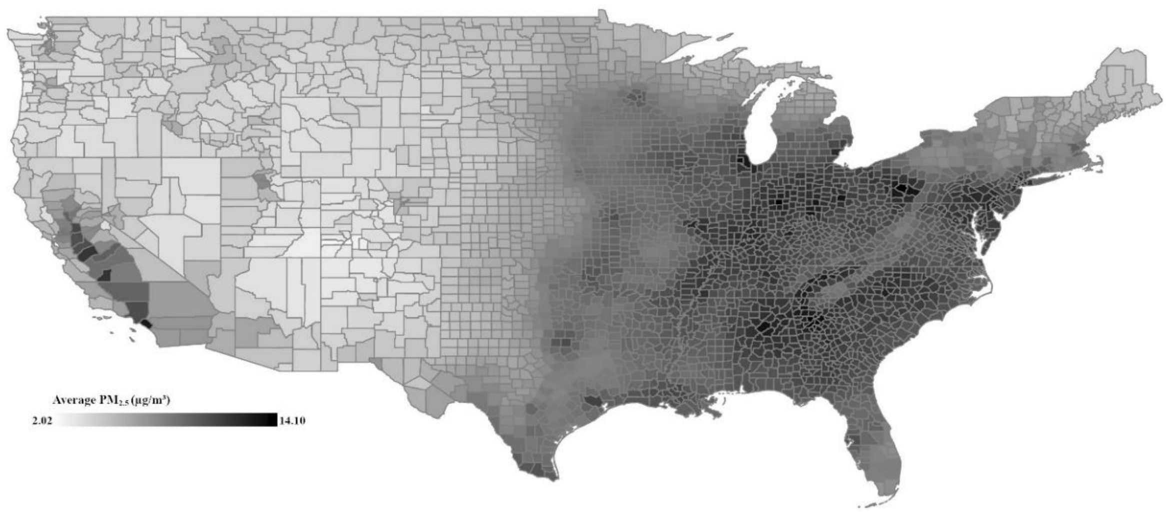
Source: Authors' illustration[9]. **Data:** NLDAS Daily Heat Index, CDC WONDER.

PM is one of the six major pollutants identified by the U.S. EPA. According to the U.S. EPA ISA for Particulate Matter Report (2019), among the various size fractions of PM (such as $PM_{10-2.5}$), the causal relationships between health effects and $PM_{2.5}$ are relatively more likely to exist. As a result, based on the literature and data availability, I use the average level of $PM_{2.5}$ (PM_{it}) to

[8] The heat index defined by NOAA and heat index data provided by NLDAS are not available for temperature below 80°F (27°C).
[9] All the maps in this dissertation uses 2010 U.S. Cartographic Boundary File downloaded from the U.S. Census Bureau (https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html). The comparison of the TIGER/Line Shapefile and Cartographic Boundary File can be found at the U.S. Census Bureau.

measure air pollution. The annual ground-level $PM_{2.5}$ data are from the Atmospheric Composition Analysis Group (ACAG)[10]. County-level mean $PM_{2.5}$ data are obtained by calculating the zonal statistics using ArcGIS Pro. **Figure 1. 3** presents the distribution of the average $PM_{2.5}$ in the U.S. from 2001 to 2011. Generally, the distribution is consistent with the U.S. EPA ISA report (2019, 2020) that the eastern areas of the country suffer a higher but more uniform level of $PM_{2.5}$ than western areas, whereas California has a significantly higher level of $PM_{2.5}$ than surrounding states[11]. The ISA concluded that using different methods to estimate $PM_{2.5}$ exposures did not affect the robustness of the positive associations between long-term $PM_{2.5}$ exposures and mortality in recent analyses (U.S. EPA, 2020).

Figure 1. 3 Distribution of Average Ground-level $PM_{2.5}$ ($\mu g/m^3$) (2001-2011)



Source: Authors' illustration. **Data:** North American Regional Estimates for Surface $PM_{2.5}$ (V4.NA.03), ACAG.

The covariates X_{it} include demographic characteristics (population size, percentage of people over 64, percentage of people under 20, and the white population percentage), urbanization

[10] Surface $PM_{2.5}$ dataset (North American Regional Estimates (V4.NA.03)) from Atmospheric Composition Analysis Group. Downloaded from https://sites.wustl.edu/acag/datasets/surface-pm2-5/#V4.NA.02.MAPLE on August 18, 2020. The shapefile of U.S. counties was downloaded from https://www2.census.gov/geo/tiger/TIGER2010/COUNTY/2010/.

[11] One possible explanation for the significantly higher level of air pollutants in the eastern areas is that there are higher levels of pollen in the wetter (greener) eastern half of the U.S. In addition, the pollen can also be a health threat. However, because the pollen has a larger size than $PM_{2.5}$ and it is not a part of $PM_{2.5}$. I do not consider the effect of pollen in this dissertation. Please find the reference from https://airquality.climate.ncsu.edu/2021/04/14/pm2-5-and-pollen/.

(percentage of urban population[12]), economic development (real per capita GDP), and health status/behaviors (obesity rates for men and women and smoking prevalence)[13]. In alternative specifications, I introduce the percentage of urban areas and urban population density to measure the degree of urbanization and use the real income per capita[14] and the poverty rate to measure economic development. Demographic variables are from the CDC WONDER database and the U.S. Census Bureau (CB). The data used to generate urbanization variables are from the U.S. CB. Economic data are collected from the U.S. Bureau of Economic Analysis (BEA) and the U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program. Health status/behaviors data are collected from the Institute for Health Metrics and Evaluation (IHME). I also include county-fixed effects and time-fixed effects, which control for the unobservable time-constant county-specific heterogeneity and the time-varying but county-constant factors such as nationwide shocks that may have occurred in a given year, respectively. **Table 1. 1** provides variable definitions and data sources, and **Table 1. 2** presents the summary statistics.

Table 1. 1 List of Variables in the Empirical Analysis

Dependent Variables			**Source**
Health Outcomes	All-cause Deaths	$Death_{All_{it}}$	CDC WONDER
	Deaths Caused by Respiratory System Diseases	$Death_{R_{it}}$	CDC WONDER
	Deaths Caused by Circulatory System Diseases	$Death_{C_{it}}$	CDC WONDER
Explanatory/Control Variables			**Source**
Meteorology	Average Daily Maximum Heat Index (F)	HI_{it}	CDC WONDER
	Number of Days with Daily Max Air Temperature ≥90F	$Temp90_{it}$	CDC WONDER
	Average Daily Precipitation (mm)	$Precip_{it}$	CDC WONDER
Air Pollution	Average Ground-level Particulate Matter ($PM_{2.5}$) (μg/m³)	PM_{it}	ACAG

[12] Yearly data of urban population density, the percentage of urban population, and the percentage of urban areas over the period 1999 to 2010 are obtained by interpolation and extrapolation using the U.S. census data and shapefiles of U.S. urbanized areas and counties for years 2000 and 2010. The shapefiles were downloaded from https://www.census.gov/geographies/mapping-files/2000/geo/carto-boundary-file.html and https://www2.census.gov/geo/tiger/TIGER2010/.
[13] Although I do not have the data about the potential averting behaviors, people's risk preferences should be correlated with other socioeconomic factors, health status, and other health-related behaviors.
[14] Real per capita income (logarithm) that is adjusted using the Bureau of Labor Statistics (BLS) Consumer Price Index retroactive series using current methods (R-CPI-U-RS) and using 1977 as the base year.

Table 1. 1 (cont'd)

Demographics	Population Size (in 10 thousand)	$Population_{it}$	CDC WONDER
	Percentage of the Young (under 20) (%)	$Under20_{it}$	U.S. CB
	Percentage of the Elderly (over 64) (%)	$Over64_{it}$	U.S. CB
	Percentage of the white population (%)	$White_{it}$	U.S. CB
Urbanization	Percentage of Urban Population (%)	$UrbanPopu_{it}$	U.S. CB
	Percentage of Urban areas (%)	$UrbanArea_{it}$	U.S. CB
	Urban Population Density (per 1000 square meters)	$Density_{it}$	U.S. CB
Economics	Real Per Capita GDP (chained 2012 dollars) (logarithm)	GDP_{it}	U.S. BEA
	Real Per Capita Income (dollars, based year:1977) (logarithm)	$Income_{it}$	U.S. BEA, U.S. BLS
	Poverty Rate (%)	$Poverty_{it}$	U.S. CB
Health	Prevalence of Obesity (Female) (%)	${Obesity_F}_{it}$	IHME
	Prevalence of Obesity (Male) (%)	${Obesity_M}_{it}$	IHME
	Prevalence of People Who Currently Smoke (%)	$Smoking_{it}$	IHME
Time FE	Year Indicator Variables	T_t	-
County FE	County Indicator Variables	C_i	-
Instrumental Variables			**Source**
Instruments	$PM_{2.5}$ from Distant Counties (μg/m³)	$DistPM_{it}$	ACAG, U.S. CB, NASA
	$PM_{2.5}$ Attributed to Lightning-caused Wildfires from Distant Counties (μg/m³)	${DistPM_{Fires}}_{it}$	ACAG, FPA FOD, U.S. CB, NASA

Table 1. 2 Summary Statistics

Variables	**Mean (2001~2011)**	**Std. Dev. (2001~2011)**	**Mean (2001)**	**Mean (2011)**
All-cause Deaths	807.25	2183.75	798.12	830.10
Deaths Caused by Respiratory System Diseases	77.78	197.42	75.96	82.12
Deaths Caused by Circulatory System Diseases	278.87	800.31	306.41	258.81
Average Daily Maximum Heat Index (F)	90.14	3.71	89.64	92.24
Number of Days with Daily Max Air Temperature ≥90F	39.80	40.47	28.60	62.30
Average Daily Precipitation (mm)	2.76	1.05	2.55	2.77

Table 1. 2 (cont'd)

Average Ground-level Particulate Matter ($PM_{2.5}$) (μg/m³)	8.72	2.65	9.25	7.95
Population Size	98184.63	311435.60	93784.86	102522.50
Percentage of the Young (under 20) (%)	26.99	3.29	28.15	25.89
Percentage of the Elderly (over 64) (%)	15.27	4.04	14.81	16.10
Percentage of the white population (%)	86.47	15.74	87.19	85.74
Percentage of Urban Population (%)	42.19	30.70	41.52	42.79
Percentage of Urban areas (%)	6.55	15.72	6.36	6.72
Urban Population Density (per 1000 square meters)	0.54	0.64	0.55	0.54
Real Per Capita GDP (chained 2012 dollars)	37956.90	26308.69	34676.84	40234.21
Real Per Capita Income (dollars) (based year:1977)	10150.38	2548.81	9623.59	10930.87
Poverty Rate (%)	15.16	6.09	13.73	17.36
Prevalence of Obesity (Female) (%)	35.74	5.83	31.40	39.13
Prevalence of Obesity (Male) (%)	33.41	4.33	28.81	37.17
Prevalence of People Who Currently Smoke (%)	25.86	4.06	27.03	24.41
$PM_{2.5}$ from Distant Counties (μg/m³)	7.56	4.66	7.66	7.17
$PM_{2.5}$ Attributed to Lightning-caused Wildfires from Distant Counties (μg/m³)	0.0005	0.0016	0.0003	0.0009
Number of Observations	32912	32912	2992	2992

3.2 Confounding and Interaction Effects Models

Because heat and air pollution are associated, and both contribute to the risks of mortality, it is necessary to consider the potential confounding of heat and air pollution to mitigate the bias of the estimators. Thus, I modify the baseline model by including air pollution (heat) as a confounder. To explore the variation of heat (air pollution) risk to people exposed to different levels of air pollution (heat), I modify the model by adding an interaction term for heat and air pollution. The regression equations of confounding and interaction effect models are:

$$E(Deaths_{it}|HI_{it}, PM_{it}, \boldsymbol{X_{it}}, C_i, T_t) = C_i \cdot \exp\big(\gamma_h HI_{it} + \gamma_p PM_{it} + \boldsymbol{\gamma} \cdot \boldsymbol{X_{it}} + T_t\big) \quad (3)$$

$$E(Deaths_{it}|HI_{it}, PM_{it}, \boldsymbol{X_{it}}, C_i, T_t) = C_i \cdot \exp\big(\delta_h HI_{it} + \delta_p PM_{it} + \delta_{hp} HI_{it} \cdot PM_{it} + \boldsymbol{\delta} \cdot \boldsymbol{X_{it}} + T_t\big) \quad (4)$$

By comparing the confounding effect and baseline models, I can determine whether the bias of estimated heat-related (or PM-related) health risk resulting from the omission of potential confounders exists.

3.3 Endogeneity of Air Pollution

When investigating the adverse impacts of $PM_{2.5}$, one of the main concerns is that $PM_{2.5}$ may still be correlated with unobserved factors in the error term even when I include a series of control variables as discussed above and the county and year-fixed effects. For instance, local governments, which enact stringent environmental policies to regulate air pollution, are more likely to provide better access to medical care. Further, the sectoral composition of a local economy may not be fully reflected in economic measures such as GDP. A larger industrial sector is associated with greater emissions of air pollutants, whereas a higher share in the service sector may result in a lower level of air pollutants. Failing to address the endogeneity problem can lead to biased estimates. Therefore, I use the Control Function (CF) approach to address this concern; however, at least one excluded exogenous variable is required (Wooldridge, 2010, 2015).

Previous economic papers put forward a variety of instrumental variables (IVs) to address the endogeneity concern of air pollution. One strategy is to construct measures of imported air pollution from distant areas as instrumental variables, considering the wind pattern and geographic distance as primary factors influencing the cross-boundary transportation of air pollutants (Bayer et al., 2009, Luechinger, 2010; Khawand, 2015; Zheng et al., 2014; Barwick et al., 2018; Yang and Zhang, 2018; Williams and Phaneuf, 2019; Zheng et al., 2019; Chen et al., 2021). Following a similar logic, Tan-Soo (2018) constructed wind- and distance-based fire hotspots as an instrument for air pollution in a study from Indonesia. In the present paper, I further combine and extend these methods and construct the wind-driven, distance-weighted imported all-source $PM_{2.5}$, and $PM_{2.5}$ attributed to lightning-caused wildfires as instrumental variables, which are defined as follows:

$$DistPM_{it} = \sum_{i \neq j} PM_{jt} * I\left(WD_{jt} = GD_{ji}\right) * \frac{1}{d_{ij}^{p}}, \qquad d_{ij}^{p} \geq 100km \quad (5)$$

$$PM_{it} = \theta_l Wildfires_{lightning_{it}} + C_i + T_t + u_{it} \quad (6)$$

$$DistPM_{Fire_{it}} = \sum_{i \neq j} \widehat{PM_{jt}} * I\left(WD_{jt} = GD_{ji}\right) * \frac{1}{d_{ij}^{f}}, \quad 30km \leq d_{ij}^{f} \leq 100km \quad (7)$$

WD_{jt} represents the dominant wind direction(s) in county *j* at year *t*, defined as the most frequent wind direction(s) within 12 months in county *j* at year *t*. GD_{ji} is the geographic direction of the vector from county *j* to county *i*. Both WD_{jt} and GD_{ji} have four categories and are defined by the quadrants, in which the dominant wind direction vector and the vector from county *j* to county *i* fall. $\widehat{PM_{jt}}$ denotes the predicted $PM_{2.5}$ attributed to lightning-caused wildfires. ${Wildfires_{lightning}}_{it}$ denotes the number of occurrences of lightning-caused wildfires with 10 ~ 999 burned acres.

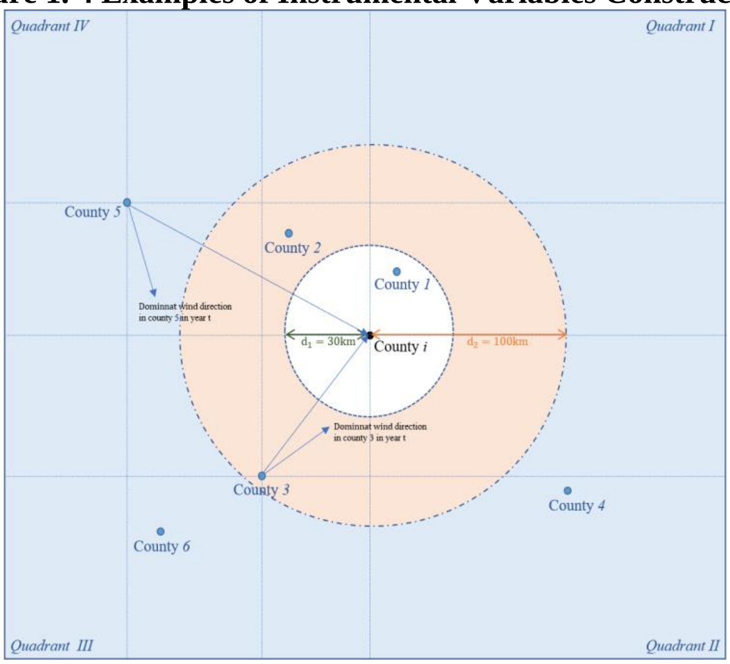
The $PM_{2.5}$ data are also from the ACAG. The wildfire data are obtained from the Fire Program Analysis fire-occurrence database (FPA FOD)[15]. To generate the distance weighting matrix, I use the 2010 TIGER/Line Shapefiles of U.S. counties from the Census Bureau to obtain the distances among counties. To take account of the wind direction effects on air pollution transportation, I collect monthly zonal and meridional wind speed data from Phase 2 of the North American Land Data Assimilation System (NLDAS-2) available from the website of the National Aeronautics and Space Administration (NASA). The county-level monthly mean wind speeds are calculated using ArcGIS Pro.

I consider two types of weights for each equation: wind direction and geographic distance. For the imported all-source $PM_{2.5}$ in county *i* at year *t*, $DistPM_{it}$, I consider the annual mean $PM_{2.5}$ imported from upwind counties that locate at least 100 km away from county *i*, and the $PM_{2.5}$ is weighted by the reciprocal of geographic distance (km), as presented by equation 5. The farther the county is located, the smaller the spillover effect that distant $PM_{2.5}$ has on local $PM_{2.5}$. For the distant $PM_{2.5}$ attributed to lightning-caused wildfires in county *i* at year *t*, ${DistPM_{Fire}}_{it}$, I consider lightning-caused wildfires with 10 ~ 999 burned acres in upwind counties that are located between

[15] The wildfire data are obtained from the Fire Program Analysis fire-occurrence database (FPA FOD) (Short, 2017). This database includes 1,880,465 wildfire events from 1992 to 2015. After excluding the observations for Puerto Rico, Alaska, and Hawaii, I have 1,835,646 observations in total. Since some of the county information (643,450 out of 1,835,646 events) are missing in the database, I map the longitude and latitude of wildfire into the 2010 TIGER/Line Shapefiles of the county from the Census Bureau to obtain the missing county information, and I obtain additional 643,446 county FIPS codes. There are four wildfire events without county information, and thus I drop these four events. Comparing the county FIPS codes in the database to the generated county FIPS codes from shapefile, there are 44,287 county-year fire events where data are unmatched; the overall unmatched rate is 3.71%. Considering that wildfires may occur near the boundary, I treat these 44,287 wildfires as occurring in both counties. Because most of the dates on which the wildfires were declared contained (or controlled) are missing and I use annual wildfire information, I rely on the discovery date of wildfires. Next, I generate the occurrences of wildfires with different sizes and by different causes, and wildfire acres for each county. I assume that there is no wildfire event if there is no record for a specific county and specific year.

30 km and 100 km away from county i. I first predict the $PM_{2.5}$ attributed to lightning-caused wildfires for each county by regressing the local $PM_{2.5}$ on the occurrences of lightning-caused wildfires and the year and county fixed effects, as presented in equation 6. The lightning-caused wildfires can be regarded as an exogenous source of local $PM_{2.5}$. I then use a similar approach to construct the imported $PM_{2.5}$ due to distant lightning-caused wildfires, as presented by Equation 7. **Figure 1. 4** shows the construction of instrumental variables with examples.

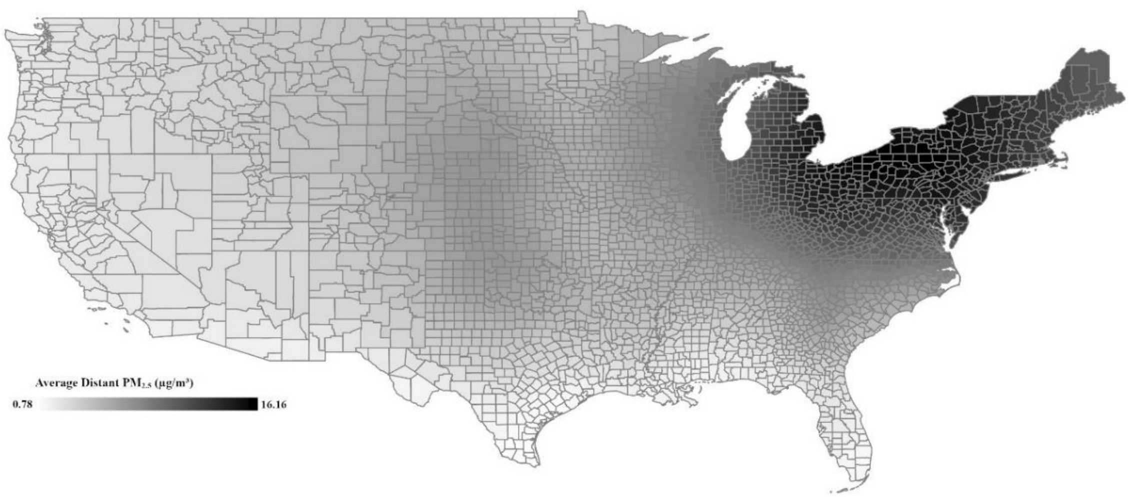
Figure 1. 4 Examples of Instrumental Variables Construction



Note: For county i, the imported all-source $PM_{2.5}$ is from the counties located outside the circle with a radius of 100km, such as counties 4, 5, and 6. For example, the vector from county 5 to county i falls in quadrant II and the dominant wind direction of county 5 in year t falls in quadrant II as well, so I assign the weight of wind direction equal to one (i.e., $I(WD_{5t} = GD_{5i}) = 1$). The impact of imported all-source $PM_{2.5}$ from county 5 is weighted by the reciprocal of the distance between county 5 and county i. The imported $PM_{2.5}$ attributed to lightning-caused wildfires for county i is from counties within a circle of 100km but outside the circle of 30km, such as county 2 and county 3. For county 3, both the vector from county 3 to county i and the dominant wind direction of county 3 fall in quadrant I (i.e., $I(WD_{3t} = GD_{3i}) = 1$). The imported $PM_{2.5}$ attributed to lightning-caused wildfires from county 3 is weighted by the reciprocal of the distance between county 3 and county i.

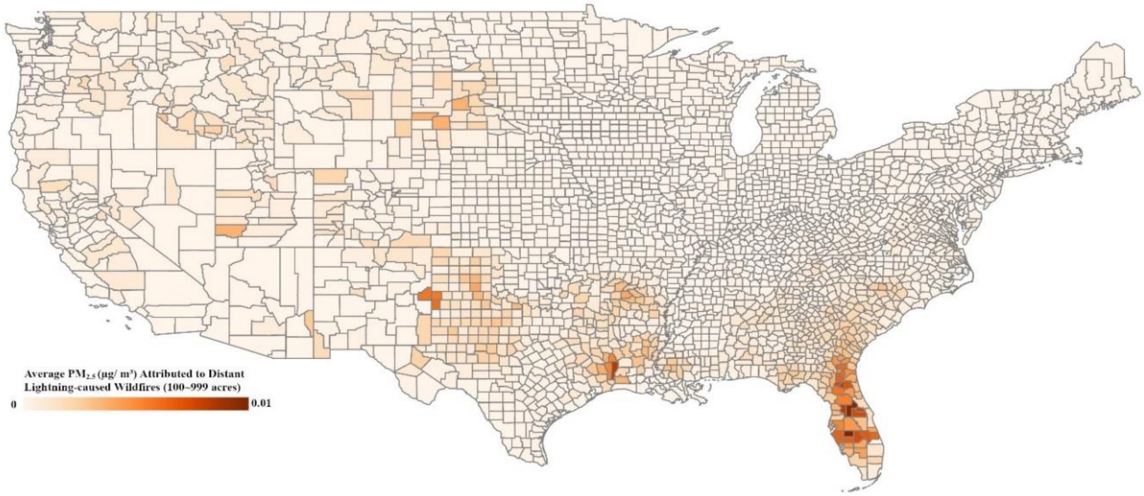
The maps in **Figure 1. 5** and **Figure 1. 6** show the spatial distributions of the average *DistPM* and *DistPM_Fire* from 2001 to 2011. The all-source $PM_{2.5}$ is imported from counties located 100 km away and the $PM_{2.5}$ attributed to lightning-caused wildfires (10 ~999 acres) is from counties located in a range of 30 ~ 100 km away. These two figures highlight the counties with more all-source $PM_{2.5}$ and $PM_{2.5}$ attributed to lightning-caused wildfires from distant counties over the 2001-2011 period. Influenced by the wind direction, the distribution of distant $PM_{2.5}$ shows some differences with the local $PM_{2.5}$, but in general, eastern areas suffered more imported all-source air pollution. The transboundary externalities of $PM_{2.5}$ attributed to lightning-caused wildfires are more significant in the western and southern areas, perhaps because of higher temperatures and more droughts.

Figure 1. 5 Distribution of Average Distant $PM_{2.5}$ (μg/m³) (2001-2011)



Next, I discuss the validity of instrumental variables (IVs). The valid IVs should satisfy two conditions: relevance and exogeneity. First, the concentration level of $PM_{2.5}$ in distant counties should be associated with the local concentration level of $PM_{2.5}$. Previous studies have found evidence of the significant transboundary air pollution spillover effects, and these spillover effects are mainly driven by wind and associated with distance (Bayer et al., 2009; Banzhaf and Chupp, 2010; Luechinger, 2010; Khawand, 2015; Zheng et al., 2014; Barwick et al., 2018; Yang and Zhang, 2018; Chen and Ye, 2019; Williams and Phaneuf, 2019; Zheng et al., 2019; Chen et al., 2021). I also consider the impact of lightning-caused wildfires on air pollution since wildfires are an important source of PM (Khawand, 2015; U.S. EPA, 2020; Burke et al., 2021). Moreover, previous findings show that smoke from smaller-size wildfires can also travel a long distance as large-size wildfires (Miller et al., 2017). Because distant wildfires should affect the air quality in distant counties and air pollutants in distant counties can be transported across counties, I also expect a significant positive association between air pollution attributed to distant wildfires and local air pollution. Through the first-stage regression of the control function approach, I can test the strength of instrumental variables. In section 3.2, I present that the instrumental variables are strongly associated with the local level of $PM_{2.5}$.

Figure 1. 6 Distribution of Average $PM_{2.5}$ ($\mu g/m^3$) Attributed to Distant Lightning-caused Wildfires (2001-2011)

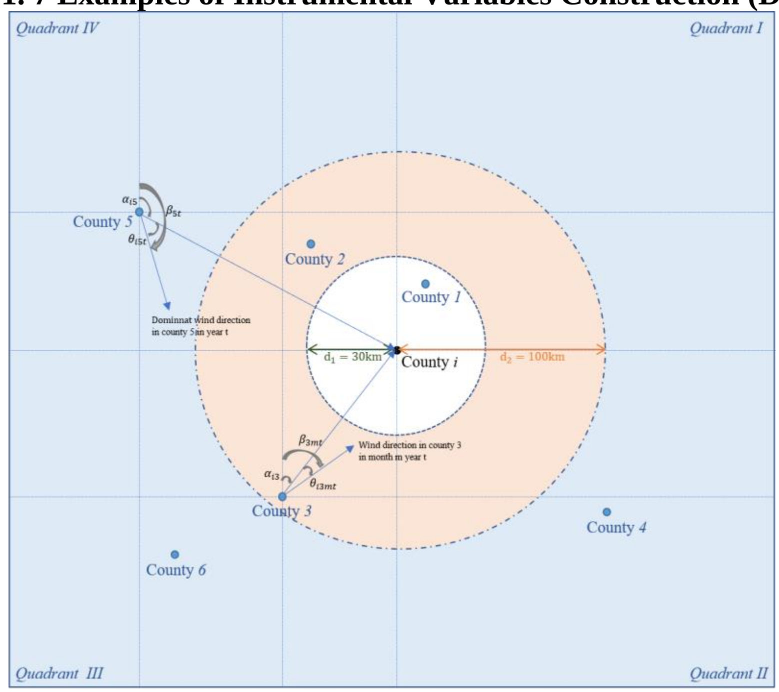


Previous studies also applied wildfire-related IVs. Khawand (2015) simulated $PM_{2.5}$ resulting from large-size wildfires and used it as the instrumental variable to estimate the $PM_{2.5}$-related health impacts and found that wildfires contribute at least 15% of ambient ground-level $PM_{2.5}$. Tan-Soo (2018) constructed wind- and distance-based forest fire hotspots instrument for $PM_{2.5}$. Based on Khawand (2015) and Tan-Soo (2018), I restrict wildfires to be those caused by an exogenous source, lightning, make use of externalities of $PM_{2.5}$ attributed to wildfires with relatively smaller size in counties at a closer distance (30 ~ 100 km), and consider the concern of the health effects of other pollutants (other than $PM_{2.5}$) emitted by wildfires. To evaluate robustness, $PM_{2.5}$ attributed to distant lightning-caused wildfires of 0.26 ~ 299 burned acres (measured using the number of occurrences) and $PM_{2.5}$ attributed to distant all-size lightning-caused wildfires (measured using the total burned acres) are also constructed. In addition, following a logic similar to Tan-Soo (2018), I construct the distant lightning-caused wildfires to substitute the predicted $PM_{2.5}$. The distant lightning-caused wildfires are defined below:

$$DistFires_{it} = \sum_{m=1}^{12} \sum_{i \neq j} Wildfires_{jmt} * I\ (cos\theta_{ijmt} > 0) * \frac{1}{d_{ij}^{f}}, \qquad 30km \leq d_{ij}^{f} \leq 100km$$

Examples of the construction of *DistFire* can be found in **Figure 1. 7**. θ_{ijmt} is the angle between the vector from county *j* to *i* and the vector of wind direction in county *j* in month *m* year *t*. Since the impact of wildfires on air pollution may be temporary, I consider the occurrences and wind direction in each month and sum up the weighted occurrences to generate the annual weighted occurrences of wildfires.

Figure 1. 7 Examples of Instrumental Variables Construction (DistFire)



Second, the instrumental variables should not directly influence the outcome variables. To minimize the likelihood that $PM_{2.5}$ in nearby counties is correlated with the variables influencing the health outcomes in a focal county, I created two buffer zones. For the first buffer zone, following the previous studies that typically considered imported air pollution from counties at a long distance as IVs to ensure the exogeneity condition, I set the radius of the first buffer zone to 100 km and only include the distant all-source $PM_{2.5}$ from counties located outside this buffer zone. Given that I do not make use of the source of $PM_{2.5}$ from counties within 100 km and because $PM_{2.5}$ attributed to lightning-caused wildfires can be regarded as an exogenous source, I consider the $PM_{2.5}$ attributed to lightning-caused wildfires that occurred within the first buffer zone. However, there is still a small possibility that lightning-caused wildfires may lead to accidental deaths. Therefore, I create the second buffer zone, which has a radius of 30 km. I further exclude counties within 30 km, given that I focus on wildfires with less than 1000 acres (about 4.05 km^2). That is, I use the $PM_{2.5}$ attributed to lightning-caused wildfires that occurred within the first buffer zone but outside the second buffer zone as the second IV. Note that most direct fire fatalities (not restricted to wildfires), about 50%-80%, are caused by smoke inhalation and not burns (Holstege, 2019; NFPA, n.d.). Also, according to the NOAA's NCEI Storm Events Database, the average direct injuries and direct fatalities caused by wildfires are about 87 and 8, respectively, per year among the 6,331 wildfires recorded from 1999 to 2017, and the deaths caused by wildfire smoke inhalation are also attributed to direct fatalities in this dataset.

There may also be a concern that other air pollutants emitted by wildfires can be transported by wind to the focal county and affect local health outcomes. Three pollutants (particulate matter, ozone, and carbon monoxide) may pose health threats during wildfire events (Stone et al., 2019). First, $PM_{10-2.5}$ (PM_{10} is comprised of $PM_{2.5}$ and $PM_{10-2.5}$) is not a major concern. Particles from wildfire smoke tend to be very small (with a size range near the wavelength of visible light (0.4-0.7 μm)), and about 90% of total particle masses consist of $PM_{2.5}$ (Stone et al., 2019), so the $PM_{2.5}$ this paper focused on is a major threat to public health (Stone et al., 2019). Second, carbon monoxide dilutes rapidly, so it is rarely a concern unless people are in very close proximity to wildfires (Stone et al., 2019). Therefore, it is improbable that carbon monoxide travels to the focal county and influences local health. Third, ozone is not directly emitted from a wildfire, but forms in the plume as wildfire smoke moves downwind (Stone et al., 2019), so ozone can be another channel through which distant wildfires influence local health. We, therefore, address this problem by constructing the predicted $PM_{2.5}$ emitted by distant lightning-caused wildfires.

In addition, exogenously determined wind direction and geographic distance further increase confidence in the exogeneity of the instrumental variables. Moreover, I consider different radius sizes of the first buffer zone to examine robustness. The radius of the first buffer zone is set to be 80 and 150 km, and then the counties included to construct the two IVs change correspondingly. To increase confidence in these instrumental variables, I follow the methodology described in Wooldridge (2010) and conduct the overidentification test, which is also applied by Wrenn et al. (2017). Section 3.2 presents the overidentification test results, which indicate that I cannot reject the hypothesis that the instrumental variables are exogenous.

Then, I use the control function approach to address the endogeneity problem. In the first stage, I use the fixed-effect model and regress the endogenous variable, air pollution, on instrumental variables, all the other exogenous variables contained in the previous models, and county and time-fixed effects to break the correlation of air pollution and unobservables affecting the dependent variables (Wooldridge, 2010). The first-stage regression equation for the confounding and interaction effect models is[16]:

$$PM_{it} = \theta_0 + \theta_h HI_{it} + \boldsymbol{\theta_{IV} IV_{it}} + \boldsymbol{\theta} \cdot \boldsymbol{X_{it}} + C_i + T_t + \varepsilon_{it} \quad (8)$$

The fixed effect residuals ($\widehat{\varepsilon_{it}}$) of the first-stage regression can be regarded as the "control" for the endogeneity of air pollution (Wooldridge, 2015). I then add $\widehat{\varepsilon_{it}}$ into equations 3 and 4 as the second-stage regressions[17]:

$$E(Deaths_{it} | HI_{it}, PM_{it}, \boldsymbol{X_{it}}, C_i, T_t) = C_i \cdot \exp\left(\gamma_h HI_{it} + \gamma_p PM_{it} + \gamma_r \widehat{\varepsilon_{it}} + \boldsymbol{\gamma} \cdot \boldsymbol{X_{it}} + T_t\right) \quad (9)$$

$$E(Deaths_{it} | HI_{it}, PM_{it}, \boldsymbol{X_{it}}, C_i, T_t) = C_i \cdot \exp\big(\delta_h HI_{it} + \delta_p PM_{it} + \delta_{hp} HI_{it} \cdot PM_{it} + \delta_r \widehat{\varepsilon_{it}} + \boldsymbol{\delta} \cdot \boldsymbol{X_{it}} + T_t\big) \quad (10)$$

Because the residuals depend on the estimates of parameters in the first-stage regression, the variance matrix estimators for parameters in the second-stage regression need to be adjusted to account for this dependence (Wooldridge, 2010). A block-bootstrap procedure (500 repetitions) draws from the entire FIPS code with replacement to correct the standard errors (Wooldridge, 2010; Schlenker and Walker, 2016). I examine the endogeneity of air pollution by testing whether the coefficient of residuals equals zero.

Following the method described in Wooldridge (2010) and applied by Wrenn et al. (2017), I add one of two IVs to the right-hand side of the second stage regression equation (*i.e.*, Equations 9 and 10), and then perform a significance test for this IV. The test result is invariant to which subset of IVs I add. This paper includes distant $PM_{2.5}$ in the second stage regression, as shown in

[16] To estimate the separate impact of air pollution in the baseline model, $Heat_{it}$ in equation 9 is excluded.
[17] Similarly, for the separate impact of air pollution in the baseline model, $Heat_{it}$ in equation 10 is excluded.

Equations 11 and 12. As before, I use the block-bootstrap procedure (500 repetitions) to correct the standard errors.

$$E(Deaths_{it}|HI_{it}, PM_{it}, \boldsymbol{X_{it}}, C_i, T_t) = C_i \cdot \exp\big(\gamma_h' HI_{it} + \gamma_p' PM_{it} + \gamma_{IV}' DistPM_{it} + \gamma_r{}' \widehat{\varepsilon_{it}} + \boldsymbol{\gamma}' \cdot \boldsymbol{X_{it}} + T_t\big) \quad (11)$$

$$E(Deaths_{it}|HI_{it}, PM_{it}, \boldsymbol{X_{it}}, C_i, T_t) = C_i \cdot \exp\big(\delta_h{}' HI_{it} + \delta_p{}' PM_{it} + \delta_{hp}{}' HI_{it} \cdot PM_{it} + \delta_{IV}{}' DistPM_{it} + \delta_r{}' \widehat{\varepsilon_{it}} + \boldsymbol{\delta}' \cdot \boldsymbol{X_{it}} + T_t\big) \quad (12)$$

The empirical strategy I use to address endogeneity also enables us to examine the transboundary spillover effect of air pollution on local health outcomes. If distant $PM_{2.5}$ significantly influences local $PM_{2.5}$ and local $PM_{2.5}$ is significantly associated with health outcomes, then it is reasonable for us to test for the existence of spillover effects. Local $PM_{2.5}$ is a mediator through which distant $PM_{2.5}$ and wildfires indirectly influence local health outcomes. To test for the existence of the spillover effect, I again use the block-bootstrap approach to test the product of the coefficients of distant $PM_{2.5}$ (all-source and from lightning-caused wildfires) in the first stage regression and the coefficient on $PM_{2.5}$ in the second stage regression (i.e., testing $\boldsymbol{\theta_{IV}} * \gamma_p$ for the confounding effect model).

4. Results

The estimation results for heat-related and PM-related fatalities using baseline models, confounding/interaction effect models, and control function approach are shown in **Table 1. 3**. The complete estimation results can be found in **Table 1. 7**, **Table 1. 8**, and **Table 1. 9** of Appendix. Based on these estimates, I calculate the additional deaths due to a one standard deviation (SD) increase in the heat index and $PM_{2.5}$, which are presented in **Table 1. 4**. **Table 1. 5** presents the estimation results for the spillover effects and the estimated extra deaths due to $PM_{2.5}$ and wildfires from distant counties. Finally, I conduct a series of robustness checks.

4.1 Baseline Models and Confounding/Interaction Effect Models

Columns (1) and (2) in **Table 1. 3** show the impacts of the heat index and $PM_{2.5}$ on mortality in the baseline models, respectively. The heat index has a statistically significant and positive relationship with all three mortality categories with a 1% level of significance. The impacts of $PM_{2.5}$ are also statistically significant on mortality for all categories, but with lower 5% and 10%

levels of significance on the deaths caused by respiratory system diseases and circulatory system diseases, respectively.

After controlling the $PM_{2.5}$/heat index as a confounder, the magnitudes of the estimates for the heat index/$PM_{2.5}$ become slightly smaller (column (4) of **Table 1. 3**). These results are generally consistent with the evidence in the existing literature[18] that failing to consider the confounding effect results in biased estimates, and $PM_{2.5}$ (heat index) is a positive confounder of the heat index ($PM_{2.5}$). However, the impacts of $PM_{2.5}$ are less significant, and the impact on deaths caused by circulatory system diseases is no longer significant.

All the interaction terms have negative signs, but they are only statistically significant on all-cause deaths and deaths caused by respiratory system diseases, as shown in column (6), which means that exposure to more severe extreme heat (or higher concentration levels of $PM_{2.5}$) results in reduced risk from $PM_{2.5}$ (or heat index). A possible explanation is the role of avoidance behavior or adaptation; during periods of extreme heat (or with a higher concentration of pollution), people may tend to stay indoors and thus have reduced exposure to particulate matter (or heat).

Overall, the results are consistent with previous literature. Positive and significant associations are observed between heat and mortality and between $PM_{2.5}$ and mortality (except that the impacts of $PM_{2.5}$ are insignificant in the confounding and interaction effect models for deaths in the circulatory system disease category). Moreover, the results confirm that the confounding effects cannot be ignored.

4.2 Endogeneity of Air Pollution

While the estimates in the baseline and confounding/interaction models are useful, previous research suggests the possibility of an existing endogeneity problem of $PM_{2.5}$. Next, I present the estimates using a control function approach to address endogeneity. First, I present evidence of the validity of the IVs. From the first-stage regression results, presented in **Table 1. 10** of the Appendix, I find that the distant all-source $PM_{2.5}$ and distant $PM_{2.5}$ attributed to lightning-caused

[18] Previous studies found that air pollution is a positive confounder of heat; failing to consider the confounding effect of air pollution may overestimate heat-related mortality risk (Rainham and Smoyer-Tomic, 2003; Analitis et al., 2014). On the other hand, previous studies of pollution-related mortality also controlled for the potential confounding effect of meteorological variables such as temperature and humidity (Katsouyanni et al. 2001; U.S. EPA, 2019b; U.S. EPA, 2020). Although omitting potential confounders was found to lead to either an overestimate or underestimate of the $PM_{2.5}$ effect, the adverse impacts of short-term $PM_{2.5}$ exposure on total mortality and some cardiovascular-related and respiratory-related health outcomes are generally consistent across studies that control for the potential weather-related confounding effects (e.g., temperature) (U.S. EPA, 2019b; U.S. EPA, 2020).

wildfires have positive and statistically significant associations with local concentrations of $PM_{2.5}$ at a 1% significance level in all the model settings. The F-statistics of these two IVs for baseline, confounding effect, and interaction effect models are greater than 400, which indicates that the IVs are strong predictors of the endogenous variable, local $PM_{2.5}$. Through overidentification tests, I also find that all the χ^2-statistics are not statistically significant, as shown in **Table 1. 3**, which means that I cannot reject the hypothesis that the IVs are exogenous.

Second, I tested the endogeneity of $PM_{2.5}$. According to the second-stage regression results in columns (3), (5), and (7) of **Table 1. 3**, all the residuals obtained from the first-stage regressions are negative and statistically significant with at least a 5% significance level, which indicates that the unobserved factors that are correlated with $PM_{2.5}$ significantly contribute to reducing the mortality of all categories. We, therefore, reject the hypothesis that $PM_{2.5}$ is exogenous in all the models. This finding also confirms that the endogeneity problem of $PM_{2.5}$ cannot be ignored when studying the adverse health impacts of $PM_{2.5}$.

After addressing the endogeneity issue, I find that the overestimation bias of heat-related risks, which is caused by the ignorance of the confounder $PM_{2.5}$, increases slightly compared to models that fail to consider endogeneity, as shown in columns (1), (4), and (5) of **Table 1. 3**. The $PM_{2.5}$ effects are now statistically significant with at least a 5% significance level in all the three models, and the magnitudes increase substantially, as demonstrated by comparisons of columns (2) and (3), columns (4) and (5), and columns (6) and (7). In addition, the overestimation bias of PM-related risk due to omitting the confounding effect of heat index increases as well, as shown by comparing columns (2) and (4) and columns (3) and (5). The interaction terms of heat and $PM_{2.5}$ are only significant in the model of respiratory system disease deaths.

Overall, the IVs have strong associations with the endogenous variable, $PM_{2.5}$, and I find statistically significant evidence of the endogeneity of $PM_{2.5}$. After addressing the endogeneity of $PM_{2.5}$, the positive association between the heat index and mortality is still consistent, and the positive association between $PM_{2.5}$ and mortality is more significant. Ignoring the endogeneity can lead to a substantial underestimate of the impact of $PM_{2.5}$ on mortality. The results further confirm that $PM_{2.5}$ (heat index) is a positive confounder of heat ($PM_{2.5}$) and the necessity of considering the joint impact of heat and air pollution. I also find that the overestimation bias caused by ignoring the potential confounder is smaller when failing to consider the endogeneity of air pollution. Therefore, the confounding effect model considering the endogeneity of $PM_{2.5}$ is the preferred model setting. If the average daily maximum heat index increases by one unit (F), then the all-cause deaths, deaths caused by respiratory system diseases, and deaths caused by circulatory system diseases will increase by 0.16%, 0.57%, and 0.23%, respectively. Meanwhile, a one-unit increase (μg/m³) of the ground-level $PM_{2.5}$ is associated with 2.11%, 2.55%, and 1.09% addition fatalities on average.

Finally, given that the data covers about 98.5% of the U.S. population, the results can be concluded to represent the population of the whole country. The average of the all-cause deaths, respiratory system disease deaths, and circulatory system disease deaths per year in the U.S. from 1999 to 2019 are around 2,543,934, 247,023, and 850,241, respectively[19]. Based on these data and the estimates of heat and air pollution impacts, I calculate the percentage changes of deaths and the extra additional deaths due to a one SD increase in the heat index and $PM_{2.5}$, as presented in **Table 1. 4**. Based on the results with consideration of the endogeneity of $PM_{2.5}$, I find that a one SD increase in the heat index and $PM_{2.5}$ results in a substantial number of additional deaths, and the bias due to ignorance of confounders is significant. In the confounding effect models, a one SD increase (≈2.71) in the heat index results in 15,136, 5,279, and 7,287 more fatalities from all-cause, respiratory system diseases, and circulatory system diseases, respectively. A one SD increase in $PM_{2.5}$ (≈2.65) will result in 146,302, 17,269, and 24,921 additional fatalities, respectively. The overestimation biases of heat risks are 2,850, 188, and 637 for the above three mortality categories. The overestimation biases of $PM_{2.5}$ fatalities are even larger at 7,123, 2,535, and 3,715, respectively.

Moreover, the literature is unclear about the relative magnitude of heat effects on respiratory and cardiovascular mortalities (Anderson and Bell, 2009; Deschênes and Moretti, 2009; D'Ippoliti et al., 2010; Huang et al., 2010). In this study, the results indicate that the magnitude of the heat-related and $PM_{2.5}$-related deaths (in terms of the percentage change) is greater for respiratory system disease deaths than circulatory system disease deaths.

4.3 Spillover Effect of Air Pollution

As discussed in Section 2.3, the empirical strategy enables us to examine the transboundary spillover effects of $PM_{2.5}$ and wildfires. The distant all-source $PM_{2.5}$ and $PM_{2.5}$ attributed to the

[19] The data is collected from CDC WONDER online database. Retrieved from https://wonder.cdc.gov/controller/datarequest/D76 on February 11, 2021.

lightning-caused wildfires indirectly influence health outcomes in the focal county by influencing local $PM_{2.5}$. The results of the spillover effect are presented in **Table 1. 5**, where I focus on the preferred model setting, the confounding effect model addressing the endogeneity issue. Similarly, I calculate the percentage changes in deaths and the extra deaths due to a one SD increase of distant all-source $PM_{2.5}$ and distant $PM_{2.5}$ attributed to the lightning-caused wildfires. First, the spillover effects are statistically significant with at least a 5% significance level. Second, the magnitudes of spillover effects are substantial. A one SD increase of distant all-cause $PM_{2.5}$ (≈4.66) is associated with 23,811, 2,779, and 3,971 additional all-cause deaths, deaths caused by respiratory system disease, and deaths caused by circulatory system disease, respectively. Note that I still have not included the $PM_{2.5}$ imported from the nearby counties within 100 km. Although I only consider $PM_{2.5}$ attributed to the lightning-caused wildfires with 10 ~ 999 burned acres and within nearby 30 ~ 100 km, the additional deaths caused by a one SD increase of distant $PM_{2.5}$ attributed to lightning-caused wildfires (≈0.0016) are substantial as well, which are 4,019, 472, and 697, respectively.

4.4 Additional Results

In this section, I briefly summarize the estimation results of the control variables. In all the specifications, the percentage of the elderly and the young are positively and statistically significantly correlated with all categories of mortality at a 1% significance level. According to U.S. Census Bureau data, the population of those aged 65 and above has been expanding at an accelerated rate in the United States from 2000 to 2019, while the number of people under the age of 20 has been growing at a slower rate. Therefore, the aging trend will be a growing public health challenge in the coming decades. In addition, the proportion of white people in a given population is often found to be negatively associated with all-cause deaths and deaths due to circulatory system diseases.

Increasing urbanization is positively associated with all mortality categories at a 1% significance level. Local economic development also plays a vital role in reducing fatalities. The logarithm of real per capita GDP is negatively and significantly associated with all-cause deaths and circulatory system disease deaths, especially after addressing endogeneity. People’s health status matters as well. A higher obesity rate is associated with higher mortality, especially for men. The obesity rate among men has a positive and significant association with all three mortality categories. In contrast, the obesity rate of the female group is only significantly associated with all-cause deaths and deaths caused by respiratory system diseases. Surprisingly, I do not find a significant association between smoking prevalence and mortality.

4.5 Robustness Evaluation

In this section, I conduct a series of robustness checks. First, I use different spatial ranges of instrumental variables. In the original analysis, I chose the counties at least 100 km away from the focal county to construct the distant all-source $PM_{2.5}$ and counties within a range of 30-100 km away to construct the distant $PM_{2.5}$ attributed to the lightning-caused wildfires. This section examines the 80 km and 150 km for distant all-source $PM_{2.5}$ and the ranges of 30-80 km and 30-150 km for distant $PM_{2.5}$ attributed to lightning-caused wildfires, respectively. The further the distant sources of $PM_{2.5}$, the more likely it is for the exogeneity condition to hold, but the weaker the association between local and distant $PM_{2.5}$. From the first stage results in **Table 1. 5** of the Appendix, I report that the magnitudes of coefficients for distant all-source $PM_{2.5}$ decrease as the radius of the first buffer zone increases. Meanwhile, as the radius of the first buffer zone increases, the range of the second buffer zone increases, which means that farther lightning-caused wildfires are included, and the magnitudes of the impacts of distant $PM_{2.5}$ that are attributed to the lightning-caused wildfires on the local $PM_{2.5}$ become smaller. However, regardless of which spatial range I choose, the F-statistics for these two IVs are greater than 350, which means that the IVs are strong predictors of local $PM_{2.5}$. As shown in **Table 1. 6**, the χ^2-statistics of the overestimation tests are not significant in all settings, which means that I cannot reject the hypothesis that the IVs constructed using different spatial ranges are exogenous. Further, I find that the estimation results using IVs of different spatial distances change slightly.

Second, I also examine robustness using lightning-caused wildfires of different sizes. In the original analysis, I choose the occurrences of lightning-caused wildfires with 10 ~ 999 acres. This section also considers the burning acres of all lightning-caused wildfires and the occurrences of the lightning-caused wildfires with 0.26~299 acres. From the first stage results in **Table 1. 6** of the Appendix, when larger distant wildfires are included, the impacts on the local $PM_{2.5}$ (columns 1 and 2) increase. The joint tests of IVs are statistically significant, and the F-statistics are over 400 for all the specifications. In addition, the IVs using different sizes of wildfires pass the overidentification test. Overall, the results are robust, and there are only slight variations in the coefficient estimates.

Third, I substitute the distant $PM_{2.5}$ attributed to the lightning-caused wildfires with the occurrences of lightning-caused wildfires in distant counties as the second IV (see **Table 1. 13** and **Table 1. 14** of Appendix). Overall, there are larger variations among the estimates compared with the estimates using distant $PM_{2.5}$ attributed to lightning-caused wildfires. Also, the overidentification test fails when considering all lightning-caused fire acres in the model of all-cause mortality. Therefore, the concern that the distant lightning-caused wildfire may influence local health outcomes by emitting other pollutants is reasonable; thus, the distant $PM_{2.5}$ attributed to the lightning-caused wildfires is the preferred IV.

Fourth, the estimates for heat and $PM_{2.5}$ are robust to a variety of alternative specifications: 1) use of the number of days with a daily maximum air temperature (≥90F) and average daily precipitation to substitute heat index[20]; 2) use of a different economic variable (real per capita income or poverty rate); 3) use of a different variable representing urbanization (percentage of urban area or urban population density) in the preferred model setting. These estimation results are available upon request.

Lastly, I exclude Georgia, New Jersey, and California samples. The mortality data in these states have some unusually high death counts for "Other ill-defined and unspecified causes of mortality" for some years within the study period, which may influence the death counts caused by respiratory system and circulatory system diseases, according to the data notes in the CDC WONDER system. Although there are some variations in the magnitudes, the overall results are robust when excluding these states. These estimation results are also available upon request.

[20] I find positive and significant effects $PM_{2.5}$ (with 1% significance level) and precipitation (with at least 10% significance level) on all the three categories of mortality, but the number of hot days (≥90F) is only significant on deaths caused by respiratory system diseases (with at 1% significance level).

Table 1. 3 Estimation Results

	Heat	Air Pollution		Confounding		Interaction	
	FE Poisson	FE Poisson	CF	FE Poisson	CF	FE Poisson	CF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Dependent Variable: All-cause Deaths						
HI	0.0019^{***}			0.0017^{***}	0.0016^{***}	0.0038^{***}	0.0033^{**}
	(4.14)			(3.65)	(3.28)	(2.79)	(2.38)
PM		0.0055^{***}	0.0221^{***}	0.0054^{***}	0.0211^{***}	0.0254^{**}	0.0366^{***}
		(6.32)	(5.59)	(6.23)	(5.40)	(2.32)	(3.16)
HI*PM						-0.0002^{*}	-0.0002
						(-1.87)	(-1.51)
Residuals			-0.0174^{***}		-0.0164^{***}		-0.0160^{***}
			(-4.73)		(-4.54)		(-4.47)
χ^2 statistics overidentification test			0.35		0.17		0.09
			[0.5546]		[0.6843]		[0.7634]
	Dependent Variable: Respiratory System Disease Deaths						
HI	0.0059^{***}			0.0058^{***}	0.0057^{***}	0.0155^{***}	0.0148^{***}
	(5.81)			(5.74)	(5.28)	(6.18)	(5.98)
PM		0.0032^{**}	0.0291^{***}	0.0028^{*}	0.0255^{***}	0.0963^{***}	0.1118^{***}
		(2.06)	(3.83)	(1.80)	(3.49)	(4.43)	(5.04)
HI*PM						-0.0011^{***}	-0.0010^{***}
						(-4.35)	(-4.11)
Residuals			-0.0269^{***}		-0.0237^{***}		-0.0213^{***}
			(-3.58)		(-3.24)		(-2.92)
χ^2 statistics for overidentification test			0.80		0.28		0.04
			[0.3718]		[0.5965]		[0.8349]
	Dependent Variable: Circulatory System Disease Deaths						
HI	0.0025^{***}			0.0024^{***}	0.0023^{***}	0.0045^{***}	0.0041^{***}
	(3.86)			(3.78)	(3.61)	(2.73)	(2.58)
PM		0.0016^{*}	0.0125^{***}	0.0014	0.0109^{**}	0.0209	0.0276^{**}
		(1.75)	(2.83)	(1.56)	(2.53)	(1.59)	(2.03)
HI*PM						-0.0002	-0.0002
						(-1.52)	(-1.35)
Residuals			-0.0114^{***}		-0.0099^{**}		-0.0095^{**}
			(-2.62)		(-2.34)		(-2.24)
χ^2-statistics for overidentification test			0.68		1.03		1.23
			[0.4094]		[0.3112]		[0.2683]

Table 1.3 (cont'd)

F-statistics for instruments			452.3070 [0.0000]		457.8425 [0.0000]		457.8425 [0.0000]
Obs.	32912	32912	32912	32912	32912	32912	32912

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) Standard errors of regressions in columns 3, 5, and 7 are obtained from the block-bootstrap (500 repetitions) procedure (drawing the entire FIPS code with replacement). (5) Time-fixed effects and county-fixed effects are included in all model specifications.

Table 1. 4 Increases in Deaths Due to A One SD Increase of Heat Index and $PM_{2.5}$

Deaths	Average Deaths (1999~2019)	A one SD (≈3.71) increases in HI			A one SD (≈2.65) increases in $PM_{2.5}$		
		Coefficient	∆Pct.	∆Deaths	Coefficient	∆Pct.	∆Deaths
Baseline models							
All-cause	2,543,934	0.0019***	0.707%	17,986	0.0221***	6.031%	153,425
Respiratory System	247,023	0.0059***	2.213%	5,467	0.0291***	8.017%	19,804
Circulatory System	850,241	0.0025***	0.932%	7,924	0.0125***	3.368%	28,636
Confounding effect models							
All-cause	2,543,934	0.0016***	0.595%	15,136	0.0211***	5.751%	146,302
Respiratory System	247,023	0.0057***	2.137%	5,279	0.0255***	6.991%	17,269
Circulatory System	850,241	0.0023***	0.857%	7,287	0.0109**	2.931%	24,921
Overestimation bias due to omitting confounder (HI or $PM_{2.5}$)							
All-cause	2,543,934	0.0003	0.112%	2,850	0.0010	0.280%	7,123
Respiratory System	247,023	0.0002	0.076%	188	0.0036	1.026%	2,535
Circulatory System	850,241	0.0002	0.075%	637	0.0016	0.437%	3,715

Note: (1) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
(2) All the models, except the baseline models for heat, consider the endogeneity of $PM_{2.5}$.

Table 1. 5 Spillover Effects Due to A One SD Increase of Distant $PM_{2.5}$ and Wildfires

Deaths	Average Deaths (1999~2019)	A one SD (≈4.66) increases in DistPM			A one SD (≈0.0016) increases in DistPM_Fire		
		Coefficient[1]	ΔPct.	Δ Deaths	Coefficient[2]	ΔPct.	ΔDeaths
All-cause	2,543,934	0.0020^{***}	0.936%	23,811	0.9838^{***}	0.158%	4,019
Respiratory System	247,023	0.0024^{***}	1.125%	2,779	1.1915^{***}	0.191%	472
Circulatory System	850,241	0.0010^{**}	0.467%	3,971	0.5101^{**}	0.082%	697

Note: (1) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

(2) Coefficient[1] is the product of the coefficient of *DistPM_Fire* in the first stage and the coefficient of $PM_{2.5}$ in the second stage, and the standard errors are obtained from the block-bootstrap (500 repetitions) procedure (drawing entire FIPS code with replacement).

(3) Coefficient[2] is the product of the coefficient of *DistPM* in the first stage and the coefficient of $PM_{2.5}$ in the second stage, and the standard errors are obtained from the block-bootstrap (500 repetitions) procedure (drawing entire FIPS code with replacement).

Table 1. 6 Estimation Results – Robustness Check

	Main Analysis	Robustness Check			
	***DistPM*(≥100km) *DistPM_Fire* (30~100km, 10 ~ 999 acres)**	***DistPM_Fire* (≥100 acres)**		***DistPM* (≥100km) & *DistPM_Fire* (30~100km)**	
		***DistPM*(≥80km) *DistPM_Fire*(30~80 km)**	***DistPM*(≥150km) *DistPM_Fire*(30~150 km)**	***DistPM_Fire* (0.26~299 acres)**	***DistPM_Fire* (All lightning-caused fire acres)**
Dependent Variable: All-cause Deaths					
HI	0.0016***	0.0016***	0.0016***	0.0016***	0.0016***
	(3.28)	(3.28)	(3.30)	(3.27)	(3.22)
PM	0.0211***	0.0208***	0.0211***	0.0211***	0.0223***
	(5.40)	(5.49)	(5.17)	(5.13)	(5.49)
Residuals	-0.0164***	-0.0161***	-0.0163***	-0.0164***	-0.0176***
	(-4.54)	(-4.62)	(-4.28)	(-4.33)	(-4.70)
χ^2-statistics	0.17	0.08	0.61	0.47	0.19
	[0.6843]	[0.7799]	[0.4330]	[0.4931]	[0.6663]
Dependent Variable: Respiratory System Disease Deaths					
HI	0.0057***	0.0057***	0.0057***	0.0057***	0.0056***
	(5.28)	(5.28)	(5.28)	(5.28)	(5.23)
PM	0.0255***	0.0250***	0.0267***	0.0249***	0.0291***
	(3.49)	(3.49)	(3.45)	(3.27)	(3.73)
Residuals	-0.0237***	-0.0232***	-0.0248***	-0.0230***	-0.0273***
	(-3.24)	(-3.25)	(-3.21)	(-3.04)	(-3.54)
χ^2-statistics	0.28	0.09	0.95	0.99	0.93
	[0.5965]	[0.7692]	[0.3303]	[0.3207]	[0.3346]

Table 1.6 (cont'd)

	Dependent Variable: Circulatory System Disease Deaths				
HI	0.0023***	0.0024***	0.0023***	0.0023***	0.0023***
	(3.61)	(3.61)	(3.60)	(3.62)	(3.61)
PM	0.0109**	0.0103**	0.0120**	0.0104**	0.0106**
	(2.53)	(2.48)	(2.55)	(2.20)	(2.35)
Residuals	-0.0099**	-0.0093**	-0.0110**	-0.0093**	-0.0096**
	(-2.34)	(-2.27)	(-2.36)	(-2.05)	(-2.15)
χ^2-statistics	1.03	1.00	0.65	1.48	1.11
	[0.3112]	[0.3171]	[0.4201]	[0.2238]	[0.2920]
F-statistics for IVs	457.8425	496.5736	386.0210	455.1311	441.9121
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Obs.	32912	32912	32912	32912	32912

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) Standard errors are obtained from the block-bootstrap (500 repetitions) procedure (drawing the entire FIPS code with replacement). (5) Time-fixed effects and county-fixed effects are included for all the model specifications. (6) The χ^2-statistics of overidentification tests are based on Wooldridge (2010), and the instrumental variable for distant $PM_{2.5}$ attributed to lightning-caused wildfires is excluded.

5. Conclusion

With anticipated more frequent extreme heat events in the U.S., people are exposed to more severe heat-related health consequences. Although the ambient concentrations of $PM_{2.5}$ exhibited a declining trend across most of the U.S., the wildfire-induced $PM_{2.5}$ became a new challenge, and $PM_{2.5}$, as an important confounder of heat, is still a significant health threat. Furthermore, air pollution (all-source and that emitted from wildfires) can be carried by the wind across boundaries and then influence distant areas. Although some papers examined the transboundary spillover effects of air pollution in developing countries, this issue has not been fully explored in the U.S. Building on the existing studies, I employ a nationwide panel dataset of the U.S. to explore the joint impacts of heat and $PM_{2.5}$ on the all-cause and cause-specific mortality with considerations of the endogeneity problem and spillover effects of $PM_{2.5}$.

Consistent with previous literature, this paper confirms that extreme heat and $PM_{2.5}$ are significantly and positively associated with all-cause deaths, deaths caused by respiratory system diseases, and deaths caused by circulatory system diseases. Failure to consider the confounding effect between heat and air pollution results in overestimates of the health impacts of heat and $PM_{2.5}$. In addition, this overestimation bias is magnified after addressing the endogeneity problem. Further, ignoring the endogeneity of $PM_{2.5}$ leads to a substantial underestimation of the magnitude of $PM_{2.5}$-related risk. Lastly, the instrumental variable strategy considers the externalities of $PM_{2.5}$ and wildfires, which offers new evidence of the existence of transboundary spillover effects of $PM_{2.5}$ and lightning-caused wildfires. With a growing trend of extreme heat events and wildfires, wildfire-induced air pollution is an important environmental disamenity, especially for the western areas in the U.S. This paper increases the understanding of the transboundary spillover effect of wildfire-induced air pollution. Overall, the evaluation confirms and builds on previous analyses, offering refined nationwide estimates of the impact of heat and air pollution on mortality.

Further, combing the estimates of the additional fatalities caused by extreme heat and $PM_{2.5}$, and the estimate of the "value of a statistical life (VSL)" for the United States, I calculate the economic cost of these two environmental problems. The estimates of the "value of a statistical life" can be used to measure the willingness to pay to achieve the risk reduction of 1.0 premature death (Freeman III et al., 2014). Currently, the U.S. EPA recommends using the central estimate of $7.4 million (2006 dollars) "in all benefits analyses that seek to quantify mortality risk reduction benefits regardless of the age, income, or other population characteristics of the affected population" [21]. Lim and Skidmore (2021) estimate that one additional heat island mitigation (HIM) measure lowers the heat index values by 0.261 F in locations that adopted HIM. Based on this finding, if the U.S. government adopted one additional HIM in all urban locations across the country, there would be a reduction in costs of more than $9 billion annually (for reductions in all-cause mortality). In addition, if the government reduced ground-level fine particulate matter by one standard deviation (≈2.65 $\mu g/m^3$) on average, then the related cost of all-cause mortality in the U.S. will decrease by more than $1 trillion per year.

Due to data limitations, I only have annual mortality data and thus use the annual measurements of heat and air pollution. In future work, it may be useful to employ the daily, weekly, or monthly data to better capture how short and medium-term variations of heat and air pollution influence health. Another concern is the potential endogeneity problem of the heat index. To clarify, heat may influence people's migration decisions. That is, heat may be associated with the unobserved demographic characteristics embedded in the error term. This issue may generate bias in the estimates. The potential endogeneity of the heat index variable and the selection of associated valid instrumental variables is an important topic for future research. Further, the association between meteorological variables and different air pollutants is complex and still under discussion in the epidemiologic and meteorologic fields. Given the expectation of increased frequency of extreme weather events and continued air pollution challenges globally, future work on these topics is worthy of further exploration.

[21] Mortality Risk Valuation. Retrieved from https://www.epa.gov/environmental-economics/mortality-risk-valuation on November 13, 2022.

REFERENCES

Anderson, B. G., & Bell, M. L. (2009). Weather-related mortality: How heat, cold, and heat waves affect mortality in the United States. *Epidemiology (Cambridge, Mass.)*, *20*(2), 205–213. https://doi.org/10.1097/EDE.0b013e318190ee08

Banzhaf, H. S., & Chupp, B. A. (2010). *Heterogeneous harm vs. spatial spillovers: Environmental federalism and US air pollution* (Working Paper No. 15666). National Bureau of Economic Research. https://doi.org/10.3386/w15666

Barreca, A. I. (2012). Climate change, humidity, and mortality in the United States. *Journal of Environmental Economics and Management*, *63*(1), 19–34. https://doi.org/10.1016/j.jeem.2011.07.004

Barwick, P. J., Li, S., Rao, D., & Zahur, N. B. (2018). *The morbidity cost of air pollution: Evidence from consumer spending in China* (SSRN Scholarly Paper No. 2999068). https://doi.org/10.2139/ssrn.2999068

Basu, R., & Ostro, B. D. (2008). A multicounty analysis identifying the populations vulnerable to mortality associated with high ambient temperature in California. *American Journal of Epidemiology*, *168*(6), 632–637. https://doi.org/10.1093/aje/kwn170

Bayer, P., Keohane, N., & Timmins, C. (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management*, *58*(1), 1–14. https://doi.org/10.1016/j.jeem.2008.08.004

Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., & Wara, M. (2021). The changing risk and burden of wildfire in the United States. *Proceedings of the National Academy of Sciences*, *118*(2), e2011048118. https://doi.org/10.1073/pnas.2011048118

Buyantuyev, A., & Wu, J. (2010). Urban heat islands and landscape heterogeneity: Linking spatiotemporal variations in surface temperatures to land-cover and socioeconomic patterns. *Landscape Ecology*, *25*(1), 17–33. https://doi.org/10.1007/s10980-009-9402-4

Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data*. Cambridge University Press.

Chen, S., Chen, Y., Lei, Z., & Tan-Soo, J.-S. (2021). Chasing clean air: Pollution-induced travels in China. *Journal of the Association of Environmental and Resource Economists*, *8*(1), 59–89. https://doi.org/10.1086/711476

Chen, X., Shao, S., Tian, Z., Xie, Z., & Yin, P. (2017). Impacts of air pollution and its spatial spillover effect on public health based on China's big data sample. *Journal of Cleaner Production*, *142*, 915–925. https://doi.org/10.1016/j.jclepro.2016.02.119

Chen, X., & Ye, J. (2019). When the wind blows: Spatial spillover effects of urban air pollution in China. *Journal of Environmental Planning and Management*, *62*(8), 1359–1376. https://doi.org/10.1080/09640568.2018.1496071

Chen, Y., Zhi, G., Feng, Y., Fu, J., Feng, J., Sheng, G., & Simoneit, B. R. T. (2006). Measurements of emission factors for primary carbonaceous particles from residential raw-coal combustion in China. *Geophysical Research Letters*, *33*(20). https://doi.org/10.1029/2006GL026966

Chindapol, S., Blair, J., Osmond, P., & Prasad, D. (2017). A suitable thermal stress index for the elderly in summer tropical climates. *Procedia Engineering*, *180*, 932–943.

Dawson, J. P., Bloomer, B. J., Winner, D. A., & Weaver, C. P. (2014). Understanding the meteorological drivers of U.S. particulate matter concentrations in a changing climate. *Bulletin of the American Meteorological Society*, *95*(4), 521–532.

de Freitas, C. R., & Grigorieva, E. A. (2017). A comparison and appraisal of a comprehensive range of human thermal climate indices. *International Journal of Biometeorology*, *61*(3), 487–512.

De Sario, M., Katsouyanni, K., & Michelozzi, P. (2013). Climate change, extreme weather events, air pollution and respiratory health in Europe. *The European Respiratory Journal*, *42*(3), 826–843. https://doi.org/10.1183/09031936.00074712

Deschênes, O. (2014). Temperature, human health, and adaptation: A review of the empirical literature. *Energy Economics*, *46*, 606–619. https://doi.org/10.1016/j.eneco.2013.10.013

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. https://doi.org/10.1257/app.3.4.152

Deschênes, O., & Moretti, E. (2009). Extreme Weather Events, Mortality, and Migration. *The Review of Economics and Statistics*, *91*(4), 659–681. https://doi.org/10.1162/rest.91.4.659

D'Ippoliti, D., Michelozzi, P., Marino, C., de'Donato, F., Menne, B., Katsouyanni, K., Kirchmayer, U., Analitis, A., Medina-Ramón, M., Paldy, A., Atkinson, R., Kovats, S., Bisanti, L., Schneider, A., Lefranc, A., Iñiguez, C., & Perucci, C. A. (2010). The impact of heat waves on mortality in 9 European cities: Results from the EuroHEAT project. *Environmental Health*, *9*(1), 37. https://doi.org/10.1186/1476-069X-9-37

Freeman III, A. M., Herriges, J. A., & Kling, C. L. (2014). *The measurement of environmental and resource values: Theory and methods*. Routledge.

Gabriel, K. M., & Endlicher, W. R. (2011). Urban and rural mortality rates during heat waves in Berlin and Brandenburg, Germany. *Environmental Pollution*, *159*(8–9), 2044–2050.

Hajian Tilaki, K. (2012). Methodological issues of confounding in analytical epidemiologic studies. *Caspian Journal of Internal Medicine*, *3*(3), 488–495.

Hammer, M. S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A. M., Hsu, N. C., Levy, R. C., Garay, M. J., Kalashnikova, O. V., Kahn, R. A., Brauer, M., Apte, J. S., Henze, D. K., Zhang, L., Zhang, Q., Ford, B., Pierce, J. R., & Martin, R. V. (2020). Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). *Environmental Science & Technology*, *54*(13), 7879–7890. https://doi.org/10.1021/acs.est.0c01764

Hayhoe, K., Wuebbles, D. J., Easterling, D. R., Fahey, D. W., Doherty, S., Kossin, J. P., Sweet, W. V., Vose, R. S., & Wehner, M. F. (2018). *Our changing climate. Impacts, risks, and adaptation in the united states: The fourth national climate assessment, volume II*.

Heal, G., & Park, J. (2016). Reflections—Temperature stress and the direct impact of climate change: A review of an emerging literature. *Review of Environmental Economics and Policy*, *10*(2), 347–362. https://doi.org/10.1093/reep/rew007

Hernandez, G., Berry, T.-A., Wallis, S., & Poyner, D. (2017). *Temperature and humidity effects on particulate matter concentrations in a sub-tropical climate during winter*.

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., & Kaufman, J. D. (2013). Long-term air pollution exposure and cardio- respiratory mortality: A review. *Environmental Health*, *12*(1), 43. https://doi.org/10.1186/1476-069X-12-43

Holstege, C. P. (2019). *Smoke Inhalation.* EMedicineHealth. https://www.emedicinehealth.com/smoke_inhalation/article_em.htm

Huang, W., Kan, H., & Kovats, S. (2010). The impact of the 2003 heat wave on mortality in Shanghai, China. *The Science of the Total Environment*, *408*(11), 2418–2420. https://doi.org/10.1016/j.scitotenv.2010.02.009

Institute for Health Metrics and Evaluation (IHME). (2013). *United States Physical Activity and Obesity Prevalence by County 2001-2011* [Data set]. Institute for Health Metrics and Evaluation (IHME). https://doi.org/10.6069/5E84-HD26

Institute for Health Metrics and Evaluation (IHME). (2014). *United States Smoking Prevalence by County 1996-2012* [Data set]. Institute for Health Metrics and Evaluation (IHME). https://doi.org/10.6069/H6ZD-KX77

Jacob, D. J., & Winner, D. A. (2009). Effect of climate change on air quality. *Atmospheric Environment*, *43*(1), 51–63.

Karlsson, M., & Ziebarth, N. R. (2018). Population health effects and health-related costs of extreme temperatures: Comprehensive evidence from Germany. *Journal of Environmental Economics and Management*, *91*, 93–117. https://doi.org/10.1016/j.jeem.2018.06.004

Katsouyanni, K., Touloumi, G., Samoli, E., Gryparis, A., Le Tertre, A., Monopolis, Y., Rossi, G., Zmirou, D., Ballester, F., Boumghar, A., Anderson, H. R., Wojtyniak, B., Paldy, A., Braunstein, R., Pekkanen, J., Schindler, C., & Schwartz, J. (2001). Confounding and effect modification in the short-term effects of ambient particles on total mortality: Results from 29 European cities within the APHEA2 project. *Epidemiology*, *12*(5), 521–531.

Khawand, C. (2015). *Air quality, mortality, and perinatal health: Causal evidence from wildfires* (No. pkh318). Job Market Papers. https://ideas.repec.org/p/jmp/jm2015/pkh318.html

Knol, M. J., & VanderWeele, T. J. (2012). Recommendations for presenting analyses of effect modification and interaction. *International Journal of Epidemiology*, *41*(2), 514–520.

Knowlton, K., Rotkin-Ellman, M., King, G., Margolis, H. G., Smith, D., Solomon, G., Trent, R., & English, P. (2009). The 2006 California heat wave: Impacts on hospitalizations and emergency department visits. *Environmental Health Perspectives*, *117*(1), 61–67. https://doi.org/10.1289/ehp.11594

Kovats, R. S., & Hajat, S. (2008). Heat stress and public health: A critical review. *Annual Review of Public Health*, *29*, 41–55. https://doi.org/10.1146/annurev.publhealth.29.020907.090843

Lavaine, E. (2015). An econometric analysis of atmospheric pollution, environmental disparities and mortality rates. *Environmental and Resource Economics*, *60*(2), 215–242. https://doi.org/10.1007/s10640-014-9765-0

Lim, J., & Skidmore, M. (2020). Heat vulnerability and heat island mitigation in the United

States. *Atmosphere*, *11*(6), Article 6. https://doi.org/10.3390/atmos11060558

Lin, S., Luo, M., Walker, R. J., Liu, X., Hwang, S.-A., & Chinery, R. (2009). Extreme high temperatures and hospital admissions for respiratory and cardiovascular diseases. *Epidemiology (Cambridge, Mass.)*, *20*(5), 738–746. https://doi.org/10.1097/EDE.0b013e3181ad5522

Luber, G., & McGeehin, M. (2008). Climate change and extreme heat events. *American Journal of Preventive Medicine*, *35*(5), 429–435. https://doi.org/10.1016/j.amepre.2008.08.021

Luechinger, S. (2010). Life satisfaction and transboundary air pollution. *Economics Letters*, *107*(1), 4–6. https://doi.org/10.1016/j.econlet.2009.07.007

Madrigano, J., Ito, K., Johnson, S., Kinney, P. L., & Matte, T. (2015). A Case-Only Study of Vulnerability to Heat Wave–Related Mortality in New York City (2000–2011). *Environmental Health Perspectives*, *123*(7), 672–678. https://doi.org/10.1289/ehp.1408178

McGeehin, M. A., & Mirabelli, M. (2001). The potential impacts of climate variability and change on temperature-related morbidity and mortality in the United States. *Environmental Health Perspectives*, *109*(suppl 2), 185–189.

Miller, N., Molitor, D., & Zou, E. (2017). Blowing smoke: Health impacts of wildfire plume dynamics. *Environmental And Resource Economics At The University Of Illinois*.

Mocko, D., & NASA/GSFC/HSL. (2012). *NLDAS Primary Forcing Data L4 Monthly 0.125 x 0.125 degree V002. 2012* [Data set]. Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5067/Z62LT6J96R4F

National Fire Protection Association (NFPA). (n.d.). *Reporter's Guide: The consequences of fire*. National Fire Protection Association. Retrieved February 22, 2021, from https://www.nfpa.org/News-and-Research/Publications-and-media/Press-Room/Reporters-Guide-to-Fire-and-NFPA/Consequences-of-fire

North America Land Data Assimilation System (NLDAS). (2012a). *Daily Air Temperatures and Heat Index, years 1979-2011* [Data set]. CDC WONDER Online Database. https://wonder.cdc.gov/EnvironmentalData.html

North America Land Data Assimilation System (NLDAS). (2012b). *Daily Precipitation years 1979-2011 on CDC WONDER Online Database* [Data set]. CDC WONDER Online Database. https://wonder.cdc.gov/EnvironmentalData.html

Pope, C. A., & Dockery, D. W. (2006). Health effects of fine particulate air pollution: Lines that connect. *Journal of the Air & Waste Management Association (1995)*, *56*(6), 709–742. https://doi.org/10.1080/10473289.2006.10464485

Rainham, D. G. C., & Smoyer-Tomic, K. E. (2003). The role of air pollution in the relationship between a heat stress index and human mortality in Toronto. *Environmental Research*, *93*(1), 9–19. https://doi.org/10.1016/s0013-9351(03)00060-4

Schlenker, W., & Walker, W. R. (2016). Airports, air pollution, and contemporaneous health. *The Review of Economic Studies*, *83*(2), 768–809. https://doi.org/10.1093/restud/rdv043

Shaposhnikov, D., Revich, B., Bellander, T., Bedada, G. B., Bottai, M., Kharkova, T., Kvasha, E., Lezina, E., Lind, T., Semutnikova, E., & Pershagen, G. (2014). Mortality related to air pollution with the moscow heat wave and wildfire of 2010. *Epidemiology (Cambridge, Mass.)*, *25*(3), 359–364. https://doi.org/10.1097/EDE.0000000000000090

Short, K. C. (2017). *Spatial wildfire occurrence data for the United States, 1992-2015 [FPA_FOD_20170508] (4th Edition)* [Data set]. Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2013-0009.4

Steadman, R. G. (1984). A universal scale of apparent temperature. *Journal of Climate and Applied Meteorology*, *23*(12), 1674–1687.

Stone, S. L., Anderko, L., Berger, M., Butler, C., Cascio, W., Clune, A., & others. (2019). Wildfire smoke: A guide for public health officials. *US Environmental Protection Agency*.

Stott, P. A., Christidis, N., Otto, F. E. L., Sun, Y., Vanderlinden, J.-P., van Oldenborgh, G. J., Vautard, R., von Storch, H., Walton, P., Yiou, P., & Zwiers, F. W. (2016). Attribution of extreme weather and climate-related events. *Wiley Interdisciplinary Reviews. Climate Change*, *7*(1), 23–41. https://doi.org/10.1002/wcc.380

Tai, A. P. K., Mickley, L. J., & Jacob, D. J. (2010). Correlations between fine particulate matter (PM2.5) and meteorological variables in the United States: Implications for the sensitivity of PM2.5 to climate change. *Atmospheric Environment*, *44*(32), 3976–3984. https://doi.org/10.1016/j.atmosenv.2010.06.060

Tan-Soo, J.-S. (2018). Valuing Air Quality in Indonesia Using Households' Locational Choices. *Environmental and Resource Economics*, *71*(3), 755–776. https://doi.org/10.1007/s10640-017-0182-z

United States Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), & National Center for Health Statistics (NCHS). (2020). *Underlying Cause of Death 1999-2019* [Data set]. CDC WONDER Online Database. https://wonder.cdc.gov/ucd-icd10.html

U.S. Bureau of Economic Analysis. (n.d.-a). *CAGDP1 Gross Domestic Product (GDP) summary by county and metropolitan area* [Data set]. U.S. Bureau of Economic Analysis. Retrieved July 13, 2020, from https://apps.bea.gov/itable/iTable.cfm?ReqID=70&step=1

U.S. Bureau of Economic Analysis. (n.d.-b). *CAINC30 Economic profile by county and metropolitan area* [Data set]. U.S. Bureau of Economic Analysis. Retrieved December 27, 2020, from https://apps.bea.gov/itable/iTable.cfm?ReqID=70&step=1

U.S. Bureau of Labor Statistics. (n.d.). *Consumer Price Index retroactive series using current methods (R-CPI-U-RS)* [Data set]. U.S. Bureau of Labor Statistics. Retrieved February 13, 2021, from https://www.bls.gov/cpi/research-series/r-cpi-u-rs-home.htm

U.S. Census Bureau. (n.d.-a). *Annual County Resident Population Estimates by Age, Sex, Race, and Hispanic Origin: April 1, 2010 to July 1, 2019* [Data set]. U.S. Census Bureau. Retrieved August 13, 2020, from https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-detail.html

U.S. Census Bureau. (n.d.-b). *Intercensal Estimates of the Resident Population by Five-Year Age Groups and Sex for Counties: April 1, 2000 to July 1, 2010* [Data set]. U.S. Census Bureau. Retrieved September 15, 2020, from https://www.census.gov/data/datasets/time-series/demo/popest/intercensal-2000-2010-counties.html

U.S. Census Bureau. (n.d.-c). *Small Area Income and Poverty Estimates (SAIPE) Program online database* [Data set]. U.S. Census Bureau. Retrieved February 11, 2020, from https://www.census.gov/programs-surveys/saipe/data/datasets.html

U.S. Environmental Protection Agency. (2008). *Urban Heat Island Basics. In: Reducing Urban Heat Islands: Compendium of Strategies. Draft.* U.S. Environmental Protection Agency. https://www.epa.gov/heat-islands/ heat-island-compendium

U.S. Environmental Protection Agency. (2019). *Integrated Science Assessment (ISA) for Particulate Matter (Final Report, Dec 2019)* (EPA/600/R-19/188). U.S. Environmental Protection Agency. https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=347534

U.S. Environmental Protection Agency. (2020). *Policy Assessment for the Review of the National Ambient Air Quality Standards for Particulate Matter*. U.S. Environmental Protection Agency. https://www.epa.gov/sites/production/files/2020-01/documents/final_policy_assessment_for_the_review_of_the_pm_naaqs_01-2020.pdf

U.S. Global Change Research Program (USGCRP). (n.d.). *Heat Waves*. Retrieved November 19, 2019, from https://www.globalchange.gov/browse/indicators/us-heat-waves

Van Donkelaar, A., Martin, R. V., Li, C., & Burnett, R. T. (2019). Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environmental Science & Technology*, *53*(5), 2595–2611.

VanderWeele, T. J. (2009). On the distinction between interaction and effect modification. *Epidemiology (Cambridge, Mass.)*, *20*(6), 863–871. https://doi.org/10.1097/EDE.0b013e3181ba333c

Williams, A. M., & Phaneuf, D. J. (2019). The morbidity costs of air pollution: Evidence from spending on chronic respiratory conditions. *Environmental and Resource Economics*, *74*(2), 571–603.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, *50*(2), 420–445.

Wrenn, D. H., Klaiber, H. A., & Newburn, D. A. (2017). Confronting price endogeneity in a duration model of residential subdivision development. *Journal of Applied Econometrics*, *32*(3), 661–682.

Wuebbles, D. J., Fahey, D. W., & Hibbard, K. A. (2017). *Climate science special report: Fourth national climate assessment, volume I*.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., & others. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres*, *117*(D3). https://doi.org/10.1029/2011JD016048

Xu, B., Yao, T., Liu, X., & Wang, N. (2006). Elemental and organic carbon measurements with a two-step heating-gas chromatography system in snow samples from the Tibetan Plateau. *Annals of Glaciology*, *43*, 257–262. https://doi.org/10.3189/172756406781812122

Yang, D., Liu, C., Peng, L., Ye, X., Zhou, J., Mao, Z., Yang, S., Kan, H., Fu, Q., & Chen, R. (2021). Associations between fine particulate matter constituents and hospital outpatient and emergency room visits in Shanghai, China. *Atmospheric Environment*, *261*, 118606.

Yang, J., & Zhang, B. (2018). Air pollution and healthcare expenditure: Implication for the benefit of air pollution control in China. *Environment International*, *120*, 443–455.

Yu, X., Lei, X., & Wang, M. (2019). Temperature effects on mortality and household adaptation: Evidence from China. *Journal of Environmental Economics and Management*, *96*, 195–212.

Zhang, H., Wang, Y., Park, T.-W., & Deng, Y. (2017). Quantifying the relationship between extreme air pollution events and extreme weather events. *Atmospheric Research*, *188*, 64–79.

Zheng, S., Cao, J., Kahn, M. E., & Sun, C. (2014). Real estate valuation and cross-boundary air pollution externalities: Evidence from Chinese cities. *The Journal of Real Estate Finance and Economics*, *48*(3), 398–414.

Zheng, S., Zhang, X., Sun, W., & Lin, C. (2019). Air pollution and elite college graduates' job location choice: Evidence from China. *The Annals of Regional Science*, *63*(2), 295–316.

APPENDIX

Table 1. 7 Estimation Results for All-cause Deaths

	Dependent variable: All-cause Deaths						
	Heat	Air Pollution		Confounding		Interaction	
	FE	FE	CF	FE	CF	FE	CF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HI	0.0019^{***}			0.0017^{***}	0.0016^{***}	0.0038^{***}	0.0033^{**}
	(4.14)			(3.65)	(3.28)	(2.79)	(2.38)
PM		0.0055^{***}	0.0221^{***}	0.0054^{***}	0.0211^{***}	0.0254^{**}	0.0366^{***}
		(6.32)	(5.59)	(6.23)	(5.40)	(2.32)	(3.16)
HI*PM						-0.0002^{*}	-0.0002
						(-1.87)	(-1.51)
Population	0.0023^{***}	0.0021^{***}	0.0021^{***}	0.0021^{***}	0.0021^{***}	0.0021^{***}	0.0021^{***}
	(4.29)	(4.08)	(2.98)	(4.12)	(3.02)	(4.08)	(3.00)
% Under20	0.0267^{***}	0.0276^{***}	0.0285^{***}	0.0267^{***}	0.0276^{***}	0.0267^{***}	0.0275^{***}
	(7.69)	(8.23)	(8.75)	(7.69)	(8.20)	(7.67)	(8.16)
% Over64	0.0494^{***}	0.0494^{***}	0.0516^{***}	0.0493^{***}	0.0513^{***}	0.0495^{***}	0.0514^{***}
	(15.13)	(15.41)	(17.10)	(15.29)	(16.93)	(15.42)	(16.97)
% White	-0.0039^{*}	-0.0041^{*}	-0.0043^{**}	-0.0042^{**}	-0.0045^{**}	-0.0043^{**}	-0.0045^{**}
	(-1.85)	(-1.94)	(-2.14)	(-2.01)	(-2.20)	(-2.04)	(-2.22)
Urban population	0.0111^{***}	0.0110^{***}	0.0111^{***}	0.0110^{***}	0.0110^{***}	0.0110^{***}	0.0110^{***}
	(15.24)	(15.35)	(15.46)	(15.36)	(15.47)	(15.38)	(15.47)
ln(GDP)	-0.0286^{*}	-0.0286^{*}	-0.0365^{**}	-0.0287^{*}	-0.0360^{**}	-0.0286^{*}	-0.0358^{**}
	(-1.75)	(-1.76)	(-2.24)	(-1.76)	(-2.21)	(-1.75)	(-2.18)
Obesity rate (F)	0.0008	0.0008	0.0016^{*}	0.0008	0.0015^{*}	0.0007	0.0015
	(0.87)	(0.91)	(1.80)	(0.87)	(1.71)	(0.81)	(1.63)
Obesity rate (M)	0.0059^{***}	0.0055^{***}	0.0052^{***}	0.0055^{***}	0.0052^{***}	0.0054^{***}	0.0051^{***}
	(6.39)	(6.13)	(5.55)	(6.10)	(5.54)	(6.00)	(5.48)
Smoking prevalence	0.0022^{**}	0.0018^{*}	0.0010	0.0016	0.0009	0.0016	0.0008
	(2.13)	(1.81)	(0.97)	(1.64)	(0.87)	(1.60)	(0.85)
Residuals			-0.0174^{***}		-0.0164^{***}		-0.0160^{***}
			(-4.73)		(-4.54)		(-4.47)
F-statistics for instruments			452.3070		457.8425		457.8425
			[0.0000]		[0.0000]		[0.0000]
χ^2 statistics for overidentification test			0.35		0.17		0.09
			[0.5546]		[0.6843]		[0.7634]
Obs.	32912	32912	32912	32912	32912	32912	32912

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) Standard errors of regressions in columns 3, 5, and 7 are obtained from the block-bootstrap (500 repetitions) procedure (drawing the entire FIPS code with replacement). (5) Time-fixed effects and county-fixed effects are included for all the model specifications.

Table 1. 8 Estimation Results for Respiratory System Disease Deaths

	Dependent variable: Respiratory System Disease Deaths						
	Heat	Air Pollution		Confounding		Interaction	
	FE	FE	CF	FE	CF	FE	CF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HI	0.0059***			0.0058***	0.0057***	0.0155***	0.0148***
	(5.81)			(5.74)	(5.28)	(6.18)	(5.98)
PM		0.0032**	0.0291***	0.0028*	0.0255***	0.0963***	0.1118***
		(2.06)	(3.83)	(1.80)	(3.49)	(4.43)	(5.04)
HI*PM						-0.0011***	-0.0010***
						(-4.35)	(-4.11)
Population	0.0006	0.0004	0.0004	0.0005	0.0005	0.0005	0.0004
	(0.85)	(0.60)	(0.54)	(0.76)	(0.69)	(0.66)	(0.61)
% Under20	0.0352***	0.0384***	0.0397***	0.0352***	0.0364***	0.0350***	0.0361***
	(7.87)	(8.77)	(9.08)	(7.87)	(8.13)	(7.83)	(8.07)
% Over64	0.0511***	0.0515***	0.0548***	0.0511***	0.0540***	0.0519***	0.0545***
	(11.80)	(11.88)	(12.94)	(11.80)	(12.85)	(11.96)	(12.90)
% White	0.0042	0.0045*	0.0041	0.0040	0.0036	0.0037	0.0034
	(1.54)	(1.65)	(1.54)	(1.46)	(1.36)	(1.35)	(1.26)
Urban population	0.0103***	0.0104***	0.0104***	0.0103***	0.0103***	0.0102***	0.0103***
	(9.29)	(9.29)	(9.15)	(9.31)	(9.17)	(9.33)	(9.15)
ln(GDP)	-0.0192	-0.0192	-0.0315	-0.0195	-0.0303	-0.0191	-0.0288
	(-0.82)	(-0.80)	(-1.33)	(-0.83)	(-1.31)	(-0.82)	(-1.24)
Obesity rate (F)	0.0041***	0.0042***	0.0055***	0.0041***	0.0052***	0.0039***	0.0049***
	(2.86)	(2.93)	(3.64)	(2.85)	(3.47)	(2.72)	(3.28)
Obesity rate (M)	0.0065***	0.0063***	0.0058***	0.0063***	0.0058***	0.0059***	0.0055***
	(4.23)	(4.15)	(3.74)	(4.15)	(3.78)	(3.87)	(3.55)
Smoking prevalence	0.0031*	0.0033*	0.0020	0.0028	0.0017	0.0026	0.0016
	(1.74)	(1.93)	(1.16)	(1.62)	(0.96)	(1.51)	(0.93)
Residuals			-0.0269***		-0.0237***		-0.0213***
			(-3.58)		(-3.24)		(-2.92)
F-statistics for instruments			452.3070		457.8425		457.8425
			[0.0000]		[0.0000]		[0.0000]
χ^2 statistics for overidentification test			0.80		0.28		0.04
			[0.3718]		[0.5965]		[0.8349]
Obs.	32912	32912	32912	32912	32912	32912	32912

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) Standard errors of regressions in columns 3, 5, and 7 are obtained from the block-bootstrap (500 repetitions) procedure (drawing the entire FIPS code with replacement). (5) Time-fixed effects and county-fixed effects are included for all the model specifications.

Table 1. 9 Estimation Results for Circulatory System Disease Deaths

	Dependent variable: Circulatory System Disease Deaths						
	Heat	Air Pollution		Confounding		Interaction	
	FE	FE	CF	FE	CF	FE	CF
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
HI	0.0025***			0.0024***	0.0023***	0.0045***	0.0041***
	(3.86)			(3.78)	(3.61)	(2.73)	(2.58)
PM		0.0016*	0.0125***	0.0014	0.0109**	0.0209	0.0276**
		(1.75)	(2.83)	(1.56)	(2.53)	(1.59)	(2.03)
HI*PM						-0.0002	-0.0002
						(-1.52)	(-1.35)
Population	0.0020***	0.0019***	0.0019***	0.0020***	0.0019***	0.0019***	0.0019***
	(3.71)	(3.62)	(2.82)	(3.66)	(2.87)	(3.63)	(2.86)
% Under20	0.0173***	0.0186***	0.0191***	0.0173***	0.0178***	0.0173***	0.0178***
	(5.37)	(5.95)	(6.37)	(5.34)	(5.75)	(5.31)	(5.69)
% Over64	0.0469***	0.0470***	0.0484***	0.0469***	0.0481***	0.0470***	0.0482***
	(15.40)	(15.51)	(16.08)	(15.35)	(15.93)	(15.47)	(16.00)
% White	-0.0040**	-0.0039*	-0.0041**	-0.0041**	-0.0043**	-0.0042**	-0.0043**
	(-1.96)	(-1.91)	(-2.00)	(-2.00)	(-2.07)	(-2.03)	(-2.09)
Urban population	0.0101***	0.0101***	0.0101***	0.0101***	0.0101***	0.0101***	0.0101***
	(12.40)	(12.42)	(12.70)	(12.38)	(12.66)	(12.39)	(12.66)
ln(GDP)	-0.0274*	-0.0274*	-0.0326**	-0.0273*	-0.0318**	-0.0272*	-0.0315**
	(-1.73)	(-1.73)	(-2.05)	(-1.72)	(-2.01)	(-1.71)	(-1.98)
Obesity rate (F)	-0.0009	-0.0009	-0.0004	-0.0010	-0.0005	-0.0010	-0.0006
	(-0.89)	(-0.87)	(-0.38)	(-0.90)	(-0.47)	(-0.93)	(-0.53)
Obesity rate (M)	0.0039***	0.0038***	0.0036***	0.0038***	0.0036***	0.0037***	0.0035***
	(3.25)	(3.22)	(2.81)	(3.19)	(2.82)	(3.13)	(2.77)
Smoking prevalence	-0.0002	-0.0001	-0.0006	-0.0003	-0.0008	-0.0004	-0.0008
	(-0.14)	(-0.08)	(-0.51)	(-0.27)	(-0.64)	(-0.30)	(-0.65)
Residuals			-0.0114***		-0.0099**		-0.0095**
			(-2.62)		(-2.34)		(-2.24)
F-statistics for instruments			452.3070		457.8425		457.8425
			[0.0000]		[0.0000]		[0.0000]
χ^2 statistics for overidentification test			0.68		1.03		1.23
			[0.4094]		[0.3112]		[0.2683]
Obs.	32912	32912	32912	32912	32912	32912	32912

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) Standard errors of regressions in columns 3, 5, and 7 are obtained from the block-bootstrap (500 repetitions) procedure (drawing the entire FIPS code with replacement). (5) Time-fixed effects and county-fixed effects are included for all the model specifications.

Table 1. 10 First-stage Results - Main Analysis

	(1) PM_{25}	(2) PM_{25}	(3) PM_{25}
Distant PM (d≥100km)	0.0942***	0.0944***	0.0944***
	(29.65)	(29.79)	(29.79)
Distant Fire (30km≤d≤100km, ≥100 acres)	46.6177***	46.6609***	46.6609***
	(10.16)	(10.18)	(10.18)
HI		-0.0030	-0.0030
		(-0.69)	(-0.69)
Population	-0.0004	-0.0004	-0.0004
	(-0.08)	(-0.08)	(-0.08)
% Under20	-0.0610***	-0.0599***	-0.0599***
	(-5.18)	(-5.01)	(-5.01)
% Over64	-0.1281***	-0.1279***	-0.1279***
	(-9.64)	(-9.63)	(-9.63)
% White	0.0079	0.0082	0.0082
	(0.83)	(0.86)	(0.86)
Urban population	-0.0060**	-0.0060**	-0.0060**
	(-2.35)	(-2.36)	(-2.36)
ln(GDP)	0.5895***	0.5873***	0.5873***
	(10.38)	(10.30)	(10.30)
Obesity rate (F)	-0.0438***	-0.0438***	-0.0438***
	(-8.39)	(-8.37)	(-8.37)
Obesity rate (M)	0.0241***	0.0241***	0.0241***
	(3.97)	(3.98)	(3.98)
Smoking prevalence	0.0401***	0.0404***	0.0404***
	(6.21)	(6.25)	(6.25)
F-statistics for instruments	452.3070	457.8425	457.8425
	[0.0000]	[0.0000]	[0.0000]
Obs.	32912	32912	32912
adj. R^2	0.511	0.511	0.511

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) All the standard errors are clustered by county. (5) Time-fixed effects and county-fixed effects are included for all the model specifications.

Table 1. 11 First-stage Results - Different Radiuses (Distant Fire ⩾100 acres)

	(1) Distant PM (d≥80km) & DistPM_Fire (30km≤d<80km)	(2) Distant PM (d≥100km) & DistPM_Fire (30km≤d<100km)	(3) Distant PM (d≥150km) & DistPM_Fire (30km≤d<150km)
Distant PM	0.0973***	0.0944***	0.0878***
	(31.24)	(29.79)	(26.90)
DistPM_Fire	51.8487***	46.6609***	36.7249***
	(8.59)	(10.18)	(12.32)
HI	-0.0032	-0.0030	-0.0027
	(-0.73)	(-0.69)	(-0.61)
Population	-0.0005	-0.0004	0.0001
	(-0.09)	(-0.08)	(0.03)
% Under20	-0.0590***	-0.0599***	-0.0609***
	(-4.95)	(-5.01)	(-5.06)
% Over64	-0.1275***	-0.1279***	-0.1282***
	(-9.63)	(-9.63)	(-9.59)
% White	0.0082	0.0082	0.0086
	(0.88)	(0.86)	(0.90)
Urban population	-0.0061**	-0.0060**	-0.0059**
	(-2.40)	(-2.36)	(-2.31)
ln(GDP)	0.5841***	0.5873***	0.5898***
	(10.27)	(10.30)	(10.26)
Obesity rate (F)	-0.0435***	-0.0438***	-0.0444***
	(-8.36)	(-8.37)	(-8.43)
Obesity rate (M)	0.0243***	0.0241***	0.0237***
	(4.04)	(3.98)	(3.87)
Smoking prevalence	0.0402***	0.0404***	0.0413***
	(6.24)	(6.25)	(6.35)
F-statistics for instruments	496.5736	457.8425	386.0210
	[0.0000]	[0.0000]	[0.0000]
Obs.	32912	32912	32912
adj. R^2	0.512	0.511	0.509

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) All the standard errors are clustered by county. (5) Time-fixed effects and county-fixed effects are included for all the model specifications.

Table 1. 12 First-stage Results - Different Fire Sizes (DistPM(d≥100km) & DistPM_Fire (30km≤d≤100km))

Dependent variable: PM2.5	(1) 0.26~299 acres	(2) ≥100 acres	(3) All lightning-caused fire acres
Distant PM	0.0936***	0.0944***	0.0932***
	(29.67)	(29.79)	(29.68)
DistPM_Fire	40.3039***	46.6609***	56.4596***
	(8.43)	(10.18)	(4.85)
HI	-0.0028	-0.0030	-0.0021
	(-0.64)	(-0.69)	(-0.47)
Population	-0.0005	-0.0004	-0.0001
	(-0.10)	(-0.08)	(-0.03)
% Under20	-0.0593***	-0.0599***	-0.0587***
	(-4.95)	(-5.01)	(-4.91)
% Over64	-0.1285***	-0.1279***	-0.1291***
	(-9.61)	(-9.63)	(-9.70)
% White	0.0081	0.0082	0.0097
	(0.85)	(0.86)	(1.03)
Urban population	-0.0064**	-0.0060**	-0.0065**
	(-2.47)	(-2.36)	(-2.51)
ln(GDP)	0.5862***	0.5873***	0.5788***
	(10.29)	(10.30)	(10.21)
Obesity rate (F)	-0.0437***	-0.0438***	-0.0431***
	(-8.34)	(-8.37)	(-8.25)
Obesity rate (M)	0.0242***	0.0241***	0.0252***
	(3.98)	(3.98)	(4.14)
Smoking prevalence	0.0411***	0.0404***	0.0416***
	(6.34)	(6.25)	(6.42)
F-statistics for instruments	455.1311	457.8425	441.9121
	[0.0000]	[0.0000]	[0.0000]
Obs.	32912	32912	32912
adj. R^2	0.510	0.511	0.510

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) All the standard errors are clustered by county. (5) Time-fixed effects and county-fixed effects are included for all the model specifications.

Table 1. 13 Robustness Check – Distant PM and Distant Wildfires as IVs

	Main Analysis	Robustness Check		
	DistPM (≥100km) & *DistPM_Fire* (30~100km, 10 ~ 999 acres)	*DistPM* (≥100km) & *DistFire* (30~100km)		
		DistFire (0.26~299 acres)	*DistFire* (10 ~ 999 acres)	*DistFire* (All lightning-caused fire acres)
Dependent Variable: All-cause Deaths				
HI	0.0016***	0.0016***	0.0017***	0.0016***
	(3.28)	(3.26)	(3.33)	(3.36)
PM	0.0211***	0.0213***	0.0198***	0.0174***
	(5.40)	(5.41)	(5.43)	(4.07)
Residuals	-0.0164***	-0.0167***	-0.0152***	-0.0126***
	(-4.54)	(-4.57)	(-4.47)	(-3.19)
χ^2**-statistics**	0.17	0.07	0.91	7.64
	[0.6843]	[0.7924]	[0.3406]	[0.0057]
Dependent Variable: Respiratory System Disease Deaths				
HI	0.0057***	0.0057***	0.0057***	0.0056***
	(5.28)	(5.30)	(5.31)	(5.26)
PM	0.0255***	0.0236***	0.0243***	0.0259***
	(3.49)	(3.38)	(3.80)	(3.73)
Residuals	-0.0237***	-0.0217***	-0.0226***	-0.0242***
	(-3.24)	(-3.15)	(-3.53)	(-3.49)
χ^2**-statistics**	0.28	1.35	0.56	0.11
	[0.5965]	[0.2447]	[0.4534]	[0.7431]
Dependent Variable: Circulatory System Disease Deaths				
HI	0.0023***	0.0023***	0.0024***	0.0024***
	(3.61)	(3.60)	(3.64)	(3.69)
PM	0.0109**	0.0118***	0.0097**	0.0073
	(2.53)	(2.58)	(2.35)	(1.50)
Residuals	-0.0099**	-0.0109**	-0.0087**	-0.0061
	(-2.34)	(-2.48)	(-2.19)	(-1.31)
χ^2**-statistics**	1.03	2.33	0.15	0.33
	[0.3112]	[0.1272]	[0.6943]	[0.5639]
F-statistics for IVs	457.8425	471.1886	477.8983	446.3813
	[0.0000]	[0.0000]	[0.0000]	[0.0000]
Obs.	32912	32912	32912	32912

Note: (1) t statistics in parentheses. (2) * p < 0.1, ** p < 0.05, *** p < 0.01. (3) p-values of test statistics in squared brackets. (4) Standard errors are obtained from the block-bootstrap (500 repetitions) procedure (drawing the entire FIPS code with replacement). (5) Time-fixed effects and county-fixed effects are included for all the model specifications. (6) The χ^2-statistics of overidentification tests are based on Wooldridge (2010), and the instrumental variable for distant $PM_{2.5}$ attributed to lightning-caused wildfires is excluded.

Table 1. 14 First-stage Results - Distant PM and Distant Wildfires as IVs (DistPM(d≥100km) & DistFire (30km≤d≤100km))

Dependent variable: $PM_{2.5}$	(1) 0.26~299 acres	(2) ≥100 acres	(3) All lightning-caused fire acres
Distant PM	0.0953***	0.0964***	0.0939***
	(29.87)	(29.98)	(29.69)
DistFire	0.3272***	1.1811***	0.0005***
	(12.06)	(14.11)	(7.93)
HI	-0.0037	-0.0037	-0.0022
	(-0.86)	(-0.85)	(-0.52)
Population	-0.0008	-0.0006	-0.0005
	(-0.15)	(-0.12)	(-0.10)
% Under20	-0.0597***	-0.0606***	-0.0584***
	(-4.98)	(-5.08)	(-4.90)
% Over64	-0.1284***	-0.1274***	-0.1277***
	(-9.65)	(-9.67)	(-9.61)
% White	0.0067	0.0069	0.0087
	(0.71)	(0.73)	(0.92)
Urban population	-0.0065**	-0.0059**	-0.0062**
	(-2.52)	(-2.34)	(-2.42)
ln(GDP)	0.5834***	0.5894***	0.5798***
	(10.26)	(10.39)	(10.24)
Obesity rate (F)	-0.0439***	-0.0435***	-0.0426***
	(-8.38)	(-8.34)	(-8.17)
Obesity rate (M)	0.0237***	0.0237***	0.0251***
	(3.91)	(3.94)	(4.15)
Smoking prevalence	0.0402***	0.0398***	0.0420***
	(6.22)	(6.18)	(6.51)
F-statistics for instruments	471.1886 [0.0000]	477.8983 [0.0000]	446.3813 [0.0000]
Obs.	32912	32912	32912
adj. R^2	0.513	0.515	0.516

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) All the standard errors are clustered by county. (5) Time-fixed effects and county-fixed effects are included for all the model specifications.

CHAPTER 2: THE IMPACTS OF WILDFIRES AND WILDFIRE-INDUCED AIR POLLUTION ON MORTALITY IN THE UNITED STATES

1. Introduction

Over the past several decades, the total acreage damaged by wildfires and the average size of wildfires in the United States (U.S.) have trended upward. According to the United States National Interagency Fire Center (U.S. NIFC), the average numbers of wildfire events and acres damaged between 1985 and 2018 are about 74,201 events and 5,330,465 acres, respectively, and the corresponding total fire suppression costs exceeded $3 billion in 2018. Moreover, there are nearly 50 million homes currently located in the wildland-urban interface (WUI) in the U.S., and this number increased by about one million every three years (Burke et al., 2021). This trend is expected to continue. Though some papers examined wildfire impacts in the field of ecology, epidemiology, and economics, relatively few studies investigate how wildfires influence human health, directly and indirectly. Clarifying the mechanisms by which wildfires influence human health and examining the extent of the direct and indirect health impacts provides additional scientific evidence to help policymakers efficiently allocate the resources to reduce wildfire risk.

This paper aims to investigate how and to what extent wildfires affect mortality directly and indirectly. The analysis is based on the U.S. nationwide county-year panel dataset, which includes variables on wildfires, air pollution, and mortality and a series of variables on meteorology, economics, demographics, urbanization, and health status and behaviors. I apply the mediation analysis approach to explore the channels through which wildfires affect all-cause mortality, mortality caused by physical health problems (respiratory system diseases and circulatory system diseases), and mortality caused by mental health problems (suicide) and estimate the total, direct, and indirect impacts of wildfires. In particular, I focus on the air pollution channel. To address the potential endogeneity of air pollution, I construct two instrumental variables: wind-based and distance-weighted imported all-source air pollution and air pollution emitted from lightning-caused wildfires. Further, I quantify the spillover effects of wildfires and examine whether the causes of wildfires influence the estimation of wildfire-related health impacts.

As a prelude to the complete set of findings, I find a significant positive association between local wildfire events and $PM_{2.5}$ concentration level. As wildfire size increases, the adverse effect of wildfire on air quality increases. Meanwhile, wildfire occurrences are significantly and positively associated with all categories of mortality. I apply the mediation analysis to distinguish the direct[22] and indirect impacts of wildfire on mortality through air pollutant emissions. I find significant positive effects of wildfire-induced $PM_{2.5}$ on all-cause mortality, mortality caused by respiratory system diseases, and mortality caused by circulatory system diseases. In contrast, the indirect impact of wildfires through $PM_{2.5}$ on suicide is not significant. I conclude a complementary mediation for these three models based on Zhao et al. (2010), Wen and Ye (2014), and the present research. However, there should be channels other than $PM_{2.5}$ through which wildfires indirectly affect mortality. The indirect impacts of wildfires through $PM_{2.5}$ account for about 58.12%, 46.98%, and 21.46% of the total effects of wildfires on all-cause mortality, mortality caused by respiratory system diseases, and mortality caused circulatory system diseases, respectively. Under the assumption that there is only one possible channel of air pollution through which wildfires influence suicide, the occurrences of wildfires with at least 100 acres are directly associated with a higher risk of suicide. In addition, I find that the direct impact of wildfires on suicide can be delayed and should be monitored in the long run. Overall, for the mortality resulting from suicide, there is no mediation effect through $PM_{2.5}$. Further, the spillover effect of wildfires is significant and positive, which is even substantially larger than the local wildfire effect. Last, although most previous studies assume wildfires as exogenous, most of the wildfires are associated with or the result of human activities. This study finds that failing to consider the causes of wildfires will lead to upwardly biased estimates of health impacts.

I contribute to the literature and provide the policy implications in the following ways. First, this paper offers a more comprehensive picture of the adverse health impacts of wildfires. Existing research mainly focuses on either the wildfire events per se or the wildfire-induced effects, among which wildfire-induced air pollution is the most widely studied. In this paper, I explore how and to what extent wildfires affect mortality both directly and indirectly and provide evidence of potential indirect health impacts of wildfires through channels other than air pollution. Second, the literature on the risks of suicide induced by wildfires and wildfire-caused disamenities is still limited. This study clarifies the sources of increased suicide risks from wildfires, which can help government officials provide post-disaster psychological support. Third, this paper delivers

[22] Since I only focus on the channel of air pollution, it is possible that the direct impact here should include the mixed direct impact and indirect impacts through channels other than air pollution.

precise estimates of wildfires' direct and indirect impacts through $PM_{2.5}$ on physical and mental health on a national scale. Previous literature typically focuses on specific areas or wildfire events or uses imprecise measures of wildfire smoke (such as dummy variables). The panel dataset used in this paper covers 2,992 U.S. counties from 2001 through 2011[23], representing about 98.5% of the U.S. population. Fourth, I examine the necessity of distinguishing between human-caused and nature-caused wildfires[24]. Current research often treated wildfires as exogenously determined. However, most wildfires are caused by human activities. According to the Fire Program Analysis fire-occurrence database (FPA FOD) (Short, 2017), about 15.1% of wildfire events were caused by lightning over the 1992 to 2015 period across the contiguous U.S. This study references whether it is necessary to distinguish between wildfire causes.

The rest of this paper is organized as follows. In the next section, I review the most relevant literature. Section three provides details on the data and methodology. Section four presents and discusses the empirical results, and the last section concludes.

2. Literature Review

2.1 Wildfires and Wildfires-induced Disamenities/Amenities

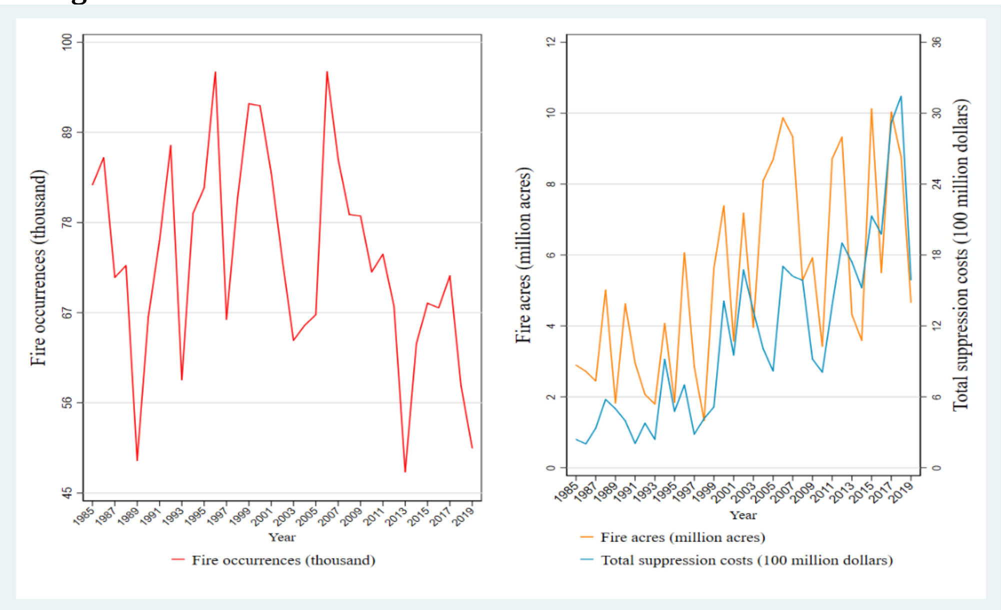
In recent decades, the annual fire acres and the total fire suppression costs exhibit an upward trend, although the frequency of wildfires is relatively stable, as shown in **Figure 2. 1**, based on wildfire statistics provided by NIFC. The upward trend is expected to continue due to more frequent and intensive extreme weather conditions such as extreme heat and drought. **Figure 2. 2** presents the distribution of average wildfire acres across the contiguous U.S. from 2001 to 2011, based on the Fire Program Analysis fire-occurrence database (Short, 2017). The data show that wildfires are heterogeneously distributed, with the western U.S. experiencing greater wildfire risk than the eastern regions. This observation is also consistent with the higher frequency of droughts in the western region.

---

[23] I focus on the period between 2001 and 2011 because of the data availability. The heat index data is only available before 2011.

[24] In the Fire Program Analysis fire-occurrence database (FPA FOD) (Short, 2017), the causes of wildfires include lightning, equipment use, smoking, campfire, debris burning, railroad, arson, children, miscellaneous, fireworks, powerline, structure, and missing/undefined.

Figure 2. 1 Wildfires Trends in the United States from 1985 to 2019



Sources: Authors' illustration. **Data:** National Interagency Fire Center, External Affairs Office[25].

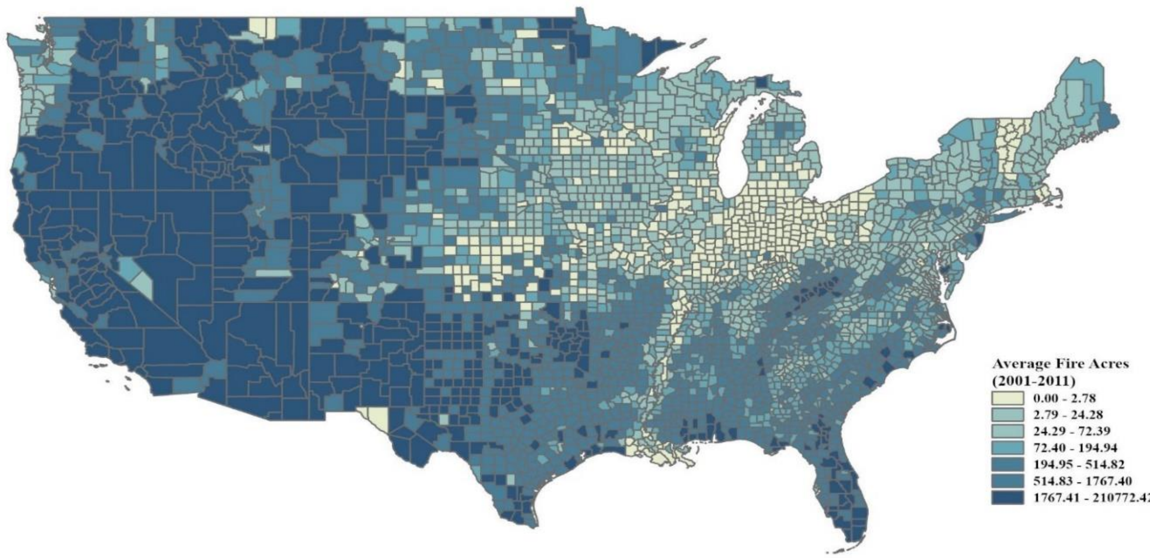
As a natural disturbance that occurs in most terrestrial ecosystems, wildfires influence the air, soil, water, fauna, flora, fuels, recreation opportunities, cultural resources, and archeology (Sandberg et al., 2002; Neary et al., 2005; Venn and Calkin, 2011; Doerr and Santín, 2016). Among these non-market forest goods and services, wildfire smoke is a major threat to human health. The burning of biomass and soil-based organic matter can generate a large amount of haze and smoke, which have complex components such as greenhouse gases, photochemically reactive compounds, sulfur dioxide (SO2), particulate matter (PM), and liquids (Neary et al., 2005; Viswanathan et al., 2006; Urbanski et al., 2008). Although the smoke components are complex, the primary pollutant is particulate matter (Sandberg et al., 2002; Stone et al., 2019). PM_{10}, particles with aerodynamic diameters generally less than or equal to 10 μm, are comprised of fine particles ($PM_{2.5}$), particles with aerodynamic diameters generally less than or equal to 2.5 μm, and coarse fractions ($PM_{10-2.5}$), particles with aerodynamic diameters generally greater than 2.5 μm and less than or equal to 10 μm (U.S. EPA, 2020). Particles from wildfire smoke tend to be very small (with a size range near

---

[25] Retrieved from https://www.nifc.gov/fireInfo/fireInfo_documents/SuppCosts.pdf on October 18, 2020. Provided by: Individual Federal Land Management Agencies. Fire totals include all private, state and federal lands in the United States for the year. Costs are provided for the FY.

the wavelength of visible light (0.4-0.7 μm)), and about 90% of total particle masses consist of $PM_{2.5}$ (Stone et al., 2019)[26]. $PM_{2.5}$ can therefore be used as the primary measurement of wildfire smoke. Wildfire is an important PM source (Dawson et al., 2014; U.S. EPA, 2020). The U.S. EPA (2020) concluded that wildfires contribute to 10% ~ 20% of primary PM emissions in the U.S. annually. Other studies, such as Khawand (2015) and Burke et al. (2021), simulated or estimated the air pollution emitted from wildfires, finding that wildfires contribute roughly 15% to 25% of $PM_{2.5}$ in the U.S., respectively. In contrast to the upward trend in wildfires, over the same period, the U.S. has experienced an overall declining trend in ambient concentrations of $PM_{2.5}$ (1990-2014) and the number of extreme $PM_{2.5}$ days (2000-2009) (H. Zhang et al., 2017; U.S. EPA, 2020). On the one hand, these downward trends reflect the effectiveness of EPA's stringent air quality regulations in recent decades (H. Zhang et al., 2017; U.S. EPA, 2020). On the other hand, it also indicates the growing relative importance of wildfires on $PM_{2.5}$ and the importance of wildfire management and education.

Figure 2. 2 Distribution of Wildfires in the United States from 2001 to 2011



Source: Authors' illustration. **Data:** Fire Program Analysis fire-occurrence database (FPA FOD).

Wildfires do not always lead to environmental disamenities. It can positively or negatively influence the soil, water, fauna, flora, fuels, recreation opportunities, cultural resources, and

[26] Sandberg et al. (2002) summarized that 90% of all smoke particles are PM_{10}, and 90% of PM_{10} is $PM_{2.5}$.

archeology (Venn and Calkin, 2011). On the downside, the change of soil structure caused by fire can reduce soil productivity, increase the vulnerability of postfire runoff and erosion, lead to postfire floods, degrade water quality, etc. However, wildfires also increase the availability of nutrients for plant growth in the short term, reduce the potential for epidemic insect and disease infestations, provide more novelty of the burnt landscape, and reduce the risk of future wildfires, etc. (Neary et al., 2005; Venn and Calkin, 2011).

2.2 Direct Health Impacts of Wildfires

The most direct health impacts from wildfires are the direct injuries and fatalities caused by wildfires. According to the National Oceanic and Atmospheric Administration's (NOAA's) National Centers for Environmental Information (NCEI) Storm Events Database, 6,331 wildfire events between 1999 and 2017 resulted in about 87 and 8 direct injuries and fatalities on average per year, respectively[27]. However, these data do not capture suicide fatalities, which may not be observed directly or in the short term after wildfires.

Most studies on the health impacts of wildfires focus on the indirect wildfire impacts, and even fewer studies explore the potential direct impacts of wildfire on mental health and suicide. Wildfires are classified as natural disasters by the U.S. EPA. Mental health risks associated with wildfires are less discussed than with other disasters such as earthquakes, floods, and tsunamis. Traumatic events can generate psychiatric pathology among those who experience such shocks (Caamano-Isorna et al., 2011). For example, property damages, injuries, and fatalities caused by wildfires can result in feelings of helplessness; children, the elderly, women, and single parents are particularly vulnerable (DeWolfe, 2000; Caamano-Isorna et al., 2011; Kõlves et al., 2013). Other vulnerability factors include pre-existing factors (such as cultural and socioeconomic background, and physical and psychological factors), natural disaster-related factors (such as type of event, the magnitude of the event, the threat to the life, and the extent of loss), and post-disaster factors (such as social support, coping skills, and secondary stressors) (DeWolfe, 2000; Kõlves et al., 2013). Previous studies found an association between wildfires and mental health issues, such as insomnia, post-traumatic stress disorder (PTSD), anxiety, depression, and risks of suicide

[27] Storm Events Database provided by National Oceanic and Atmospheric Administration's (NOAA's) National Centers for Environmental Information (NCEI) is available at https://www.ncdc.noaa.gov/stormevents/ftp.jsp. The definitions and examples of direct injuries, indirect injuries, direct fatalities, and indirect fatalities caused by wildfires can be found at https://www.nws.noaa.gov/directives/sym/pd01016005curr.pdf.

(McDermott et al., 2005; Caamano-Isorna et al., 2011; Papadatou et al., 2012; Psarros et al., 2017; Brown et al., 2019). In addition, a delayed increase in suicidal behaviors was found after a "honeymoon phase"[28]. Thus, potential suicidal behavior should be monitored several years after the disaster (Kõlves et al., 2013).

2.3 Indirect Health Impacts of Wildfires

Wildfires produce various non-market forest goods and services so wildfires can affect human health through multiple channels. Based on the NOAA's NCEI Storm Events Database, there are about 19 and 2 indirect injuries and deaths caused by wildfires per year on average, respectively, among the 6,331 wildfire events recorded from 1999 to 2017. One example of indirect fatalities/injuries included in this database is "all vehicular accidents caused by reduced visibility due to smoke". However, these statistics do not track the chronic disease's morbidity and mortality and suicides induced by wildfire-caused disamenities. Also, the amenities produced by wildfires, such as providing more recreational opportunities and reducing future carbon emissions, may provide health benefits.

Wildfire smoke is the focus of this present research among the range of wildfire-caused disamenities and potential amenities. Wildfire smoke poses a threat to human health, and PM is the principal pollutant of concern from wildfire smoke (Stone et al., 2019). The health effects of PM are linked to the size of the particles. Large particles with aerodynamic diameters greater than 10 micrometers can irritate the eyes, nose, and throat but do not usually reach the lungs (Stone et al., 2019). In contrast, small particles, PM_{10}, $PM_{10\text{-}2.5}$, and $PM_{2.5}$, can be inhaled into the lungs and thus affect the lungs, heart, and blood vessels. $PM_{2.5}$ is the most significant risk among these small particles since $PM_{2.5}$ can reach deep into the lungs and may even enter the bloodstream (Stone et al., 2019). Previous studies documented the adverse impacts of PM on morbidity and premature mortality (Dominici et al., 2006; Brook et al., 2010; Hoek et al., 2013; Dawson et al., 2014; Khawand, 2015; H. Zhang et al., 2017; U.S. EPA, 2019). Besides, the U.S. EPA Integrated Science Assessment (ISA) for Particulate Matter Report (2019) concluded that causal relationships between exposure to $PM_{2.5}$ and respiratory effects are likely to exist. This evidence of morbidity

[28] There are six phases that people may experience before, during, and after the natural disasters: the pre-disaster phase, the impact phase, the heroism (rescue) phase, the honeymoon (remedy) phase, the disillusionment phase, and the reorganization (reconstruction and recovery) phase (Pasnau and Fawzy, 1989; DeWolfe, 2000; Kõlves et al., 2012).

provides biological plausibility for cause-specific mortality (such as mortality caused by respiratory and cardiovascular diseases) and ultimately total mortality (U.S. EPA, 2019).

Many recent papers explored the health impacts of $PM_{2.5}$ attributed to wildfires. Previous research found a strong association between wildfire smoke and respiratory morbidity and mortality, but the effects on cardiovascular morbidity showed mixed results (Liu et al., 2017; Dittrich and McCallum, 2020). Also, the wildfires-specific $PM_{2.5}$ was more toxic than equal doses from other sources (or ambient $PM_{2.5}$) and was associated with a higher respiratory effect than non-wildfire $PM_{2.5}$ (Kochi et al., 2010; Dittrich and McCallum, 2020; Aguilera et al., 2021).

Not only can wildfires trigger mental health problems, but wildfire-induced $PM_{2.5}$ may also lead to mental health symptoms and an increased risk of suicide. However, many current papers focus on all-source $PM_{2.5}$ rather than wildfire-specific $PM_{2.5}$. Compared to the physical impacts of $PM_{2.5}$, the literature on the psychological effects is limited and further research is still required. However, emerging evidence still shows associations between PM and adverse mental health outcomes (such as anxiety, depression, bipolar disorder, psychosis, and suicide) (Bakian et al.,2015; Lin et al., 2016; X. Zhang et al., 2017; Gładka et al., 2018; Lu et al., 2018; Braithwaite et al., 2019). Previous studies found that PM is associated with mental health symptoms such as depression and psychosis by influencing the nervous system and thus increasing the risks of suicide (Gładka et al., 2018; Braithwaite et al., 2019). People may perceive higher health risks from wildfire-specific air pollution than air pollution from urban air pollution sources (Kochi et al., 2010). Further, people who experience wildfires are likely to suffer higher levels of stress due to this perception. The risks of suicide were heterogeneous between females and males and among people with different education levels (Lin et al., 2016).

In addition to suicide, air pollution may increase crime and other antisocial behavior (Lu et al., 2018; Burkhardt et al., 2019; Burkhardt et al., 2020), thus increasing all-cause mortality. Potential pathways are that the air pollutants can induce anxiety, generate a sense of anonymity, diminish moral appropriateness, and spur impulsive aggressive behavior (Lu et al., 2018; Burkhardt et al., 2019).

2.4 Existing Studies on Wildfire Smoke and Health Impacts of Wildfires

One of the challenges of studies on wildfire-induced air pollution is distinguishing air pollution from wildfire and non-wildfire sources. Many previous studies focused on a specific wildfire event within a particular region. These studies explored the health impacts of wildfire smoke by linking variations in air pollutants or air quality and health outcomes. Two typically applied methods are the time-series and historical control methods (Kochi et al., 2010). The historical control method compares the aggregate adverse health outcomes between the study and control periods but is not ideal for detecting relatively small health impacts (Kochi et al., 2010). In contrast, time-series analysis does not have sufficient statistical power since the wildfire event may not last for a long enough period (Kochi et al., 2010). Also, it fails to consider the transboundary effects and long-run effects of wildfire smoke. Since these studies focus on specific wildfire events, the variation in air quality and health outcomes in the particular areas were only observed before, during, and after the event. However, wildfire-induced air pollutants can be carried by wind and influence a broader spatial scale. Further, the direct impact of wildfires on suicide risks may be delayed (Kõlves et al., 2013), and the U.S. EPA also confirmed the adverse health impacts of long-term exposure to $PM_{2.5}$ (U.S. EPA, 2019).

Air pollution data used in previous research are typically from a local/nearby monitoring station. Recent studies used satellite-based data or statistical and geographical techniques to measure the ambient $PM_{2.5}$ and wildfire smoke across broader spatial scales. Many papers used satellite-based smoke exposure data from the NOAA Hazard Mapping System (HMS). However, it measures the wildfire smoke with an indicator variable (thin, medium, or thick) and does not precisely measure air pollutant levels. As a result, it cannot be used to estimate the precise health impact of wildfire smoke corresponding to a specific air pollution exposure. Despite this, previous studies made good use of this dataset and derived various wildfire smoke measurements. Jones (2018) linked the smoke plume data with the county location and used the frequency of smoke days as the primary measurement for wildfire smoke. Miller et al. (2017) used a dummy variable to indicate whether a given day in the areas defined by zip code was covered by smoke. By further combing wildfire occurrence data from major wildland management agencies and the air pollution data from the EPA, Miller et al. (2017) found that wind patterns influenced the spread of wildfire smoke, and that smoke exposure was significant and transient. Also, similar to large-size wildfires, smoke from smaller-size wildfires can travel long distances (Miller et al., 2017).

One approach to estimate wildfire-caused smoke is using chemical transport models (CTMs) such as GEOS-Chem. This method is typically used in studies of a single wildfire event, and it is computationally demanding and requires surmounting several major uncertainties in the pathway between source and receptor (Burke et al., 2021; Aguilera et al., 2021). Liu et al. (2017)

applied the GEOS-Chem model to estimate the daily $PM_{2.5}$ attributable to wildfires in 561 western U.S. counties and found a significant association between short-term high wildfire-specific $PM_{2.5}$ and respiratory admissions. Instead of CTMs, some papers apply statistical approaches and the HMS data to estimate wildfire-specific smoke. These approaches do not require heavy computational effort and offer the advantage of modeling daily wildfire-specific $PM_{2.5}$ with a fine resolution over a long study period and extensive area (Aguilera et al., 2021). Aguilera et al. (2020) employed four approaches to isolate the health impacts of wildfire-specific $PM_{2.5}$ and found that the estimates of the wildfire-specific $PM_{2.5}$ impact on respiratory admissions are similar though varying in amplitude. Burke et al. (2021) considered both fire and smoke information in HMS and constructed a statistical model to predict $PM_{2.5}$ across the nation with and without wildfire smoke; the difference between the two predictions is the $PM_{2.5}$ attributed to smoke.

This paper aims to study both the direct wildfire effects and the indirect effects of wildfire through $PM_{2.5}$, so I apply a mediation analysis approach, as presented in Section 3.2. I offer the following hypotheses based on the literature discussed above.

Hypothesis 1: Wildfires can indirectly increase the mortality caused by chronic diseases such as respiratory and circulatory system diseases through emitting wildfire smoke.

Hypothesis 2: Wildfires can directly increase suicide risks, and there may be a delayed impact.

3. Data and Method

This section introduces the data and method employed in this paper. As a general overview, I model fatalities as a function of several factors, including wildfire and wildfire smoke, meteorology, economics**,** demographics, urbanization, and health status and behaviors. The data I use in the evaluation are described next.

3.1 Data

Table 2. 1 provides the definitions and data sources of the variables used in empirical analysis, and the summary statistics are presented in **Table 2. 2**. I focus on all-cause deaths, physical health outcomes (deaths caused by respiratory system disease (International Classification of Diseases, Revision 10 (ICD-10) code: J00-J98), deaths caused by circulatory system disease (ICD-10 code: I00- I99)), and mental health outcome (deaths caused by suicide). County-level data for annual mortality in the United States are from the Center for Disease Control and Prevention (CDC) WONDER online database.

The wildfire data are from the Fire Program Analysis fire-occurrence database (FPA FOD)

(Short, 2017). This database includes 1,880,465 wildfire events from 1992 to 2015. After excluding the observations for Puerto Rico, Alaska, and Hawaii, I have 1,835,646 observations in total[29]. I then generate wildfire occurrences and acres for each county by different sizes and causes. I assume that there is no wildfire event if there is no record for a specific county and specific year.

$PM_{2.5}$ is the principal pollutant of health concern from wildfire smoke; about 90% of total wildfire smoke particle mass emitted from wildfires consists of $PM_{2.5}$ (Stone et al., 2019). In this paper, I use the annual concentration estimates of ground-level $PM_{2.5}$ data collected from the Atmospheric Composition Analysis Group (ACAG)[30] as the primary measurement of wildfire smoke[31]. **Figure 2. 3** presents the distribution of the average $PM_{2.5}$ in the United States from 2001 to 2011. Generally, this map is consistent with the U.S. EPA ISA report (2019, 2020) that the eastern areas suffered a higher but more uniform level of $PM_{2.5}$ than western areas, whereas California has a significantly higher level of $PM_{2.5}$ than surrounding western states.

I also include control variables on meteorology, economic and demographic factors, urbanization, and health status and behaviors. For the meteorological variables, I control for the average daily maximum heat index, average daily precipitation, and average daily sunlight. Meteorological variables are associated with and may also contribute to various health problems. The heat index, precipitation, and sunlight data are obtained from the CDC WONDER system and are initially from the North America Land Data Assimilation System (NLDAS)[32]. The heat index measures "how hot it really feels when relative humidity is factored in with the actual air

[29] Since some of the county information (643,450 out of 1,835,646 events) are missing in the database, I map the longitude and latitude of wildfire into the 2010 TIGER/Line Shapefiles of the county from the Census Bureau to obtain the missing county information, and I obtain additional 643,446 county FIPS codes. There are still four wildfire events without county information, and thus I drop these four events. Comparing the county FIPS codes in the database to the generated county FIPS codes from shapefile, there are 44,287 county information unmatched, and the overall unmatched rate is about 3.71%, which is very low. Considering that the wildfires may occur near the boundary, I treat these 44,287 wildfires as occurring in both counties. Because most of the dates on which the wildfires were declared contained (or controlled) are missing and I use annual wildfire information, I rely on the discovery date of wildfires.

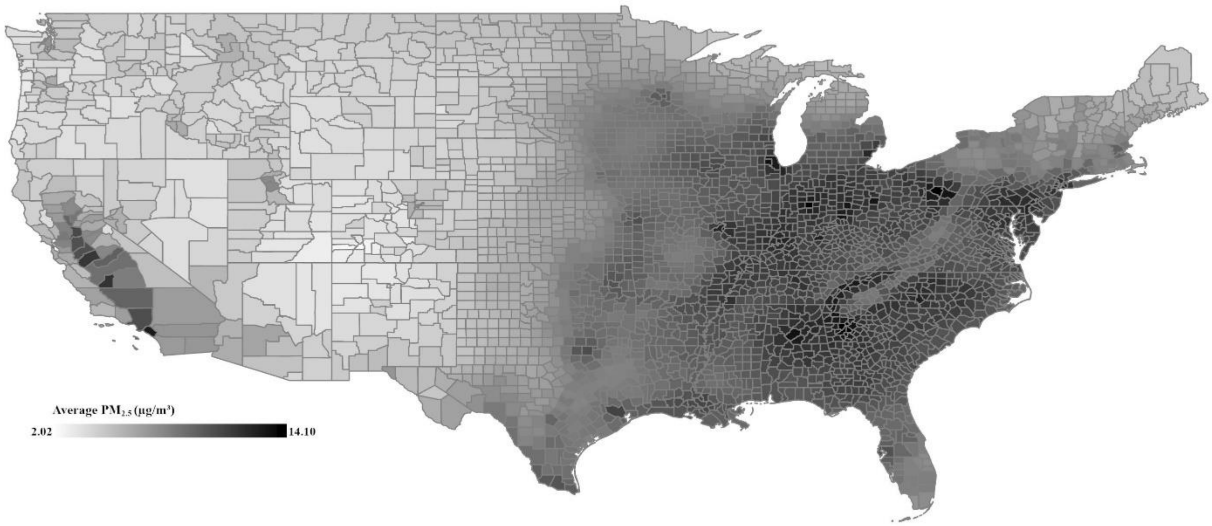
[30] Surface $PM_{2.5}$ dataset (North American Regional Estimates (V4.NA.03)) from Atmospheric Composition Analysis Group. Downloaded from https://sites.wustl.edu/acag/datasets/surface-pm2-5/#V4.NA.02.MAPLE on August 18, 2020. The shapefile of U.S. counties was downloaded from https://www2.census.gov/geo/tiger/TIGER2010/COUNTY/2010/.

[31] The county-level mean $PM_{2.5}$ data is obtained by calculating the zonal statistics using ArcGIS Pro.

[32] The heat index defined by NOAA and heat index data provided by NLDAS are not available for temperature below 80°F (27°C).

temperature"[33]. The heat index obtained from the CDC WONDER system is calculated based on Steadman (1979). After simplification, only the ambient dry bulb temperature and relative humidity were included in the heat index equation[34].

Figure 2. 3 Average Ground-level Particulate Matter ($PM_{2.5}$) (μg/m³) (2001-2011)



Source: Authors' illustration. **Data:** North American Regional Estimates for Surface $PM_{2.5}$ (V4.NA.03), ACAG.

Real per capita GDP (logarithm of chained 2012 dollars) measures local economic development, and the data are collected from the U.S. Bureau of Economic Analysis (BEA). Demographic variables include population size (in 10,000s), percentage of people under the age of 20, percentage of people over the age of 64, and percentage of people who are white, all of which are from the CDC WONDER database and the U.S. Census Bureau (CB). The percentage of the urban population measures urbanization and the data used are collected from the U.S. CB [35]. Health status and risky behaviors may also influence fatalities, especially fatalities induced by

[33] Heat Index. The U.S. National Oceanic and Atmospheric Administration National Weather Service (NOAA NWS). Retrieved from https://www.weather.gov/safety/heat-index on March 28, 2021.
[34] North America Land Data Assimilation System (NLDAS) Daily Air Temperatures and Heat Index, 1979 – 2011, on CDC WONDER. Retrieved from https://wonder.cdc.gov/wonder/help/Climate/ta_htindx.PDF on November 22, 2019.
[35] Yearly data of the percentage of urban population over the period 1999 to 2010 are obtained by interpolation and extrapolation using the U.S. census data and shapefiles of U.S. urbanized areas and counties for years 2000 and 2010. The shapefiles were downloaded from https://www.census.gov/geographies/mapping-files/2000/geo/carto-boundary-file.html and https://www2.census.gov/geo/tiger/TIGER2010/.

diseases. Therefore, I control for female and male obesity prevalence and the prevalence of people who smoke, which are collected from the Institute for Health Metrics and Evaluation (IHME).

To address the endogeneity of air pollution, I construct the imported all-source $PM_{2.5}$ and the $PM_{2.5}$ caused by lightning-caused wildfires from distant counties as instrumental variables. These variables are created using exogenous sources: wind direction, geographic distance, and lightning phenomenon. The construction approach and discussion of their validity are presented in Section 3.2.2. To obtain the distances among counties, I use the 2010 TIGER/Line Shapefiles of U.S. counties from the Census Bureau and generate the distance weighting matrix. To incorporate the wind direction effects on the transport of air pollution, I collect the monthly zonal and meridional wind speeds data from Phase 2 of the North American Land Data Assimilation System (NLDAS-2) on the website of the National Aeronautics and Space Administration (NASA). The county-level monthly zonal and meridional wind speeds are calculated using ArcGIS Pro.

I obtained a balanced panel dataset of 11 years spanning 2001 to 2011, covering 2,992 counties in the 48 U.S. contiguous states and Washington, DC. This dataset covers about 98.5% of the U.S. population. I also include county-fixed effects and time-fixed effects, which control for the unobservable time-constant county-specific heterogeneity and the time-varying but county-constant factors such as nationwide shocks that may have occurred in a given year, respectively.

Table 2. 1 List of Variables in the Empirical Analysis

Dependent Variables			**Source**
Health Outcomes	All-cause Deaths	$Death_{All_{it}}$	CDC WONDER
	Deaths Caused by Respiratory System Diseases	$Death_{R_{it}}$	CDC WONDER
	Deaths Caused by Circulatory System Diseases	$Death_{C_{it}}$	CDC WONDER
	Deaths Caused by Suicide	$Death_{s_{it}}$	CDC WONDER
Explanatory/Control Variables			**Source**
Wildfires	Annual Total Wildfire Occurrences	$Fire_{Occ_{it}}$	FPA FOD
	Annual Total Wildfire Acres	$Fire_{Acres_{it}}$	FPA FOD
	Annual Total Wildfire Occurrences Caused by Lightning	$FireLightning_{Occ_{it}}$	FPA FOD
	Annual Total Wildfire Acres Caused by Lightning	$Firelightning_{Acre_{it}}$	FPA FOD
Air Pollution	Average Ground-level Particulate Matter (μg/m³)	PM_{it}	ACAG

Table 2. 1 (cont'd)

Meteorology	Average Daily Maximum Heat Index (F)	HI_{it}	CDC WONDER
	Average Daily Precipitation (mm)	$Precip_{it}$	CDC WONDER
	Average Daily Sunlight (KJ/m²) (in thousands)	$Sunlight_{it}$	CDC WONDER
Economics	Real Per Capita GDP (chained 2012 dollars) (logarithm)	GDP_{it}	U.S. BEA
	Real Per Capita Income (dollars, based year:1977) (logarithm)	$Income_{it}$	U.S. BEA, U.S. BLS
	Poverty Rate (%)	$Poverty_{it}$	U.S. CB
Demographics	Population Size (in 10 thousand)	$Population_{it}$	CDC WONDER
	Percentage of the Young (under 20) (%)	$Under20_{it}$	U.S. CB
	Percentage of the Elderly (over 64) (%)	$Over64_{it}$	U.S. CB
	Percentage of the White (%)	$White_{it}$	U.S. CB
Urbanization	Percentage of Urban Population (%)	$UrbanPopu_{it}$	U.S. CB
	Percentage of Urban areas (%)	$UrbanArea_{it}$	U.S. CB
	Urban Population Density (per 1000 square meters)	$Density_{it}$	U.S. CB
Health	Prevalence of Obesity (Female) (%)	$Obesity_{F_{it}}$	IHME
	Prevalence of Obesity (Male) (%)	$Obesity_{M_{it}}$	IHME
	Prevalence of People Who Currently Smoke (%)	$Smoking_{it}$	IHME
Time FE	Year Indicator Variables	T_t	-
County FE	County Indicator Variables	C_i	-
Instrumental Variables			**Source**
Instruments	$PM_{2.5}$ from Distant Counties (μg/m³)	$DistPM_{it}$	ACAG, U.S. CB, NASA
	$PM_{2.5}$ Attributed to Lightning-caused Wildfires from Distant Counties (μg/m³)	$DistPM\_Fire_{it}$	ACAG, FPA FOD, U.S. CB, NASA

Table 2. 2 Summary Statistics

Variables	**Mean (2001~2011)**	**Std. Dev. (2001~2011)**	**Mean (2001)**	**Mean (2011)**
All-cause Deaths	807.25	2183.75	798.12	830.10
Deaths Caused by Respiratory System Diseases	77.78	197.42	75.96	82.12
Deaths Caused by Circulatory System Diseases	278.87	800.31	306.41	258.81
Deaths Caused by Suicide	11.25	29.81	10.03	12.96
Annual Total Wildfire Occurrences	27.70	53.04	28.87	29.48
Annual Wildfire Occurrences (≥100 acres)	0.84	2.56	0.75	1.19

Table 2. 2 (cont'd)

Annual Total Wildfire Acres	1738.59	15376.82	1238.13	2808.87
Annual Total Wildfire Occurrences Caused by Lightning	4.04	18.15	4.63	4.10
Annual Total Wildfire Acres Caused by Lightning	875.92	12370.81	721.97	1074.30
Average Ground-level Particulate Matter ($PM_{2.5}$) (μg/m³)	8.72	2.65	9.25	7.95
Average Daily Maximum Heat Index (F)	90.14	3.71	89.64	92.24
Average Daily Sunlight (KJ/m²)	16235.11	1584.03	16094.85	16606.98
Average Daily Precipitation (mm)	2.76	1.05	2.55	2.77
Real Per Capita GDP (chained 2012 dollars)	37956.90	26308.69	34676.84	40234.21
Population Size	98184.63	311435.60	93784.86	102522.50
Percentage of the Young (under 20) (%)	26.99	3.29	28.15	25.89
Percentage of the Elderly (over 64) (%)	15.27	4.04	14.81	16.10
Percentage of the White (%)	86.47	15.74	87.19	85.74
Percentage of Urban Population (%)	42.19	30.70	41.52	42.79
Prevalence of Obesity (Female) (%)	35.74	5.83	31.40	39.13
Prevalence of Obesity (Male) (%)	33.41	4.33	28.81	37.17
Prevalence of People Who Currently Smoke (%)	25.86	4.06	27.03	24.41
Number of Observations	32912	32912	2992	2992

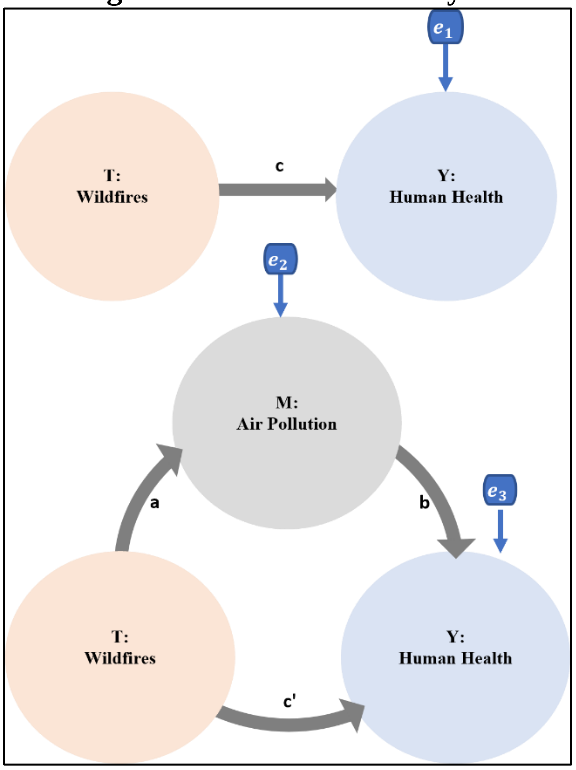
3.2 Method

3.2.1 Direct and Indirect Impacts of Local Wildfires

More recent papers applied a mediation analysis approach to conduct mechanism/channel studies (Alan et al., 2018; Pace et al., 2022; Shi, 2022). In this present paper, I also apply this approach to study how wildfires affect human health. Wildfires can, directly and indirectly, affect health through air pollution and other environmental considerations. Thus, these wildfire-caused environmental factors can be regarded as the mediating variables or mechanisms through which wildfires affect health outcomes. The relationship is shown in **Figure 2. 4** and is represented by equations (1), (2), and (3). The direct impact of wildfires, c', and the total indirect/mediation impact of wildfires, ab, sum to yield the total effect of wildfires on health, c. Since there may be more than one mediating variable, the total mediation effect equals the sum of each path's mediation effects. The total effect equals the sum of the mediation effect and the direct effect. Since I only discuss the channel of air pollution in this paper, if the mediation effects of other channels exist, then c' should also include the indirect effects of other channels.

$$Y = cT + e_1 \quad (1)$$
$$M = aT + e_2 \quad (2)$$
$$Y = c'T + bM + e_3 \quad (3)$$

Figure 2. 4 Mediation Analysis



The most well-known and widely used way to conduct mediation analysis is the causal step approach as outlined in the classic paper of Baron and Kenny (1986), which requires the following three criteria for the establishment of mediation: (1) c must be statistically significantly different from zero; (2) a must be statistically significantly different from zero; and (3) b must be statistically significantly different from zero and c' is no longer statistically significant from zero. This approach has found to have the following problems: (1) the inferences about the indirect effects is based on the outcomes of hypothesis tests on a and b, rather than the estimate of the indirect effect ab, and the separate test of a and b are of the low power; (2) this approach involves several null hypothesis tests but only one inferential test of the indirect effect is needed; (3) whether or not the mediation effect exists is just a qualitative claim, which is not based on a quantification of the indirect effect and does not carry information on uncertainty that can be reflected by a confidence interval; (4) the total effect, c, does not need to be statistically significantly different from zero because of the potential suppression (inconsistent mediation or competitive mediation), the lower power of the test of total effects, and the existence of subpopulations in which the total effects have different signs, etc. (MacKinnon et al., 2000; Zhao et al., 2010; Wen and Ye, 2014; Hayes, 2018).

Recent studies further developed this approach and put forward some new procedures and tests for the mediation analysis (MacKinnon et al., 2000; Zhao et al., 2010; Imai et al., 2011; Wen and Ye, 2014). One important recommendation is that the establishment of the mediation effect should rely on the significance of the indirect effect, ab, only, instead of the separate test of a and b, the indirect effect is examined directly by the testing H_0: $ab = 0$ (Zhao et al., 2010). Many approaches have been put forward, including the Sobel test, bootstrap, the Monte Carlo, and the distribution of the product approaches[36]. The bootstrap approach is more widely recommended (Zhao et al., 2010; Wen and Ye, 2014; Hayes, 2018). Several bootstrap methods are typically used: The percentile bootstrap, bias-corrected bootstrap, and bias-corrected and accelerated bootstrap. Although the bias-corrected bootstrap and bias-corrected and accelerated bootstrap are better than the percentile bootstrap in principle, these approaches have an elevated risk of Type I error under certain conditions (Wen and Ye, 2014; Hayes, 2018).

Another important development is the identification of the causal mediation effect. According to Imai et al (2011), "sequential ignorability" assumptions are needed to identify the causal mediation relationship, which can be written as equations (4) and (5). First, given a series of control variables, wildfire occurrence is independent of health outcomes and air pollution. Second, given

---

[36] One approach recommended by Baron and Kenny (1986) is the Sobel test, which is called the normal theory approach. However, the problems are that this method assumes a normal sampling distribution of ab and that this method has lower power and generates less accurate confidence intervals than other methods (Zhao et al., 2010; Hayes, 2018). Other methods to test indirect effects include bootstrap, the Monte Carlo, and the distribution of the product approaches. Compared to the Sobel test, the bootstrap method does not assume the sampling distribution of ab and has higher power than the normal theory approach (Hayes, 2018). The Monte Carlo approach relies on the fact that the distribution of ab is not normal, but the sampling distributions of a and b tend to be nearly normal (Hayes, 2018). In contrast, the distribution of the product approach relies on a mathematical approximation of the sampling distribution of the product (Hayes, 2018). Further, note that the bootstrap and Monte Carlo approaches are simulation-based and generate asymmetric confidence intervals that are preferred when the sampling distribution of the estimators is asymmetric (the sampling distribution of ab is usually asymmetric) (Hayes, 2018).

a series of control variables and wildfire occurrence, air pollution is independent of health outcomes.

$$\{Y_i(m,t), M_i(t)\} \perp\!\!\!\perp T_i | X_i = x \quad (4)$$
$$Y_i(m,t) \perp\!\!\!\perp M_i(t) | T_i = t, X_i = x \quad (5)$$

Following the procedures presented in equations 1, 2, and 3, I first estimate the total impacts of wildfires. Since the outcome variables are the number of annual deaths in each county, priority is the count data models. The distributions of death data have characteristics of non-negativity, discreteness, and left skewness, so I employ a conditional Fixed-effect Poisson quasi-maximum likelihood model. This model is robust to an arbitrary misspecified distribution and any serial correlation so long as the conditional mean is correctly specified (Cameron and Trivedi, 2013; Wooldridge, 2010). The regression equation for the total impact of wildfires is shown in equation (6).

$$E(Deaths_{it} | Fire_{it}, X_{it}, C_i, T_t) = C_i \cdot \exp(\alpha_T Fire_{it} + \boldsymbol{\alpha} \cdot \boldsymbol{X_{it}} + T_t) \quad (6)$$

I use the number of annual wildfire occurrences (at least 100 acres) as the primary measurement of wildfires for the baseline analysis. The covariates, $\boldsymbol{X_{it}}$, include variables on the meteorology, economics, demographics, degree of urbanization, health status/behaviors, and the percentage of lightning-caused wildfires, as introduced in Section 3.1. I include the percentage of lightning-caused wildfires because most wildfires are caused by human activities, which may be correlated with health outcomes. County fixed effects, C_i, and year fixed effects, T_t, are included as well.

If equations (5) and (6) hold, I can then estimate the average causal mediation effect through equations (7) and (8). I test whether local wildfires significantly influence local air quality in the next step, as shown in equation (7). Similarly, I use the number of annual wildfire occurrences (at least 100 acres) as the primary measurement of wildfires as the baseline analysis. The number of annual wildfire occurrences with other fire sizes and the annual wildfire acres are used as alternative substitutes for robustness checks. The covariates, $\boldsymbol{E_{it}}$, include all the control variables in equation (6) but the health status and behavior variables. As previously noted, county-fixed effects, C_i, and the time-fixed effects, T_t, are included.

$$PM_{it} = \gamma_0 + \gamma_1 Fire_{it} + \boldsymbol{\gamma} \cdot \boldsymbol{E_{it}} + T_t + C_i + u_{it} \quad (7)$$

I include $PM_{2.5}$ in equation (6) in the last step, as shown in equation (8). Then I can use the percentile bootstrap to test the indirect effect, $\gamma_1 * \beta_P$, directly. Since I use the Fixed-effect Poisson model, $\gamma_1 * \beta_P$ is an approximation of indirect impact. The formula for the indirect effect of one additional wildfire event (at least 100 acres) on fatalities is $(exp\ (\gamma_1 * \beta_P) - 1) * 100\%$, which can be approximated by $\gamma_1 * \beta_P$.

$$E(Deaths_{it} | Fire_{it}, PM_{it}, \boldsymbol{X}_{it}, C_i, T_t) = C_i \cdot \exp(\beta_D Fire_{it} + \beta_P PM_{it} + \boldsymbol{\beta} \cdot \boldsymbol{X_{it}} + \ T_t\ ) \quad (8)$$

Given that previous studies show a delayed effect of wildfires on suicide, I also explore the longer-term effect of wildfires by adding extra 1-year lags of wildfires and the percentage of lightning-caused wildfires.

3.2.2 Addressing Endogeneity of $PM_{2.5}$

However, the second assumption, i.e., equation (5) may not hold, because $PM_{2.5}$ may still be correlated with unobserved factors in the error term even when I have added a series of control variables above along with the county and year fixed effects. For instance, an efficient government may enact more stringent environmental policies to control air pollution and provide better access to medical care. In addition, the sectoral composition of an economy may not be fully captured by GDP. For example, a higher share of economic activity in the industrial sector can lead to larger emissions of air pollutants. Another potential unobservable can stem from avoidance behaviors. People concerned about detrimental local air quality may choose to move to other counties. These omitted demographic characteristics may be associated with local health outcomes as well.

To address this problem, Imai et al. (2011) suggested applying an instrumental variable approach. Because I use a Fixed-effect Poisson model in equation (8), I then apply the Control Function (CF) approach to address this concern. At least one excluded exogenous variable is required (Wooldridge, 2010, 2015). Previous literature explored a variety of instrumental variables for air pollution. In this paper, I follow the instrumental variable construction method adopted in previous papers such as Bayer et al. (2009), Zheng et al. (2014), Tan-Soo (2018), Barwick et al.

(2018), Yang and Zhang (2018), and Chen et al. (2021). I construct the wind-driven distance-weighted imported all-source $PM_{2.5}$ and $PM_{2.5}$ attributed to the lightning-caused wildfires as instrumental variables. These two instrumental variables are defined as follows:

$$DistPM_{it} = \sum_{i \neq j} PM_{jt} * I(WD_{jt} = GD_{ji}) * \frac{1}{d_{ij}^{p}}, \qquad d_{ij}^{p} \geq 100km \quad (9)$$

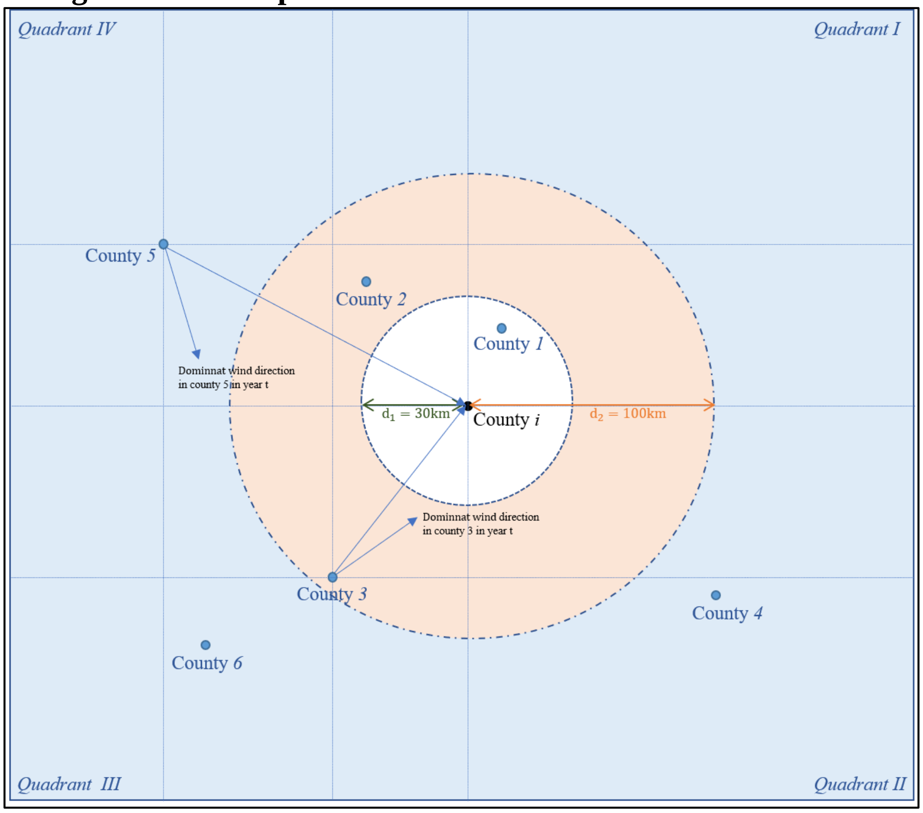
$$PM_{it} = \pi_l Wildfires\_lightning_{it} + C_i + T_t + u_{it} \quad (10)$$

$$DistPM\_Fire_{it} = \sum_{i \neq j} \widehat{PM_{jt}} * I(WD_{jt} = GD_{ji}) * \frac{1}{d_{ij}^{f}}, \quad 30km \leq d_{ij}^{f} \leq 100km \quad (11)$$

WD_{jt} represents the dominant wind direction(s) in county *j* at year *t*, which is defined as the most frequent wind direction(s) within 12 months in county *j* at year *t*. GD_{ji} is the geographic direction of the vector from county *j* to county *i*. Both WD_{jt} and GD_{ji} have four categories and are defined by the quadrants, in which the dominant wind direction vector and the vector from county *j* to county *i* fall. $\widehat{PM_{jt}}$ denotes the predicted $PM_{2.5}$ attributed to lightning-caused wildfires. $Wildfires\_lightning_{it}$ denotes the number of occurrences of lightning-caused wildfires with 10 ~ 999 burned acres. For each equation, I consider two types of weights, which are based on wind direction and geographic distance. For the imported all-source $PM_{2.5}$ in county *i* at year *t*, $DistPM_{it}$, I consider the annual mean $PM_{2.5}$ imported from upwind counties located at least 100 km away from county *i*. The $PM_{2.5}$ is weighted by the reciprocal of geographic distance (km), as presented by equation (9). The farther the county locates, the smaller the spillover effect of distant $PM_{2.5}$ has on local $PM_{2.5}$. For the distant $PM_{2.5}$ attributed to lightning-caused wildfires in county *i* at year *t*, $DistPM\_Fire_{it}$, I consider the lightning-caused wildfires with 10 ~ 999 burned acres in upwind counties located between 30 km and 100 km away from county *i*. I first predict the $PM_{2.5}$ attributed to lightning-caused wildfires for each county by regressing the local $PM_{2.5}$ on the occurrences of lightning-caused wildfires and the year and county-fixed effects. The lightning-caused wildfires can be regarded as an exogenous source of local $PM_{2.5}$, as presented in equation (10). Then, I use a similar approach to construct the imported $PM_{2.5}$ due to distant lightning-caused wildfires, as presented by equation (11).

To evaluate the robustness, $PM_{2.5}$ attributed to distant lightning-caused wildfires with 0.26 ~ 299 burned acres and $PM_{2.5}$ attributed to distant all-size lightning-caused wildfires are also constructed[37]. Moreover, I consider different radius sizes of the buffer zones to check the robustness. The radius of the first buffer zone is 80 and 150 km, while the range of the second buffer zone is 30-80 km and 30-150 km, respectively.

Figure 2. 5 Examples of Instrumental Variables Construction



Note: For county *i*, the imported all-source $PM_{2.5}$ is from the counties located outside the circle with a radius of 100 km, such as county 4, county 5, and county 6. For example, the vector from county 5 to county *i* falls in quadrant II and the dominant wind direction of county 5 in year *t* falls in quadrant II as well, so I assign the weight of wind direction to be one (i.e., $\mathrm{I}(WD_{5t} = GD_{5i}) = 1$). The impact of imported all-source $PM_{2.5}$ from county 5 is weighted by the reciprocal of the distance between county 5 and county *i*. The imported $PM_{2.5}$ attributed to lightning-caused wildfires for county *i* is from counties within a circle of 100km but outside the circle of 30km such as county 2 and county 3. For county 3, both the vector from county 3 to county *i* and the dominant wind direction of county 3 fall in quadrant I (i.e., $\mathrm{I}(WD_{3t} = GD_{3i}) = 1$). The imported $PM_{2.5}$ attributed to lightning-caused wildfires from county 3 is weighted by the reciprocal of the distance between county 3 and county *i*.

---

[37] The distant lightning-caused wildfires with 0.26 ~ 299 burned acres (0.001~ 1.210 km^2) is measured using the number of occurrences and the distant all-size lightning-caused wildfires is measured using the total burned acres. In addition, the number of lightning-caused wildfire events with fire acres greater than 30*30 km^2 (about 222,395 acres) from 2001 to 2011 in the dataset is only 27 (there are 134,481 lightning-caused wildfire events in total).

Figure 2. 6 Distribution of Average Distant $PM_{2.5}$ (μg/m³) (2001-2011)

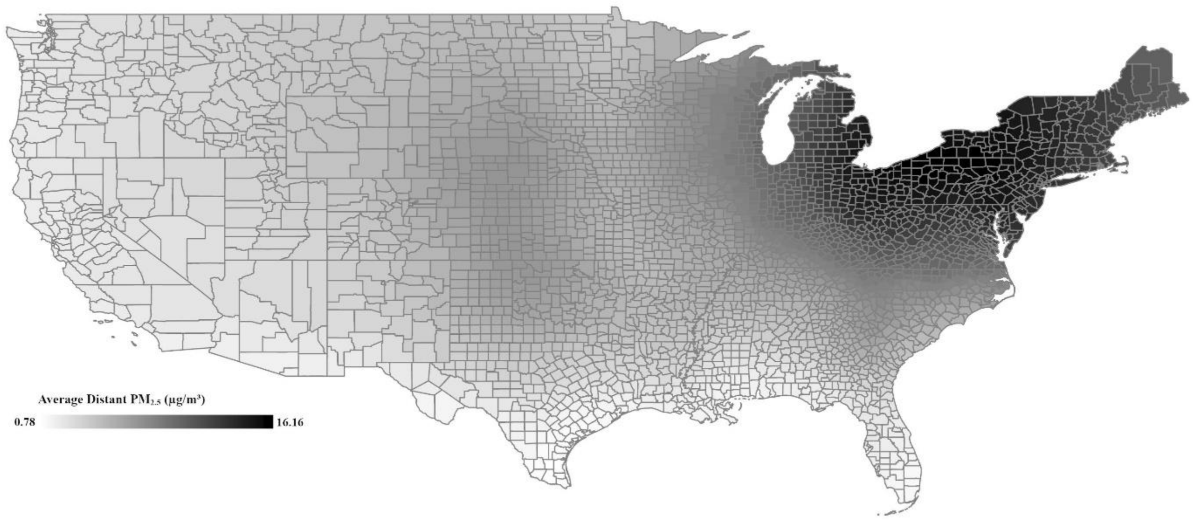


Figure 2. 7 Distribution of Average $PM_{2.5}$ (μg/m³) Attributed to Distant Lightning-caused Wildfires (2001-2011)

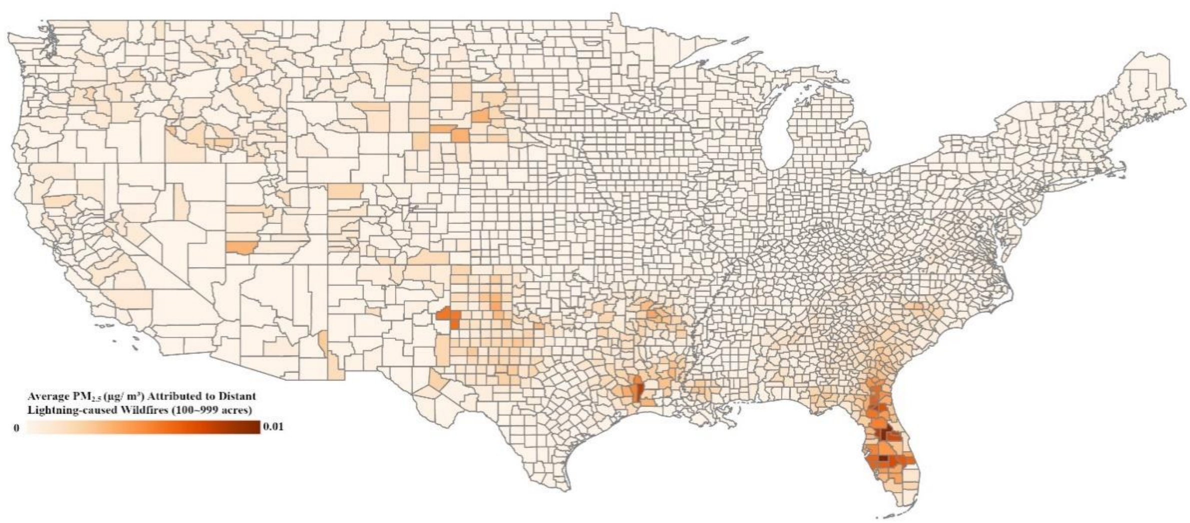


Figure 2. 5 shows examples to illustrate the construction of instrumental variables. **Figure 2. 6** and **Figure 2. 7** show the spatial distributions of the average *DistPM* and *DistPM_Fire* from 2001 to 2011. The all-source $PM_{2.5}$ is imported from counties that locate 100 km away and the $PM_{2.5}$ attributed to lightning-caused wildfires (10 ~ 999 acres) is imported from counties located in a range of 30 ~ 100 km away. These two figures highlight the counties that suffered the more all-source $PM_{2.5}$ and $PM_{2.5}$ attributed to lightning-caused wildfires from distant counties over the 2001-2011 period. Influenced by the wind direction, the distribution of distant $PM_{2.5}$ shows some

differences with the local $PM_{2.5}$, but in general, eastern areas suffered more imported all-source air pollution. The transboundary externalities of wildfire-caused $PM_{2.5}$ are more significant in the western and southern areas. It is also consistent with the higher temperature and more droughts in the west U.S. However, considering that Florida experienced a significantly higher level of $PM_{2.5}$ resulting from lightning-caused wildfires than other states, we also conducted a robustness examination excluding Florida.

Next, I discuss the validity of these instrumental variables. The valid IVs should satisfy two conditions: relevance and exogeneity. First, $PM_{2.5}$ in other counties should be associated with $PM_{2.5}$ in the focal county. Previous studies have found evidence of the significant transboundary spillover effect of air pollution, and this spillover effect is mainly driven by the wind and associated with distance (Bayer et al., 2009; Banzhaf and Chupp, 2010; Luechinger, 2010; Khawand, 2015; Zheng et al., 2014; Barwick et al., 2018; Yang and Zhang, 2018; Chen and Ye, 2019; Williams and Phaneuf, 2019; Zheng et al., 2019; Chen et al., 2021). I also consider the impact of lightning-caused wildfires on air pollution since wildfire is an important PM source (Khawand, 2015; U.S. EPA, 2020). Moreover, the previous finding shows that similar to large-sized wildfires, smoke from smaller-sized wildfires can also travel long distances (Miller et al., 2017). Because distant wildfires should affect the air quality in distant counties and distant air pollutants can be carried by the wind across counties, I expect a significant positive association between air pollution attributed to distant wildfires and local air pollution. The F-statistics of instrumental variables obtained in the first-stage regression of the control function approach (as presented in Section 4) indicate that these two instrumental variables are strong predictors of $PM_{2.5}$.

Second, the instrumental variables should not directly influence the outcome variables. To minimize the likelihood that nearby $PM_{2.5}$ is correlated with the variables influencing the health outcomes in the focal county, I create two buffer zones and exclude counties outside the buffer zones. Based on previous studies, I set the radius of the buffer zone to be 100 km for the distant all-source $PM_{2.5}$. Given that I do not use any information on counties located near 100 km, I create the second instrumental variable, which considers the transboundary externalities of $PM_{2.5}$ attributed to lightning-caused wildfires at the second buffer zone. The second buffer zone is 30 km away from the focal counties but within the first buffer zone of 100 km. The lightning-caused wildfires are an exogenous natural cause, so the $PM_{2.5}$ emitted by these wildfire events is very unlikely to be associated with local characteristics that influence local health outcomes. However, there is still a small possibility that lightning-caused wildfires may lead to accidental deaths. Therefore, I exclude the counties within 30km, given that I focus on wildfires with less than 1000 acres (about 4.05 km^2). In addition, most direct fire fatalities (not restricted to wildfires), about 50%-80%, are caused by smoke inhalation and not burns (Holstege, 2019; NFPA, n.d.). According to the NOAA's NCEI Storm Events Database, as discussed above, the number of direct injuries and fatalities from wildfires is relatively small. The fatalities caused by wildfire smoke inhalation are also attributed to direct fatalities in this dataset.

Previous studies also applied wildfire-related IVs. Khawand (2015) simulated $PM_{2.5}$ resulting from large-size wildfires and used it as the instrumental variable to estimate the $PM_{2.5}$-related health impacts and found that wildfires contribute to at least 15% of ambient ground-level $PM_{2.5}$. Tan-Soo (2017) constructed a wind- and distance-based forest fire hotspots instrument for $PM_{2.5}$. Based on their works, I restrict wildfires to be those caused by the exogenous source, lightning. Also, I address the concern that other pollutants (other than $PM_{2.5}$) emitted by wildfires may also be carried by wind and affect local health outcomes. Three pollutants (particulate matter, ozone, and carbon monoxide) are major concerns during wildfire smoke events (Stone et al., 2019). First, $PM_{2.5}$ is the primary threat to public health[38] (Stone et al., 2019). Second, carbon monoxide dilutes rapidly, so it is rarely a concern unless people are in very close proximity to the wildfires (Stone et al., 2019). Thus, carbon monoxide is unlikely to travel to the focal county and influence local health. Third, ozone is not emitted from a wildfire, but forms in the plume as wildfire smoke moves downwind (Stone et al., 2019). Thus, ozone can be another channel through which distant wildfires influence local health. Therefore, I consider the contribution of wildfires on $PM_{2.5}$ and address the concern of other air pollutants by using the predicted $PM_{2.5}$ emitted by distant lightning-caused wildfires.

In addition to excluding the nearby counties and using the exogenous natural cause, lightning, other factors used to construct instrumental variables are the exogenous wind direction and geographic distance. These factors further ensure the exogeneity of the instrumental variables. To further increase confidence in these instrumental variables, following the methodology described in Wooldridge (2010), I conduct an overidentification test, which is also applied by

[38] Particles from wildfire smoke tend to be very small (with a size range near the wavelength of visible light (0.4-0.7 μm)), and about 90% of total particle masses consist of $PM_{2.5}$ (U.S. EPA, 2019a), so $PM_{10-2.5}$ (PM_{10} is comprised of $PM_{2.5}$ and $PM_{10-2.5}$) is not a major concern.

Wrenn et al. (2017). Section 4 presents the overidentification test results, which indicate that I cannot reject the hypothesis that the instrumental variables are exogenous.

I, therefore, use the instrumental variables and employ the control function approach. In the first stage, I use the fixed-effect model and regress the endogenous variable, air pollution, on distant $PM_{2.5}$, distant lightning-caused wildfires, all the other exogenous variables contained in the previous models, and county and time-fixed effects. This regression can break the correlation between the endogenous variable and unobservable factors affecting the dependent variables (Wooldridge, 2010). The first-stage regression equation is:

$$PM_{it} = \theta_0 + \theta_f Fire_{it} + \boldsymbol{\theta_{IV} IV_{it}} + \boldsymbol{\theta} \cdot \boldsymbol{X_{it}} + C_i + \ T_t + \varepsilon_{it} \quad (12)$$

The fixed effect residuals ($\widehat{\varepsilon_{it}}$) of the first-stage regression can be regarded as the "control" for the endogeneity of air pollution (Wooldridge, 2015). I then add $\widehat{\varepsilon_{it}}$ into equation (6) as the second-stage regressions:

$$E(Deaths_{it} | Fire_{it}, PM_{it}, \boldsymbol{X}_{it}, C_i, T_t)$$
$$= C_i \cdot \exp(\beta_D{}' Fire_{it} + \beta_P{}' PM_{it} + \beta_r \widehat{\varepsilon_{it}} \ + \boldsymbol{\beta}' \cdot \boldsymbol{X_{it}} + \ T_t \ ) \quad (13)$$

Since the residuals depend on the estimates of parameters in the first-stage regression, the variance matrix estimators for parameters in the second-stage regression need to be adjusted to account for this dependence (Wooldridge, 2010). A block-bootstrap procedure (500 repetitions) draws from the entire FIPS code with replacement to correct the standard errors (Wooldridge, 2010; Schlenker and Walker, 2016). I examine the endogeneity of air pollution by testing whether the coefficient of residuals is equal to zero. Using the two-stage control function approach, the indirect effect equals $\theta_f * \beta_P{}'$. Again, since I use the Fixed-effect Poisson model, $\theta_f * \beta_P{}'$ is just the approximation of indirect impact. The formula for the indirect effect of one additional wildfire event (at least 100 acres) on fatalities is $(exp\,(\theta_f * \ \beta_P{}') - 1) * 100\%$, which can be approximated by $\theta_f * \beta_P{}'$.

Following the methodology described in Wooldridge (2010) and applied by Wrenn et al. (2017), I add one of two IVs to the right-hand side of the second stage regression equation (i.e., equations (9) and (10), and then perform a significance test for this IV. The test result is invariant to which subset of IVs I add. I include the distant $PM_{2.5}$ to the second stage regression, as presented in equation (11). To correct the standard errors, as before, I use the block-bootstrap procedure (500 repetitions).

$$E(Deaths_{it}|Fire_{it}, PM_{it}, \boldsymbol{X_{it}}, C_i, T_t) \\ = C_i \cdot \exp\left(\varphi_h Fire_{it} + \varphi_p PM_{it} + \varphi_{IV} DistPM_{it} + \varphi_r \widehat{\varepsilon_{it}} + \boldsymbol{\varphi} \cdot \boldsymbol{X_{it}} + T_t\right) \quad (14)$$

3.2.3 Spillover Effect of Wildfires

To study the spillover effects, I need to construct variables representing how many wildfires occur in upwind counties. Therefore, by combing the monthly wildfire and wind direction data, I calculate the wildfire occurrences or acres in the upwind counties. Since the effect of wildfire on local air quality decreases as travel distance increases, I use the reciprocal of geographic distance as the decay factor. The upwind wildfires are defined as follows:

$$UpwindFires_{it} = \sum_{m=1}^{12} \sum_{i \neq j} Fires_{jmt} * I\ (cos\theta_{ijmt} > 0) * \frac{1}{d_{ij}} \quad (15).$$

where d_{ij} is the distance between the counties *i* and *j*. θ_{ijmt} is the angle between the vector from county *j* to *i* and the vector of wind direction in county *j* in month *m* year *t*. I consider two types of weights based on wind direction and geographic distance. The upwind wildfires in county *i* at year *t*, $UpwindFires_{it}$, is weighted by the reciprocal of geographic distance (km), d_{ij}, and I consider the wildfire events only when the angle between the vector of monthly wind direction and the vector of geographic location (i.e., θ_{ijmt}) is less than 90 degrees (i.e., I($cos\theta_{i3mt} > 0) = 1$). Since the impact of wildfires on air pollution may be temporary, I consider the occurrences and wind direction in each month and sum up the weighted occurrences to generate the annual weighted occurrences of wildfires. Similarly, wildfires can be measured using the number of annual wildfire occurrences with different fire sizes and the annual wildfire acres. I examine how the upwind wildfires influence the local $PM_{2.5}$ first, as shown in Equation (8), and the spillover effect of upwind wildfires equals $\gamma_1{}' * \beta_P$ ($\gamma_1{}' * \beta_P{}'$ if considering the endogeneity issue).

$$PM_{it} = \gamma_0' + \gamma_1' Fire_{it} + \gamma_2' UpwindFires_{it} + \boldsymbol{\gamma'} \cdot \boldsymbol{E_{it}} + T_t + C_i + u_{it} \quad (16)$$

3.2.4 Distinguishing Wildfire Causes

Previous studies usually assume that wildfire is exogenous and do not consider the potential endogeneity of wildfire-caused air pollution when studying the health impacts of wildfires and wildfire smoke. In the previous section, I control for the percentage of lightning-caused wildfires. This section tests whether there are significant differences in the health impacts with and without considering the causes of wildfires. The null hypothesis is that the indirect impacts of wildfires should not be significantly different with and without including the percentage of lightning-caused wildfires. To check the robustness, I include the percentage of the wildfires caused by powerlines and railroads among the human-caused wildfires and test whether the indirect impacts will significantly differ under these three specifications.

4. Results

4.1 Direct and Indirect Impacts of Local Wildfires

The baseline results using mediation analysis are presented in **Table 2. 3**. To examine the delayed impact of wildfires on suicide, I add a one-year lag term for wildfires and the percentage of lightning-caused wildfires. The estimation results on the all-cause deaths and deaths caused by suicide are presented in **Table 2. 4**. I then estimate the increases in fatalities due to a one standard deviation (SD) (≈2.56) increase in wildfires events (≥ 100 acres), as shown in **Table 2. 5**.

I first examine the total impact of wildfires on mortality, which is positive and significant for all the models, as shown in Panel A of **Table 2. 3** (please see **Table 2. 9** of the Appendix for the complete results). Note that a significant total impact is not necessary to establish the (mediation) indirect effect. Given that the data covers about 98.5% population in the U.S., the results represent the population of the whole country. The annual average of all-cause deaths, deaths caused by respiratory system disease, deaths caused by circulatory system disease, and the deaths caused by suicide in the U.S. from 1999 to 2019 are around 2,543,934, 247,023, 850,241, and 37,756, respectively. Based on the estimates of total impact, I estimate the additional deaths due to a one standard deviation increase in wildfires events, as presented in **Table 2. 5**. A one standard deviation (≈2.56) increases in wildfires (with at least 100 acres) events occur, there will be an additional 7,835, 1,077, 3,052, and 97 fatalities in each of the above underlying fatality categories.

Next, as presented in Panel B of **Table 2. 3**, I find statistically significant and positive associations between wildfires and $PM_{2.5}$, regardless of what measurements of wildfires are used (please see **Table 2. 10** of the Appendix for the complete results). As the wildfire size increases, the magnitude of wildfire impact increases significantly. Also, note that the percentage of the lightning-caused (human-caused) wildfires is positively (negatively) associated with $PM_{2.5}$, especially when I focus on larger-size wildfires. This result suggests that it is necessary to control the causes of wildfires because the percentage of lightning-caused (human-caused) wildfires is correlated with factors that influence the health outcomes, such as economic activity and local government effectiveness.

Panel C of **Table 2. 3** presents the direct and indirect impacts of wildfires. The complete results can be found in Appendix. Since I only focus on the channel of air pollution, the coefficient for wildfires includes both direct impacts of wildfires and indirect impacts through other channels (if they exist). First, according to the first-stage regression results, the F-statistic is about 438.7 (see Appendix for complete results), so the instrumental variables are strong predictors of local $PM_{2.5}$. In addition, the overidentification tests show that I cannot reject the null hypothesis that the instrumental variables are exogenous.

Second, I find evidence for the endogeneity of $PM_{2.5}$ in the models for all-cause deaths, deaths caused by respiratory system diseases, and deaths caused by circulatory system diseases since the residuals in the above three models are negative and significant. This result also indicates that the omitted factors associated with $PM_{2.5}$ are negatively associated with mortality. Therefore, I focus on the results using the control function approach for the three models, as presented in columns 2, 4, and 6. For suicide, I focus on the result using the Fixed-effect Poisson regression, as presented in column 7.

Third, this paper distinguishes between the adverse health impact directly from wildfires and indirectly from wildfire-induced $PM_{2.5}$. For all-cause deaths, deaths caused by respiratory system diseases, and deaths caused by circulatory system diseases, the mediation (indirect) effects of wildfire-induced $PM_{2.5}$ are statistically significant and positive. After addressing the endogeneity problem of $PM_{2.5}$, the magnitudes of indirect effect increase substantially, whereas the magnitudes of coefficients of wildfires decrease. Since it is very unlikely that people who died from underlying causes of respiratory and circulatory system diseases are killed in the wildfires at the same time, wildfires are expected to have no direct impact on mortality caused by respiratory and circulatory system diseases. Thus, the significant and positive coefficients of wildfires indicate that there may be other potential channels through which wildfires influence human health. Based on the work of Zhao et al. (2010) and Wen and Ye (2014), I conclude that there is a complementary mediation, and there should be other channels through which wildfires influence human health. This question will be explored in future research. Similarly, I estimate the additional fatalities due to the increases of $PM_{2.5}$ emitted by a one standard deviation (≈2.56) additional wildfires (with at least 100 acres) events. The additional wildfires are indirectly associated with 4,554, 506, and 655 additional all-cause deaths, deaths caused by respiratory system diseases, and deaths caused by circulatory system diseases, respectively. In addition, the indirect impacts of wildfires via $PM_{2.5}$ take up about 58.12%, 46.98%, and 21.46%, respectively, of the total impact of wildfires on the three categories of mortality.

For suicide, although the coefficient of wildfires is significant at a 10% significance level (column 7, Panel C of **Table 2. 3**), the indirect impact of wildfire-induced $PM_{2.5}$ that is tested using percentile bootstrap is not significant. Combing the results of the significant total effect of wildfires on suicide (column 4, Panel A of **Table 2. 3**), I conclude that wildfires should directly influence the risks of suicide or through wildfire-induced environmental goods other than $PM_{2.5}$. Further, under the assumption that $PM_{2.5}$ is the only channel and that the total impact of wildfires only includes the direct impact and the indirect effect of $PM_{2.5}$, I also explore the delayed direct effect of wildfires and whether the direct effect of wildfires varies with wildfire size. I include the lag terms of wildfires and the percentage of lightning-caused wildfires to identify the potential longer-term impact of wildfires on suicide, as presented in column 2 of **Table 2. 4**. Wildfires that occurred in the previous year have a larger and more significant impact on suicide fatalities, increasing all-cause deaths. This result is consistent with the literature that suicidal behavior may be delayed after natural disasters and should be monitored in the longer term. I then change the size of wildfires. Instead of using wildfires with at least 100 acres, I consider the larger-size wildfires, the wildfires with at least 10 acres, 300 acres, and 1000 acres. As shown in column 3 of **Table 2. 4**, when I include smaller-size wildfires, the direct impact decreases substantially and becomes insignificant. In contrast, if I only focus on larger wildfires, the direct impacts on suicide are larger and more significant (see columns 4 and 5 of **Table 2. 4**). Therefore, wildfire size matters; larger wildfires significantly increase the risk of suicide. The complete results regarding suicide are presented in Appendix.

4.2 Spillover Effects of Wildfires

As presented in **Table 2. 6**, I find that $PM_{2.5}$ emitted by wildfires (≥ 100 acres) in other counties is often transported by wind and influences air quality in the focal county, thus leading to higher mortality. According to the results in Panel A, there are statistically significant positive impacts of upwind wildfires on local $PM_{2.5}$ for all the measurements I use for the upwind wildfires. Meanwhile, the impacts of local wildfires are still significant and positive, except for the total occurrences of wildfires. Since the total occurrences include some extremely small wildfires (0~0.25 acres) and these small wildfires' contributions to $PM_{2.5}$ are trivial, the total number of occurrences is a less preferred measurement. Please see Appendix for complete results on the spillover and local effects of wildfires on air pollution.

The spillover effects are presented in Panel B of **Table 2. 6**. There are positive and significant spillover effects of upwind wildfires for all the mortality categories, except for mortality caused by suicide. I then compare the spillover and local effects. I calculate the percentage changes of the mortality due to a one standard deviation change in local and upwind wildfires, the elasticities evaluated at mean levels of local and upwind wildfires, and the percentage change in mortality due to change in local and upwind wildfires from 2001 to 2011. Regardless of which method I use, I find consistent and substantially larger impacts of upwind wildfires than local wildfires. This result suggests that distant wildfires can pose substantially more significant detrimental impacts on human health through $PM_{2.5}$ emissions and transportation. Therefore, wildfire management is essential for local governments as well as broader regional or national governments.

4.3 Robustness Analysis

This section conducts a series of robustness checks on the construction of instrumental variables and the necessity of considering wildfire causes. First, I use different spatial ranges for the instrumental variables. In the baseline analysis, I chose counties at least 100 km away from the focal county to construct the distant all-source $PM_{2.5}$ and counties within a range of 30-100 km away to construct the $PM_{2.5}$ attributed to distant wildfires. In this section, I examine the 80 km and 150 km for distant $PM_{2.5}$ and the ranges of 30-80 km and 30-150 km for $PM_{2.5}$ attributed to distant wildfires. The further the distant sources of $PM_{2.5}$, the more likely it is for the exogeneity condition to hold, but the weaker the association between local and distant $PM_{2.5}$. From the first-stage results in **Table 2. 12** of the appendix, I find that the t-statistics and the magnitudes of coefficients for distant $PM_{2.5}$ decrease as the radius of the first buffer zone increases. Meanwhile, as the radius of the first buffer zone increases, the range of the second buffer zone increases, which means that more lightning-caused wildfires are included. Then, the t-statistics of distant wildfires increase, and the impacts on the local $PM_{2.5}$ decrease. However, regardless of which spatial range I choose, the F-statistics for distant $PM_{2.5}$ and distant wildfires are large at more than 350, and the t-statistics for each instrumental variable are greater than 4.5. It means that the IVs are strong predictors of local $PM_{2.5}$. Applying the overidentification tests, I cannot reject the null hypothesis that the IVs constructed using different spatial ranges are exogenous. Further, I find that the estimation results using IVs of different spatial distances only change slightly.

Second, in addition to changing the spatial ranges, I also examine robustness using lightning-caused wildfires of different sizes. In the original analysis, I choose the occurrences of lightning-caused wildfires with 10 ~ 999 acres. I also consider the burning acres of all lightning-caused wildfires and the occurrences of the lightning-caused wildfires with 0.26~299 acres. From the first-stage results in **Table 2. 12** of the appendix, I find that when the more enormous distant wildfires are included, the impacts of distant wildfires on local $PM_{2.5}$ and the t-statistics increase (columns 1 and 5). The joint test of distant $PM_{2.5}$ and distant wildfires are statistically significant, and the F-statistics are over 400 for all the specifications. In addition, the IVs using different sizes of wildfires pass the overidentification test. Overall, the results are robust, and there are only slight variations in the coefficient estimates.

Third, given that Florida suffered a significantly higher level of imported $PM_{2.5}$ caused by lightning-caused wildfires, I conducted a robustness examination excluding the samples of Florida. Overall, the results are robust for all-cause mortality, mortality caused by respiratory system diseases, and mortality caused by circulatory system diseases, and the magnitudes only change slightly.

Last, I use percentile bootstrap to examine the differences in estimates with and without considering wildfire causes. According to **Table 2. 8**, I find that the magnitudes of indirect impacts on all-cause deaths and deaths caused by respiratory system diseases are overestimated if the percentage of lightning-caused wildfires is not included (Panels 1 vs. 2). Based on the main analysis, I further control the percentage of the wildfires caused by powerlines and railroads among the human-caused wildfires. Again, the magnitudes of indirect impacts for the above two mortality categories are overestimated (Panels 1 vs. 3). In addition, there are no significant differences between controlling for only one category of wildfire cause (Panel 2) and controlling more than one category (Panel 3). This evaluation suggests that it is necessary to consider wildfire causes, and it should be sufficient to distinguish between lightning-caused and human-caused wildfires.

Table 2. 3 Mediation Analysis Results

Panel A	Step 1: Total Effect of Wildfires on Mortality							
Dependent variable: Deaths	All-cause (1)		Respiratory (2)		Circulatory (3)		Suicide (4)	
Fire Occurrences	0.0012***	-	0.0017***	-	0.0014***	-	0.0010*	-
(≥ 100 acres)	(5.39)		(5.55)		(4.24)		(1.80)	
% Lighting-caused	0.00003	-	0.00003	-	0.00006	-	-0.00007	-
Wildfires(≥ 100 acres)	(0.80)		(0.57)		(0.96)		(-0.64)	

Panel B	Step 2: Effects of Wildfires on Air Pollution						
Dependent variable: $PM_{2.5}$	Fire acres (in 10 thousand)	Total occurrences	Fires (≥ 0.26 acres)	Fires (≥ 10.0 acres)	Fires (≥ 100 acres)	Fires (≥ 300 acres)	Fires (≥ 1000 acres)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Wildfires	0.0464***	0.0008***	0.0024***	0.0089***	0.0291***	0.0505***	0.0878***
	(3.84)	(3.22)	(7.34)	(8.45)	(6.15)	(5.39)	(4.24)
% Lighting-caused	0.0010***	0.0004	0.0006	0.0012***	0.0014***	0.0017***	0.0016***
Wildfires	(3.42)	(0.98)	(1.35)	(3.67)	(4.34)	(4.33)	(3.17)

Table 2. 3 (cont'd)

Panel C	Step 3: Direct and Indirect Effects of Wildfires on Mortality							
Dependent variable:	All-cause		Respiratory		Circulatory		Suicide	
Deaths	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE	CF	FE	CF	FE	CF	FE	CF
Fire Occurrences	0.0012***	0.0007***	0.0017***	0.0010**	0.0014***	0.0012***	0.0010*	0.0004
(≥ 100 acres)	(5.78)	(2.59)	(5.32)	(2.18)	(4.32)	(2.86)	(1.93)	(0.50)
% Lighting-caused	0.00002	-0.000003	0.00002	-0.000007	0.00005	0.00004	-0.00008	-0.00011
Wildfires(≥ 100 acres)	(0.48)	(-0.09)	(0.42)	(-0.12)	(0.90)	(0.69)	(-0.75)	(-0.97)
PM	0.0058***	0.0226***	0.0037**	0.0280***	0.0019**	0.0110**	0.0051	0.0260*
	(7.18)	(5.44)	(2.37)	(3.66)	(2.09)	(2.40)	(1.53)	(1.68)
Residuals		-0.0176***		-0.0253***		-0.0095**		-0.0217
		(-4.52)		(-3.29)		(-2.13)		(-1.41)
Indirect effects (ab)	0.00017***	0.00068***	0.00011**	0.00084***	0.00006**	0.00033**	0.00015	0.00078
χ^2 statistics for	-	0.05	-	0.05	-	0.62	-	1.35
overidentification test	-	[0.8234]	-	[0.8223]	-	[0.4308]	-	[0.2446]
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y	Y	Y
N	32912	32912	32912	32912	32912	32912	32890	32890

Note: (1) t statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) Standard errors of regressions using the control function (CF) approach (columns 2, 4, 6, and 8 in Panel C) are obtained from the block-bootstrap (500 repetitions) procedure (drawing the entire FIPS code with replacement).

Table 2. 4 Additional Results for Suicide

Deaths	(1) All-cause (Fire≥ 100 acres)	(2) Suicide (Fire≥ 100 acres)	(3) Suicide (Fire≥ 10 acres)	(4) Suicide (Fire≥ 300 acres)	(5) Suicide (Fire≥ 1000 acres)
Fire Occurrences	0.0011^{***}	0.0008	0.00003	0.0022^{**}	0.0031^{**}
	(5.42)	(1.46)	(0.08)	(2.52)	(2.37)
L. Fire Occurrences	0.0006^{***}	0.0014^{***}			
	(3.36)	(3.00)			
% Lighting-caused Wildfires	0.00002	-0.0001	-0.0002^{*}	-0.0002^{**}	-0.0001
	(0.63)	(-0.61)	(-1.82)	(-1.96)	(-1.04)
L. % Lighting-caused Wildfires	0.00002	0.0001			
	(0.71)	(0.80)			
PM	0.0060^{***}	0.0056^{*}	0.0050	0.0052	0.0050
	(7.59)	(1.68)	(1.50)	(1.58)	(1.50)
Indirect effects (ab)	0.0002^{***}	0.0002	0.00004	0.0003	0.0004
L. Indirect effects (L. ab)	0.00002	0.00002			
Time FE	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y
N	32912	32890	32890	32890	32890

Note: (1) t statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2. 5 Increases in Deaths Due to A One SD (≈2.56) Additional Wildfires Event (⩾ 100 acres)

Deaths	Average Deaths (1999~2019)	Total Effects			Indirect Effects (Wildfire-induced $PM_{2.5}$)			% of The Channel of $PM_{2.5}$
		Coefficient	ΔPercentage	Δ Deaths	Coefficient	ΔPercentage	ΔDeaths	Percentage
All-cause	2,543,934	0.0012***	0.308%	7,835	0.0007***	0.179%	4,554	58.12%
Respiratory	247,023	0.0017***	0.436%	1,077	0.0008***	0.205%	506	46.98%
Circulatory	850,241	0.0014***	0.359%	3,052	0.0003**	0.077%	655	21.46%
Suicide	37,756	0.0010*	0.256%	97	-	-	-	-

Note: (1) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (2) Wildfires are measured by the number of wildfire occurrences (at least 100 acres). (3) ΔPercentage=(exp(coefficient*2.56)-1)*100%.

Table 2. 6 Comparisons of Spillover and Local Effects of Wildfires

Panel A	Effects of Wildfires on Air Pollution						
Dependent variable:	Fire acres (in 10 thousand)	Total occurrences	Fires (≥ 0.26 acres)	Fires (≥ 10.0 acres)	Fires (≥ 100 acres)	Fires (≥ 300 acres)	Fires (≥ 1000 acres)
$PM_{2.5}$	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Local Wildfires	0.0408***	-0.00004	0.0010***	0.0052***	0.0164***	0.0305***	0.0606***
	(3.49)	(-0.17)	(2.99)	(4.72)	(3.43)	(3.21)	(2.92)
Upwind Wildfires	0.5073***	0.0132***	0.0169***	0.0524***	0.2421***	0.5142***	1.0410***
	(9.00)	(21.52)	(22.63)	(19.85)	(21.47)	(21.54)	(18.56)

Table 2. 6 (cont'd)

Panel B	**Indirect Effects of Wildfires (≥ 100 acres) on Mortality**							
Dependent variable:	**All-cause**		**Respiratory**		**Circulatory**		**Suicide**	
Deaths	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Indirect effects (ab)	**FE**	**CF**	**FE**	**CF**	**FE**	**CF**	**FE**	**CF**
Source of $PM_{2.5}$ (Wildfires)								
Local	0.0001^{***}	0.0004^{***}	0.0001^{**}	0.0005^{**}	0.00003^{*}	0.0002^{**}	0.0001	0.0004
Distant Upwind	0.0014^{***}	0.0055^{***}	0.0009^{**}	0.0068^{***}	0.00047^{**}	0.0027^{**}	0.0012	0.0063^{*}
	Percentage changes of deaths due to a one standard deviation of local wildfires and upwind wildfires (≈2.56 & 0.83, respectively)							
Local	0.026%	0.102%	0.026%	0.128%	0.008%	0.051%	0.026%	0.102%
Distant Upwind	0.116%	0.457%	0.075%	0.564%	0.039%	0.224%	0.100%	0.523%
	Elasticities evaluated at the data means (≈0.84 & 1.57, respectively)							
Local	0.008%	0.034%	0.008%	0.042%	0.003%	0.017%	0.008%	0.034%
Distant Upwind	0.220%	0.864%	0.141%	1.068%	0.074%	0.424%	0.188%	0.989%
	Percentage changes of deaths due to changes in mean local and upwind wildfires from 2001 to 2011(≈0.44 & 1.08, respectively)							
Local	0.004%	0.018%	0.004%	0.022%	0.001%	0.009%	0.004%	0.018%
Distant Upwind	0.151%	0.594%	0.097%	0.734%	0.051%	0.292%	0.130%	0.680%
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y	Y	Y
N	32912	32912	32912	32912	32912	32912	32890	32890

Note: (1) t statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) Standard errors of regressions using the control function (CF) approach (columns 2, 4, 6, and 8 in Panel B) are obtained from the block-bootstrap (500 repetitions) procedure (drawing the entire FIPS code with replacement).

Table 2. 7 Robustness Check

	Main Analysis	Robustness Check			
	DistPM (≥100km) & *DistPM_Fire* (30~100km, 10 ~ 999 acres)	*DistPM_Fire* (≥100 acres)		*DistPM* (≥100km) & *DistPM_Fire* (30~100km)	
		DistPM (≥80km) & *DistPM_Fire* (30~80km)	*DistPM* (≥150km) & *DistPM_Fire* (30~150km)	*DistPM_Fire* (All lightning-caused fire acres)	*DistPM_Fire* (0.26~299 acres)
Dependent Variable: All-cause Deaths					
Fire Occurrences (≥ 100 acres)	0.0007***	0.0008***	0.0007**	0.0007***	0.0007**
	(2.59)	(2.67)	(2.46)	(2.67)	(2.56)
PM	0.0226***	0.0221***	0.0235***	0.0226***	0.0226***
	(5.44)	(5.51)	(5.22)	(5.68)	(5.35)
Residuals	-0.0176***	-0.0170***	-0.0183***	-0.0175***	-0.0175***
	(-4.52)	(-4.58)	(-4.34)	(-4.73)	(-4.45)
Indirect effects (ab)	0.0007***	0.0007***	0.0007***	0.0007***	0.0007***
χ^2 statistics	0.05	0.04	0.00	0.01	0.04
	[0.8234]	[0.8405]	[0.9969]	[0.9078]	[0.8457]
Dependent Variable: Deaths Caused by Respiratory System Disease					
Fire Occurrences (≥ 100 acres)	0.0010**	0.0010**	0.0009*	0.0010**	0.0010**
	(2.18)	(2.28)	(1.95)	(2.15)	(2.20)
PM	0.0280***	0.0269***	0.0308***	0.0292***	0.0277***
	(3.66)	(3.66)	(3.69)	(3.87)	(3.60)
Residuals	-0.0253***	-0.0242***	-0.0280***	-0.0265***	-0.0250***
	(-3.29)	(-3.28)	(-3.34)	(-3.51)	(-3.23)
Indirect effects (ab)	0.0008***	0.0008***	0.0009***	0.0009***	0.0009***
χ^2 statistics	0.05	0.09	0.00	0.71	0.04
	[0.8223]	[0.7683]	[0.9771]	[0.3978]	[0.8393]

Table 2. 7 (cont'd)

	Dependent Variable: Deaths Caused by Circulatory System Disease				
Fire Occurrences (≥ 100 acres)	0.0012***	0.0012***	0.0011***	0.0012***	0.0012***
	(2.86)	(2.92)	(2.77)	(2.95)	(2.85)
PM	0.0110**	0.0104**	0.0119**	0.0112***	0.0108**
	(2.40)	(2.39)	(2.30)	(2.58)	(2.31)
Residuals	-0.0095**	-0.0089**	-0.0103**	-0.0097**	-0.0093**
	(-2.13)	(-2.10)	(-2.06)	(-2.29)	(-2.05)
Indirect effects (ab)	0.0003**	0.0003**	0.0004**	0.0003**	0.0003**
χ^2 statistics	0.62	0.48	0.54	0.53	0.77
	[0.4308]	[0.4882]	[0.4606]	[0.4666]	[0.3817]
F-statistics for instruments	438.7	478.2	365.4	435.6	438.4
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
N	32912	32912	32912	32912	32912

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) p-values of test statistics in squared brackets. (4) Standard errors are obtained from the block-bootstrap (500 repetitions) procedure (drawing the entire FIPS code with replacement). (5) Time-fixed effects and county-fixed effects are included for all the model specifications. (6) The χ^2 statistics of overidentification tests are based on Wooldridge (2010), and the instrumental variable for distant wildfires is excluded. (7) Since the endogeneity problem is not found in the model of suicide, the results are not included here.

Table 2. 8 Distinguish Wildfire Causes

Dependent variable: Deaths	All-cause		Respiratory		Circulatory		Suicide	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	FE	CF	FE	CF	FE	CF	FE	CF
	Panel 1: Not Include % Lighting-caused Wildfires							
Indirect effects	0.00018***	0.00073***	0.00012**	0.00090***	0.00006**	0.00035**	0.00016	0.00083
	Panel 2: Include % Lighting-caused Wildfires							
Indirect effects	0.00017***	0.00068***	0.00011**	0.00084***	0.00006**	0.00033**	0.00015	0.00078
	Panel 3: Include % Lighting-caused Wildfires + % Wildfires Caused by Powerline and Railroad							
Indirect effects	0.00017***	0.00067***	0.00011**	0.00084***	0.00006**	0.00033**	0.00015	0.00080
	Test differences – Panel 1 vs. Panel 2							
Indirect effects	0.00001***	0.00005***	0.00001**	0.00006***	0.00001*	0.00002	0.00001	0.00005
(z-statistics)	(3.42)	(3.10)	(2.05)	(2.64)	(1.72)	(1.57)	(0.98)	(1.27)
% change	5.6%	6.8%	8.3%	6.7%	16.7%	5.7%	12.5%	6.0%
	Test differences – Panel 1 vs. Panel 3							
Indirect effects	0.00001***	0.00006***	0.00001**	0.00006**	0.00001*	0.00002	0.00001	0.00003
(z-statistics)	(3.54)	(3.36)	(1.98)	(2.30)	(1.77)	(1.50)	(0.81)	(0.62)
% change	5.6%	8.2%	8.3%	6.7%	16.7%	5.7%	6.7%	3.6%
	Test differences – Panel 2 vs. Panel 3							
Indirect effects	0.00000	0.00001*	-0.00000	-0.00000	0.00000	0.00000	-0.00000	-0.00002
(z-statistics)	(0.61)	(1.98)	(-0.38)	(-0.11)	(0.29)	(0.60)	(-0.64)	(-0.90)
% change	0%	1.4%	0%	0%	0%	0%	0%	2.4%
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y	Y	Y
N	32912	32912	32912	32912	32912	32912	32890	32890

Note: (1) t statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) Standard errors of regressions using the control function (CF) approach (columns 2, 4, 6, and 8) are obtained from the block-bootstrap (500 repetitions) procedure (drawing entire FIPS code with replacement).

5. Conclusions

With the trend of more frequent extreme weather events, the U.S. is expected to experience an increasing number of wildfires. Although many studies have explored the detrimental effects of wildfires, the mechanisms through which wildfires influence human health physically and psychologically have not been fully explored. Clarifying the mechanisms is useful because wildfires occur in many terrestrial ecosystems and induce various environmental changes that influence human health. Previous research typically focused on wildfires or wildfire-induced air pollution but did not determine the degree to which wildfires or wildfires-induced environmental changes cause health impacts. To meet the challenge of increasing wildfire exposure, clarifying these channels helps policymakers allocate resources more effectively and enact related policies to mitigate adverse direct and indirect wildfire impacts. For direct health impacts, the statistics of direct injuries and fatalities do not reflect the mental health impacts of wildfires in the longer run. The result indicates that wildfires, especially large-size wildfires, directly increase suicide with a delayed impact. Therefore, allocating additional resources for longer-term mental health services after large-size wildfires may help to save lives.

This study also finds that wildfire-induced $PM_{2.5}$ is positively associated with all-cause mortality, mortality caused by circulatory system diseases, and mortality caused by respiratory system diseases. Through applying the mediation analysis, I confirm that $PM_{2.5}$ can be regarded as a mediator through which local wildfires influence human health. In addition, compared to local wildfires, distant wildfires play a substantially more detrimental role in human health via emitting $PM_{2.5}$. In addition, wildfire-induced $PM_{2.5}$ plays a significant role in the total adverse impacts of wildfires on mental health. I find that the detrimental health impacts of wildfires on suicide may only stem from the occurrence of large-size wildfire events and earlier wildfire events. In contrast to the evidence of adverse mental health impact of all-source $PM_{2.5}$ in the literature, I find no evidence that wildfire-induced $PM_{2.5}$ leads to higher suicide risk.

With the increased fire suppression costs in recent decades, the government faces a tradeoff between suppression costs and the costs of larger burned areas. This study provides the policymaker with estimates of health losses from wildfires and a reference for future research on the cost-benefit analysis of wildfire suppression. In addition, the estimates of direct and indirect health impacts help the policymaker set priorities for post-fire management, such as post-fire forest restoration, post-fire mental health services, and post-fire air pollution control. According to the ecological evidence, there appear to be multiple channels through which wildfires affect human health, such as water pollution and post-fire flooding. The result provides indirect evidence of the existence of these other channels and offers a reference for future research. In previous studies, wildfires are typically treated as exogenous events, but human activities cause most wildfires. The present study considers the cause of wildfires, finding that the risks of wildfire-induced air pollution may be overestimated if researchers fail to consider the causes of wildfires. Last, due to data limitations, I used annual mortality data, and thus wildfires and wildfire-induced air pollution are also measured at the annual level. Future work is needed to explore the short- and medium-term effects of wildfires by applying daily, weekly, or monthly data.

REFERENCES

Aguilera, R., Corringham, T., Gershunov, A., & Benmarhnia, T. (2021). Wildfire smoke impacts respiratory health more than fine particles from other sources: Observational evidence from Southern California. *Nature Communications*, *12*(1), 1–8.

Alan, S., Ertac, S., & Mumcu, I. (2018). Gender stereotypes in the classroom and effects on achievement. *Review of Economics and Statistics*, *100*(5), 876–890.

Bakian, A. V., Huber, R. S., Coon, H., Gray, D., Wilson, P., McMahon, W. M., & Renshaw, P. F. (2015). Acute air pollution exposure and risk of suicide completion. *American Journal of Epidemiology*, *181*(5), 295–303.

Banzhaf, H. S., & Chupp, B. A. (2010). *Heterogeneous harm vs. spatial spillovers: Environmental federalism and US air pollution* (Working Paper No. 15666). National Bureau of Economic Research. https://doi.org/10.3386/w15666

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Barwick, P. J., Li, S., Rao, D., & Zahur, N. B. (2018). *The morbidity cost of air pollution: Evidence from consumer spending in China* (SSRN Scholarly Paper No. 2999068). https://doi.org/10.2139/ssrn.2999068

Bayer, P., Keohane, N., & Timmins, C. (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management*, *58*(1), 1–14. https://doi.org/10.1016/j.jeem.2008.08.004

Braithwaite, I., Zhang, S., Kirkbride, J. B., Osborn, D. P., & Hayes, J. F. (2019). 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*(12), 126002.

Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Siscovick, D., Smith, S. C., Whitsel, L., Kaufman, J. D., & American Heart Association Council on Epidemiology and Prevention, Council on the Kidney in Cardiovascular Disease, and Council on Nutrition, Physical Activity and Metabolism. (2010). Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation*, *121*(21), 2331–2378. https://doi.org/10.1161/CIR.0b013e3181dbece1

Brown, M. R., Agyapong, V., Greenshaw, A. J., Cribben, I., Brett-MacLean, P., Drolet, J., McDonald-Harker, C., Omeje, J., Mankowsi, M., Noble, S., & others. (2019). Significant PTSD and other mental health effects present 18 months after the Fort McMurray wildfire: Findings from 3,070 grades 7–12 students. *Frontiers in Psychiatry*, 623.

Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., & Wara, M. (2021). The changing risk and burden of wildfire in the United States. *Proceedings of the National Academy of Sciences*, *118*(2), e2011048118. https://doi.org/10.1073/pnas.2011048118

Burkhardt, J., Bayham, J., Wilson, A., Berman, J. D., O'Dell, K., Ford, B., Fischer, E. V., & Pierce, J. R. (2020). The relationship between monthly air pollution and violent crime across the United States. *Journal of Environmental Economics and Policy*, *9*(2), 188–205.

Burkhardt, J., Bayham, J., Wilson, A., Carter, E., Berman, J. D., O'Dell, K., Ford, B., Fischer, E. V., & Pierce, J. R. (2019). The effect of pollution on crime: Evidence from data on particulate matter and ozone. *Journal of Environmental Economics and Management*, *98*, 102267.

Caamano-Isorna, F., Figueiras, A., Sastre, I., Montes-Martínez, A., Taracido, M., & Piñeiro-Lamas, M. (2011). Respiratory and mental health effects of wildfires: An ecological study in Galician municipalities (north-west Spain). *Environmental Health*, *10*(1), 1–9.

Cameron, A. C., & Trivedi, P. K. (2013). *Regression Analysis of Count Data*. Cambridge University Press.

Chen, S., Chen, Y., Lei, Z., & Tan-Soo, J.-S. (2021). Chasing clean air: Pollution-induced travels in China. *Journal of the Association of Environmental and Resource Economists*, *8*(1), 59–89. https://doi.org/10.1086/711476

Chen, X., & Ye, J. (2019). When the wind blows: Spatial spillover effects of urban air pollution in China. *Journal of Environmental Planning and Management*, *62*(8), 1359–1376. https://doi.org/10.1080/09640568.2018.1496071

Dawson, J. P., Bloomer, B. J., Winner, D. A., & Weaver, C. P. (2014). Understanding the meteorological drivers of U.S. particulate matter concentrations in a changing climate. *Bulletin of the American Meteorological Society*, *95*(4), 521–532.

DeWolfe, D. J. (2000). *Training manual for mental health and human service workers in major disasters*. US Department of Health and Human Services, Substance Abuse and Mental ….

Dittrich, R., & McCallum, S. (2020). How to measure the economic health cost of wildfires–A systematic review of the literature for northern America. *International Journal of Wildland Fire*, *29*(11), 961–973.

Doerr, S. H., & Santín, C. (2016). Global trends in wildfire and its impacts: Perceptions versus realities in a changing world. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1696), 20150345.

Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., & Samet, J. M. (2006). Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *JAMA*, *295*(10), 1127–1134. https://doi.org/10.1001/jama.295.10.1127

Gładka, A., Rymaszewska, J., & Zatoński, T. (2018). Impact of air pollution on depression and suicide. *Int J Occup Med Environ Health*, *31*(6), 711–721.

Hammer, M. S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A. M., Hsu, N. C., Levy, R. C., Garay, M. J., Kalashnikova, O. V., Kahn, R. A., Brauer, M., Apte, J. S., Henze, D. K., Zhang, L., Zhang, Q., Ford, B., Pierce, J. R., & Martin, R. V. (2020). Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). *Environmental Science & Technology*, *54*(13), 7879–7890. https://doi.org/10.1021/acs.est.0c01764

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., & Kaufman, J. D. (2013). Long-term air pollution exposure and cardio- respiratory mortality: A review. *Environmental Health*, *12*(1), 43. https://doi.org/10.1186/1476-069X-12-43

Holstege, C. P. (2019). *Smoke Inhalation.* EMedicineHealth. https://www.emedicinehealth.com/smoke_inhalation/article_em.htm

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, *105*(4), 765–789.

Institute for Health Metrics and Evaluation (IHME). (2013). *United States Physical Activity and Obesity Prevalence by County 2001-2011* [Data set]. Institute for Health Metrics and Evaluation (IHME). https://doi.org/10.6069/5E84-HD26

Institute for Health Metrics and Evaluation (IHME). (2014). *United States Smoking Prevalence by County 1996-2012* [Data set]. Institute for Health Metrics and Evaluation (IHME). https://doi.org/10.6069/H6ZD-KX77

Jones, B. A. (2018). Willingness to pay estimates for wildfire smoke health impacts in the US using the life satisfaction approach. *Journal of Environmental Economics and Policy*, *7*(4), 403–419.

Khawand, C. (2015). *Air quality, mortality, and perinatal health: Causal evidence from wildfires* (No. pkh318). Job Market Papers. https://ideas.repec.org/p/jmp/jm2015/pkh318.html

Kochi, I., Donovan, G. H., Champ, P. A., & Loomis, J. B. (2010). The economic cost of adverse health effects from wildfire-smoke exposure: A review. *International Journal of Wildland Fire*, *19*(7), 803–817.

Kõlves, K., Kõlves, K. E., & De Leo, D. (2013). Natural disasters and suicidal behaviours: A systematic literature review. *Journal of Affective Disorders*, *146*(1), 1–14.

Lin, G.-Z., Li, L., Song, Y.-F., Zhou, Y.-X., Shen, S.-Q., & Ou, C.-Q. (2016). The impact of ambient air pollution on suicide mortality: A case-crossover study in Guangzhou, China. *Environmental Health*, *15*(1), 1–8.

Liu, J. C., Wilson, A., Mickley, L. J., Dominici, F., Ebisu, K., Wang, Y., Sulprizio, M. P., Peng, R. D., Yue, X., Son, J.-Y., & others. (2017). Wildfire-specific fine particulate matter and risk of hospital admissions in urban and rural counties. *Epidemiology (Cambridge, Mass.)*, *28*(1), 77.

Lu, J. G., Lee, J. J., Gino, F., & Galinsky, A. D. (2018). Polluted morality: Air pollution predicts criminal activity and unethical behavior. *Psychological Science*, *29*(3), 340–355.

Luechinger, S. (2010). Life satisfaction and transboundary air pollution. *Economics Letters*, *107*(1), 4–6. https://doi.org/10.1016/j.econlet.2009.07.007

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, *1*(4), 173–181.

McDermott, B. M., Lee, E. M., Judd, M., & Gibbon, P. (2005). Posttraumatic stress disorder and general psychopathology in children and adolescents following a wildfire disaster. *The Canadian Journal of Psychiatry*, *50*(3), 137–143.

Miller, N., Molitor, D., & Zou, E. (2017). Blowing smoke: Health impacts of wildfire plume dynamics. *Environmental And Resource Economics At The University Of Illinois*.

Mocko, D., & NASA/GSFC/HSL. (2012). *NLDAS Primary Forcing Data L4 Monthly 0.125 x 0.125 degree V002. 2012* [Data set]. Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5067/Z62LT6J96R4F

National Fire Protection Association (NFPA). (n.d.). *Reporter's Guide: The consequences of fire*. National Fire Protection Association. Retrieved February 22, 2021, from https://www.nfpa.org/News-and-Research/Publications-and-media/Press-Room/Reporters-Guide-to-Fire-and-NFPA/Consequences-of-fire

Neary, D. G., Ryan, K. C., & DeBano, L. F. (2005). Wildland fire in ecosystems: Effects of fire on soils and water. *Gen. Tech. Rep. RMRS-GTR-42-Vol. 4. Ogden, UT: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. 250 p., 42*.

North America Land Data Assimilation System (NLDAS). (2012). *Daily Air Temperatures and Heat Index, years 1979-2011* [Data set]. CDC WONDER Online Database. https://wonder.cdc.gov/EnvironmentalData.html

Pace, N., Sebastian, A., Daidone, S., Prifti, E., & Davis, B. (2022). Mediation analysis of the impact of the Zimbabwe Harmonized Social Cash Transfer Programme on food security and nutrition. *Food Policy*, *106*, 102190.

Papadatou, D., Giannopoulou, I., Bitsakou, P., Bellali, T., Talias, M. A., & Tselepi, K. (2012). Adolescents' reactions after a wildfire disaster in Greece. *Journal of Traumatic Stress*, *25*(1), 57–63.

Pasnau, R. (1989). Stress and psychiatry. *Comprehensive Textbook of Psychiatry*.

Psarros, C., Theleritis, C., Economou, M., Tzavara, C., Kioulos, K. T., Mantonakis, L., Soldatos, C. R., & Bergiannaki, J.-D. (2017). Insomnia and PTSD one month after wildfires: Evidence for an independent role of the "fear of imminent death." *International Journal of Psychiatry in Clinical Practice*, *21*(2), 137–141. https://doi.org/10.1080/13651501.2016.1276192

Sandberg, D. V. (2003). *Wildland fire in ecosystems: Effects of fire on air*. US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Shi, X. (2022). Moving out but not for the better: Health consequences of interprovincial rural-urban migration in China. *Health Economics*, *31*(4), 555–573.

Short, K. C. (2017). *Spatial wildfire occurrence data for the United States, 1992-2015 [FPA_FOD_20170508] (4th Edition)* [Data set]. Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2013-0009.4

Steadman, R. G. (1984). A universal scale of apparent temperature. *Journal of Climate and Applied Meteorology*, *23*(12), 1674–1687.

Stone, S. L., Anderko, L., Berger, M., Butler, C., Cascio, W., Clune, A., & others. (2019). Wildfire smoke: A guide for public health officials. *US Environmental Protection Agency*.

Tan-Soo, J.-S. (2018). Valuing Air Quality in Indonesia Using Households' Locational Choices. *Environmental and Resource Economics*, *71*(3), 755–776. https://doi.org/10.1007/s10640-

017-0182-z

United States Department of Health and Human Services (US DHHS), Centers for Disease Control and Prevention (CDC), & National Center for Health Statistics (NCHS). (2020). *Underlying Cause of Death 1999-2019* [Data set]. CDC WONDER Online Database. https://wonder.cdc.gov/ucd-icd10.html

Urbanski, S. P., Hao, W. M., & Baker, S. (2008). Chemical composition of wildland fire emissions. *Developments in Environmental Science, 8*, 79–107.

U.S. Bureau of Economic Analysis. (n.d.-a). *CAGDP1 Gross Domestic Product (GDP) summary by county and metropolitan area* [Data set]. U.S. Bureau of Economic Analysis. Retrieved July 13, 2020, from https://apps.bea.gov/itable/iTable.cfm?ReqID=70&step=1

U.S. Bureau of Economic Analysis. (n.d.-b). *CAINC30 Economic profile by county and metropolitan area* [Data set]. U.S. Bureau of Economic Analysis. Retrieved December 27, 2020, from https://apps.bea.gov/itable/iTable.cfm?ReqID=70&step=1

U.S. Bureau of Labor Statistics. (n.d.). *Consumer Price Index retroactive series using current methods (R-CPI-U-RS)* [Data set]. U.S. Bureau of Labor Statistics. Retrieved February 13, 2021, from https://www.bls.gov/cpi/research-series/r-cpi-u-rs-home.htm

U.S. Census Bureau. (n.d.-a). *Annual County Resident Population Estimates by Age, Sex, Race, and Hispanic Origin: April 1, 2010 to July 1, 2019* [Data set]. U.S. Census Bureau. Retrieved August 13, 2020, from https://www.census.gov/data/datasets/time-series/demo/popest/2010s-counties-detail.html

U.S. Census Bureau. (n.d.-b). *Intercensal Estimates of the Resident Population by Five-Year Age Groups and Sex for Counties: April 1, 2000 to July 1, 2010* [Data set]. U.S. Census Bureau. Retrieved September 15, 2020, from https://www.census.gov/data/datasets/time-series/demo/popest/intercensal-2000-2010-counties.html

U.S. Census Bureau. (n.d.-c). *Small Area Income and Poverty Estimates (SAIPE) Program online database* [Data set]. U.S. Census Bureau. Retrieved February 11, 2020, from https://www.census.gov/programs-surveys/saipe/data/datasets.html

U.S. Environmental Protection Agency. (2019). *Integrated Science Assessment (ISA) for Particulate Matter (Final Report, Dec 2019)* (EPA/600/R-19/188). U.S. Environmental Protection Agency. https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=347534

U.S. Environmental Protection Agency. (2020). *Policy Assessment for the Review of the National Ambient Air Quality Standards for Particulate Matter*. U.S. Environmental Protection Agency. https://www.epa.gov/sites/production/files/2020-01/documents/final_policy_assessment_for_the_review_of_the_pm_naaqs_01-2020.pdf

Van Donkelaar, A., Martin, R. V., Li, C., & Burnett, R. T. (2019). Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environmental Science & Technology*, *53*(5), 2595–2611.

Venn, T. J., & Calkin, D. E. (2011). Accommodating non-market values in evaluation of wildfire management in the United States: Challenges and opportunities. *International Journal of Wildland Fire*, *20*(3), 327–339.

Viswanathan, S., Eria, L., Diunugala, N., Johnson, J., & McClean, C. (2006). An analysis of effects of San Diego wildfire on ambient air quality. *Journal of the Air & Waste Management Association*, *56*(1), 56–67.

Wen, Z., & Ye, B. (2014). Analyses of mediating effects: The development of methods and models. *Advances in Psychological Science*, *22*(5), 731.

Williams, A. M., & Phaneuf, D. J. (2019). The morbidity costs of air pollution: Evidence from spending on chronic respiratory conditions. *Environmental and Resource Economics*, *74*(2), 571–603.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources*, *50*(2), 420–445.

Wrenn, D. H., Klaiber, H. A., & Newburn, D. A. (2017). Confronting price endogeneity in a duration model of residential subdivision development. *Journal of Applied Econometrics*, *32*(3), 661–682.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., & others. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres*, *117*(D3). https://doi.org/10.1029/2011JD016048

Xu, B., Yao, T., Liu, X., & Wang, N. (2006). Elemental and organic carbon measurements with a two-step heating-gas chromatography system in snow samples from the Tibetan Plateau. *Annals of Glaciology*, *43*, 257–262. https://doi.org/10.3189/172756406781812122

Yang, J., & Zhang, B. (2018). Air pollution and healthcare expenditure: Implication for the benefit of air pollution control in China. *Environment International*, *120*, 443–455.

Zhang, H., Wang, Y., Park, T.-W., & Deng, Y. (2017). Quantifying the relationship between extreme air pollution events and extreme weather events. *Atmospheric Research*, *188*, 64–79.

Zhang, X., Zhang, X., & Chen, X. (2017). Happiness in the air: How does a dirty sky affect mental health and subjective well-being? *Journal of Environmental Economics and Management*, *85*, 81–94.

Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*(2), 197–206.

Zheng, S., Cao, J., Kahn, M. E., & Sun, C. (2014). Real estate valuation and cross-boundary air pollution externalities: Evidence from Chinese cities. *The Journal of Real Estate Finance and Economics*, *48*(3), 398–414.

Zheng, S., Zhang, X., Sun, W., & Lin, C. (2019). Air pollution and elite college graduates' job location choice: Evidence from China. *The Annals of Regional Science*, *63*(2), 295–316.

APPENDIX

Table 2. 9 Total Effect of Wildfires on Mortality

	(1)	(2)	(3)	(4）
Deaths	**All-cause**	**Respiratory**	**Circulatory**	**Suicide**
Fire Occurrences	0.0012^{***}	0.0017^{***}	0.0014^{***}	0.0010^{*}
(≥ 100 acres)	(5.39)	(5.55)	(4.24)	(1.80)
% Lighting-caused Wildfires	0.00003	0.00003	0.00006	-0.00007
(≥ 100 acres)	(0.80)	(0.57)	(0.96)	(-0.64)
Heat Index	0.0022^{***}	0.0066^{***}	0.0030^{***}	-0.0004
	(4.89)	(6.29)	(4.72)	(-0.23)
Sunlight	-0.0016	-0.0074^{**}	-0.0045^{*}	0.0028
	(-0.86)	(-2.05)	(-1.87)	(0.37)
Precipitation	-0.0009	-0.0005	-0.0008	0.0013
	(-0.81)	(-0.24)	(-0.60)	(0.29)
Population	0.0023^{***}	0.0007	0.0021^{***}	0.0020^{***}
	(4.40)	(1.04)	(3.91)	(3.14)
% Under 20	0.0262^{***}	0.0346^{***}	0.0167^{***}	0.0081
	(7.64)	(7.71)	(5.34)	(1.16)
% Over 64	0.0493^{***}	0.0510^{***}	0.0467^{***}	0.0219^{***}
	(15.12)	(11.77)	(15.43)	(3.42)
% White	-0.0039^{*}	0.0042	-0.0040^{**}	-0.0037
	(-1.90)	(1.58)	(-2.01)	(-1.09)
GDP	-0.0317^{*}	-0.0245	-0.0315^{**}	0.0228
	(-1.95)	(-1.05)	(-2.00)	(0.75)
% Urban Population	0.0111^{***}	0.0104^{***}	0.0101^{***}	0.0131^{***}
	(15.44)	(9.39)	(12.51)	(7.27)
% Obesity (Female)	0.0008	0.0041^{***}	-0.0010	-0.0026
	(0.86)	(2.87)	(-0.98)	(-1.17)
% Obesity (Male)	0.0058^{***}	0.0064^{***}	0.0038^{***}	0.0046^{*}
	(6.50)	(4.17)	(3.26)	(1.91)
% Smoking Prevalence	0.0021^{**}	0.0030^{*}	-0.0003	-0.0044
	(2.08)	(1.72)	(-0.21)	(-1.47)
Time FE	Y	Y	Y	Y
County FE	Y	Y	Y	Y
N	32912	32912	32912	32890

Note: (1) t statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2. 10 Effects of Wildfires on Air Pollution

Dependent variable: $PM_{2.5}$	(1) Fire acres (in 10 thousand)	(2) Total occurrences	(3) Fires (≥ 0.26 acres)	(4) Fires (≥ 10.0 acres)	(5) Fires (≥ 100 acres)	(6) Fires (≥ 300 acres)	(7) Fires (≥ 1000 acres)
Wildfires	0.0464***	0.0008***	0.0024***	0.0089***	0.0291***	0.0505***	0.0878***
	(3.84)	(3.22)	(7.34)	(8.45)	(6.15)	(5.39)	(4.24)
% Lighting-caused Wildfires	0.0010***	0.0004	0.0006	0.0012***	0.0014***	0.0017***	0.0016***
	(3.42)	(0.98)	(1.35)	(3.67)	(4.34)	(4.33)	(3.17)
Heat Index	-0.0044	-0.0042	-0.0037	-0.0039	-0.0040	-0.0043	-0.0042
	(-1.01)	(-0.95)	(-0.86)	(-0.90)	(-0.91)	(-0.98)	(-0.96)
Sunlight	0.2593***	0.2586***	0.2484***	0.2462***	0.2499***	0.2548***	0.2570***
	(18.21)	(17.46)	(16.89)	(16.88)	(17.26)	(17.72)	(17.87)
Precipitation	-0.1635***	-0.1597***	-0.1504***	-0.1503***	-0.1556***	-0.1588***	-0.1621***
	(-19.86)	(-18.66)	(-17.94)	(-18.00)	(-18.83)	(-19.38)	(-19.85)
Population	-0.0016	-0.0041	-0.0054	-0.0027	-0.0019	-0.0013	-0.0014
	(-0.34)	(-0.80)	(-0.97)	(-0.56)	(-0.39)	(-0.26)	(-0.29)
% Under 20	-0.0700***	-0.0704***	-0.0698***	-0.0695***	-0.0690***	-0.0684***	-0.0677***
	(-5.62)	(-5.60)	(-5.60)	(-5.62)	(-5.59)	(-5.54)	(-5.47)
% Over 64	-0.1384***	-0.1377***	-0.1346***	-0.1354***	-0.1344***	-0.1352***	-0.1352***
	(-9.77)	(-9.66)	(-9.48)	(-9.59)	(-9.52)	(-9.60)	(-9.59)
% White	0.0108	0.0097	0.0084	0.0099	0.0104	0.0108	0.0107
	(1.09)	(0.97)	(0.84)	(1.00)	(1.05)	(1.09)	(1.08)
GDP	0.6597***	0.6585***	0.6534***	0.6557***	0.6594***	0.6640***	0.6611***
	(11.10)	(11.08)	(11.07)	(11.13)	(11.16)	(11.24)	(11.18)
% Urban Population	-0.0055**	-0.0053*	-0.0048*	-0.0048*	-0.0051*	-0.0051*	-0.0054**
	(-2.03)	(-1.93)	(-1.77)	(-1.77)	(-1.88)	(-1.89)	(-1.96)
_cons	2.3822**	2.4669**	2.6505**	2.5309**	2.4013**	2.2654*	2.2668*
	(1.98)	(2.04)	(2.20)	(2.11)	(2.00)	(1.89)	(1.89)

Table 2. 10 (cont'd)

Time FE	Y	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y	Y
N	32912	32912	32912	32912	32912	32912	32912
adj. R^2	0.516	0.512	0.514	0.515	0.516	0.516	0.517

Note: (1) t statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) All the standard errors are cluster (county)-robust.

Table 2. 11 Direct and Indirect Effects of Wildfires on Mortality

Deaths	All-cause		Respiratory		Circulatory		Suicide	
	FE	CF	FE	CF	FE	CF	FE	CF
Fire Occurrences	0.0012***	0.0007***	0.0017***	0.0010**	0.0014***	0.0012***	0.0010*	0.0004
(≥ 100 acres)	(5.78)	(2.59)	(5.32)	(2.18)	(4.32)	(2.86)	(1.93)	(0.50)
% Lighting-caused	0.00003	-0.000003	0.00003	-0.000007	0.00006	0.00004	-0.00007	-0.00011
Wildfires (≥ 100 acres)	(0.48)	(-0.09)	(0.42)	(-0.12)	(0.90)	(0.69)	(-0.75)	(-0.97)
PM	0.0058***	0.0226***	0.0037**	0.0280***	0.0019**	0.0110**	0.0051	0.0260*
	(7.18)	(5.44)	(2.37)	(3.66)	(2.09)	(2.40)	(1.53)	(1.68)
Heat Index	0.0021***	0.0022***	0.0066***	0.0067***	0.0029***	0.0030***	-0.0004	-0.0003
	(4.52)	(4.33)	(6.23)	(5.92)	(4.65)	(4.47)	(-0.23)	(-0.14)
Sunlight	-0.0037*	-0.0080***	-0.0087**	-0.0149***	-0.0053**	-0.0076***	0.0010	-0.0044
	(-1.91)	(-3.80)	(-2.34)	(-3.60)	(-2.12)	(-2.93)	(0.14)	(-0.52)
Precipitation	0.0004	0.0028***	0.0003	0.0038*	-0.0004	0.0009	0.0024	0.0053
	(0.36)	(2.63)	(0.14)	(1.67)	(-0.30)	(0.61)	(0.54)	(1.05)
Population	0.0022***	0.0023***	0.0007	0.0007	0.0020***	0.0021***	0.0019***	0.0019**
	(4.27)	(3.21)	(0.94)	(0.95)	(3.85)	(3.08)	(3.03)	(2.16)
% Under 20	0.0262***	0.0272***	0.0346***	0.0359***	0.0167***	0.0172***	0.0082	0.0094
	(7.68)	(8.30)	(7.71)	(8.04)	(5.32)	(5.80)	(1.18)	(1.31)
% Over 64	0.0492***	0.0513***	0.0509***	0.0541***	0.0467***	0.0478***	0.0219***	0.0246***
	(15.29)	(17.13)	(11.77)	(12.80)	(15.39)	(16.17)	(3.43)	(3.62)
% White	-0.0042**	-0.0044**	0.0040	0.0038	-0.0041**	-0.0042**	-0.0039	-0.0040
	(-2.05)	(-2.18)	(1.49)	(1.44)	(-2.05)	(-2.09)	(-1.16)	(-1.15)
GDP	-0.0323**	-0.0411**	-0.0252	-0.0380	-0.0316**	-0.0363**	0.0224	0.0114
	(-2.00)	(-2.53)	(-1.08)	(-1.64)	(-2.00)	(-2.29)	(0.74)	(0.39)
% Urban Population	0.0110***	0.0111***	0.0103***	0.0104***	0.0101***	0.0101***	0.0131***	0.0131***
	(15.57)	(15.72)	(9.40)	(9.26)	(12.50)	(12.81)	(7.27)	(7.15)
% Obesity (Female)	0.0007	0.0015*	0.0041***	0.0052***	-0.0010	-0.0006	-0.0025	-0.0016
	(0.84)	(1.72)	(2.86)	(3.50)	(-0.99)	(-0.63)	(-1.14)	(-0.68)
% Obesity (Male)	0.0054***	0.0051***	0.0061***	0.0057***	0.0037***	0.0035***	0.0044*	0.0039
	(6.24)	(5.68)	(4.07)	(3.69)	(3.16)	(2.83)	(1.83)	(1.59)

Table 2. 11 (cont'd)

% Smoking Prevalence	0.0015	0.0008	0.0027	0.0016	-0.0005	-0.0009	-0.0048*	-0.0057*
	(1.57)	(0.81)	(1.55)	(0.92)	(-0.37)	(-0.69)	(-1.67)	(-1.88)
Residuals		-0.0176***		-0.0253***		-0.0095**		-0.0217
		(-4.52)		(-3.29)		(-2.13)		(-1.41)
χ^2 statistics for overidentification test		0.05 [0.8234]		0.05 [0.8223]		0.62 [0.4308]		1.35 [0.2446]
Indirect effects (ab)	0.00017***	0.00068***	0.00011**	0.00084***	0.00006**	0.00033**	0.00015	0.00078
(95% Conf. Interval)	(0.00011, 0.00025)	(0.00039, 0.00105)	(0.00001, 0.00019)	(0.00037, 0.00142)	(2.75e-06, 0.00011)	(0.00005, 0.00064)	(-0.00006, 0.00034)	(-0.00015, 0.00179)
Time FE	Y	Y	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y	Y	Y
N	32912	32912	32912	32912	32912	32912	32890	32890

Note: (1) t statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. (3) Standard errors of regressions in columns (2), (4), and (6) are obtained from the block-bootstrap (1000 repetitions) procedure (drawing the entire FIPS code with replacement). (4) p-values of test statistics in squared brackets. (5) Indirect effect ab is the product of the impact of wildfires on PM and the impact of PM on fatalities, which are estimated by percentile bootstrap test of ab.

Table 2. 12 The First-stage Regression Results

Dependent variable: $PM_{2.5}$	(1) *DistPM* (≥100km) & *DistPM_Fire* (30~100km, ≥100 acres)	(2) *DistPM* (≥80km) & *DistPM_Fire* (30~80km, ≥100 acres)	(3) *DistPM* (≥150km) & *DistPM_Fire* (30~150km, ≥100 acres)	(4) *DistPM* (≥100km) & *DistPM_Fire* (30~100km, All lightning-caused fire acres)	(5) *DistPM* (≥100km) & *DistPM_Fire* (30~100km, 0.26~299 acres)
DistPM_Fire	16.8333***	17.7112***	12.8862***	46.5938***	11.0356***
	(5.11)	(4.64)	(5.80)	(4.30)	(3.60)
DistPM	0.0984***	0.1013***	0.0920***	0.0982***	0.0981***
	(29.59)	(30.91)	(27.00)	(29.49)	(29.57)
Fire Occurrences (≥ 100 acres)	0.0301***	0.0304***	0.0297***	0.0302***	0.0309***
	(6.37)	(6.45)	(6.27)	(6.45)	(6.56)
Heat Index	-0.0162***	-0.0167***	-0.0151***	-0.0160***	-0.0162***
	(-3.79)	(-3.93)	(-3.51)	(-3.77)	(-3.81)
Sunlight	0.2514***	0.2544***	0.2472***	0.2561***	0.2549***
	(17.17)	(17.43)	(16.84)	(17.66)	(17.40)
Precipitation	-0.1452***	-0.1455***	-0.1459***	-0.1491***	-0.1470***
	(-17.77)	(-17.80)	(-17.82)	(-18.32)	(-18.01)
Population	-0.0044	-0.0046	-0.0039	-0.0044	-0.0045
	(-0.86)	(-0.89)	(-0.76)	(-0.86)	(-0.87)
% Under 20	-0.0624***	-0.0617***	-0.0637***	-0.0621***	-0.0622***
	(-5.08)	(-5.03)	(-5.15)	(-5.05)	(-5.06)
% Over 64	-0.1262***	-0.1254***	-0.1276***	-0.1264***	-0.1264***
	(-9.30)	(-9.27)	(-9.33)	(-9.31)	(-9.30)
% White	0.0009	0.0006	0.0016	0.0011	0.0009
	(0.09)	(0.06)	(0.16)	(0.12)	(0.09)
GDP	0.6323***	0.6305***	0.6349***	0.6286***	0.6322***
	(10.84)	(10.83)	(10.83)	(10.81)	(10.85)

Table 2. 12 (cont'd)

% Urban Population	-0.0057**	-0.0057**	-0.0057**	-0.0058**	-0.0058**
	(-2.14)	(-2.16)	(-2.13)	(-2.18)	(-2.18)
% Obesity (Female)	-0.0397***	-0.0395***	-0.0402***	-0.0394***	-0.0396***
	(-7.36)	(-7.34)	(-7.40)	(-7.30)	(-7.33)
% Obesity (Male)	0.0223***	0.0224***	0.0223***	0.0227***	0.0224***
	(3.63)	(3.65)	(3.60)	(3.68)	(3.63)
% Smoking Prevalence	0.0353***	0.0349***	0.0364***	0.0352***	0.0355***
	(5.38)	(5.34)	(5.52)	(5.37)	(5.42)
% Lighting-caused Wildfires(≥ 100 acres)	0.0014***	0.0015***	0.0014***	0.0014***	0.0015***
	(4.42)	(4.51)	(4.28)	(4.51)	(4.62)
_cons	3.1791**	3.1507**	3.1913**	3.1012**	3.1268**
	(2.56)	(2.54)	(2.56)	(2.50)	(2.52)
Time FE	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y
F-statistics for instruments	438.6988	478.2270	365.3887	435.6396	438.4049
	[0.0000]	[0.0000]	[0.0000]	[0.0000]	[0.0000]
N	32912	32912	32912	32912	32912

Note: (1) t statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2. 13 Additional Results for Suicide

Deaths	(1) All-cause (Fire≥ 100 acres)	(2) Suicide (Fire≥ 100 acres)	(3) Suicide (Fire≥ 30 acres)	(4) Suicide (Fire≥ 300 acres)	(5) Suicide (Fire≥ 1000 acres)
Fire Occurrences	0.0011***	0.0008	0.00003	0.0022**	0.0031**
(≥ 100 acres)	(5.42)	(1.46)	(0.08)	(2.52)	(2.37)
L. Fire Occurrences	0.0006***	0.0014***			
(≥ 100 acres)	(3.36)	(3.00)			
% Lighting-caused Wildfires	0. 00002	-0.0001	-0.0002*	-0.0002**	-0.0001
(≥ 100 acres)	(0.63)	(-0.61)	(-1.82)	(-1.96)	(-1.04)
L. % Lighting-caused Wildfires	0.00002	0.0001			
(≥ 100 acres)	(0.71)	(0.80)			
PM	0.0060***	0.0056*	0.0050	0.0052	0.0050
	(7.59)	(1.68)	(1.50)	(1.58)	(1.50)
Heat Index	0.0020***	-0.0007	-0.0004	-0.0005	-0.0005
	(4.41)	(-0.39)	(-0.22)	(-0.27)	(-0.27)
Sunlight	-0.0034*	0.0019	0.0017	0.0009	0.0009
	(-1.74)	(0.24)	(0.22)	(0.12)	(0.12)
Precipitation	0.0004	0.0026	0.0018	0.0022	0.0022
	(0.40)	(0.58)	(0.41)	(0.50)	(0.50)
Population	0.0022***	0.0018***	0.0018***	0.0018***	0.0019***
	(4.24)	(3.01)	(2.93)	(2.94)	(3.00)
% Under 20	0.0260***	0.0079	0.0086	0.0079	0.0081
	(7.63)	(1.14)	(1.22)	(1.14)	(1.16)
% Over 64	0.0491***	0.0218***	0.0217***	0.0217***	0.0216***
	(15.26)	(3.42)	(3.40)	(3.42)	(3.38)
% White	-0.0042**	-0.0040	-0.0039	-0.0040	-0.0039
	(-2.06)	(-1.18)	(-1.15)	(-1.18)	(-1.14)
GDP	-0.0342**	0.0173	0.0247	0.0213	0.0224
	(-2.12)	(0.57)	(0.81)	(0.70)	(0.73)
% Urban Population	0.0111***	0.0131***	0.0130***	0.0131***	0.0131***
	(15.59)	(7.29)	(7.25)	(7.29)	(7.28)
% Obesity (Female)	0.0007	-0.0025	-0.0025	-0.0025	-0.0025
	(0.81)	(-1.16)	(-1.13)	(-1.16)	(-1.14)
% Obesity (Male)	0.0054***	0.0043*	0.0045*	0.0043*	0.0043*
	(6.28)	(1.80)	(1.86)	(1.82)	(1.83)
% Smoking Prevalence	0.0014	-0.0050*	-0.0049*	-0.0049*	-0.0048*
	(1.48)	(-1.75)	(-1.67)	(-1.70)	(-1.67)

Table 2. 13 (cont'd)

Indirect effects (ab)	0.0002^{***}	0.0002	0.00004	0.0003	0.0004
(95% Conf. Interval)	(0.00012, 0.00025)	(-0.00004, 0.00035)	(-0.00002, 0.00010)	(-0.00008, 0.00061)	(-0.00017, 0.00107)
L. Indirect effects (L. ab)	0.00002	0.00002			
(95% Conf. Interval)	(-5.43e-06, 0 .00006)	(-7.84e-06, 0.00007)			
Time FE	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y
N	32912	32890	32890	32890	32890

Note: (1) t statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 2. 14 Spillover and Local Effects of Wildfires on Air Pollution

Dependent variable: $PM_{2.5}$	(1) Fire acres (in 10 thousand)	(2) Total occurrences	(3) Fires (≥ 0.26 acres)	(4) Fires (≥ 10.0 acres)	(5) Fires (≥ 100 acres)	(6) Fires (≥ 300 acres)	(7) Fires (≥ 1000 acres)
Wildfires	0.0408***	-0.00004	0.0010***	0.0052***	0.0164***	0.0305***	0.0606***
	(3.49)	(-0.17)	(2.99)	(4.72)	(3.43)	(3.21)	(2.92)
% Lighting-caused Wildfires	0.0009***	0.0003	0.0004	0.0012***	0.0014***	0.0016***	0.0014***
	(3.17)	(0.58)	(1.05)	(3.61)	(4.42)	(4.18)	(2.82)
Upwind Wildfires	0.5073***	0.0132***	0.0169***	0.0524***	0.2421***	0.5142***	1.0410***
	(9.00)	(21.52)	(22.63)	(19.85)	(21.47)	(21.54)	(18.56)
Heat Index	-0.0082*	-0.0029	-0.0036	-0.0041	-0.0080*	-0.0113***	-0.0118***
	(-1.90)	(-0.67)	(-0.85)	(-0.95)	(-1.85)	(-2.64)	(-2.73)
Sunlight	0.2406***	0.1962***	0.1908***	0.1996***	0.2072***	0.2293***	0.2329***
	(17.49)	(13.09)	(12.91)	(13.94)	(14.94)	(16.71)	(16.93)
Precipitation	-0.1522***	-0.1280***	-0.1143***	-0.1198***	-0.1249***	-0.1264***	-0.1345***
	(-18.17)	(-14.51)	(-13.29)	(-14.02)	(-14.75)	(-14.95)	(-16.16)
Population	-0.0003	-0.0006	-0.0014	0.0009	0.0016	0.0016	0.0007
	(-0.06)	(-0.13)	(-0.27)	(0.18)	(0.34)	(0.35)	(0.15)
% Under 20	-0.0689***	-0.0749***	-0.0750***	-0.0771***	-0.0718***	-0.0676***	-0.0645***
	(-5.70)	(-6.20)	(-6.27)	(-6.40)	(-6.19)	(-5.91)	(-5.62)
% Over 64	-0.1275***	-0.1280***	-0.1245***	-0.1253***	-0.1122***	-0.1086***	-0.1098***
	(-9.09)	(-9.30)	(-9.14)	(-9.18)	(-8.37)	(-8.13)	(-8.17)
% White	0.0105	0.0111	0.0096	0.0104	0.0109	0.0112	0.0111
	(1.09)	(1.16)	(1.02)	(1.09)	(1.18)	(1.23)	(1.21)
GDP	0.6417***	0.6157***	0.6056***	0.6158***	0.6025***	0.6059***	0.6108***
	(10.99)	(10.65)	(10.59)	(10.77)	(10.68)	(10.78)	(10.82)

Table 2. 14 (cont'd)

% Urban Population	-0.0044*	-0.0036	-0.0031	-0.0034	-0.0030	-0.0029	-0.0033
	(-1.67)	(-1.38)	(-1.21)	(-1.31)	(-1.18)	(-1.15)	(-1.29)
_cons	2.8514**	2.8731**	3.2265***	3.2162***	3.2237***	2.9501***	2.9034***
	(2.43)	(2.46)	(2.80)	(2.78)	(2.87)	(2.64)	(2.58)
Time FE	Y	Y	Y	Y	Y	Y	Y
County FE	Y	Y	Y	Y	Y	Y	Y
N	32912	32912	32912	32912	32912	32912	32912
adj. R^2	0.519	0.518	0.520	0.520	0.524	0.526	0.525

Note: (1) t statistics in parentheses.
(2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
(3) All the standard errors are cluster (county)-robust.

CHAPTER 3: THE IMPACTS OF WILDFIRES AND WILDFIRE-INDUCED AIR POLLUTION ON HOUSE PRICES IN THE UNITED STATES

1. Introduction

Wildfire risk is growing all around the world. In recent decades, the annual total burned acres and suppression expenses in the United States (U.S.) present an upward trend. According to the U.S. National Interagency Fire Center, there was an average of 73,524 wildland fires and 5,311,434 burned acres in the U.S. from 1985 to 2019, with annual total suppression costs of over $1 billion. In addition to these significant forest damage and suppression expenses, wildfires also directly result in property and infrastructure damage, injuries, fatalities, etc. The average total property damage caused by wildfires is around $403 million per year[39], based on the National Oceanic and Atmospheric Administration's (NOAA) National Centers for Environmental Information (NCEI) Storm Events Database. In 2018, 36 large-loss fires (those generating at least $10 million in property damage) occurred, totaling more than $12.91 billion in direct property damage and losses (Badger and Foley, 2019). In addition, from 1902 to 2017, 1,128 people died from wildfires (Neary and Leonard, 2019).

These wildfires-caused economic and health losses are expected to increase in the future because more people are living in or near the wildland-urban interface (WUI) and because more severe and large-scale wildfires are being caused by increasingly frequent and intense weather events, such as extreme heat and droughts. The WUI currently has a population of over 50 million people, and it grows by about one million people every three years (Burke et al., 2021). As more people reside in or near the wildland-urban interface, human-caused wildfires are more likely to occur, and people and homes in or near the WUI are more exposed to flames than in other places (Radeloff et al., 2018). Furthermore, only 29.4 percent of the sample houses in this study have never experienced wildfires within 80 kilometers in the last five years. Only 10.2 and 38.6 percent of the houses in the sample did not experience any wildfires in the nearby 80 and 30 kilometers since 1992, respectively. Therefore, why more people are choosing to move in or near the WUI?

---

[39] The average total property damages caused by wildfires is calculated based on available estimates of property damages provided by the Storm Events Database, which includes 6,331 wildfire events reported from 1999 to 2017. The database can be found at the National Oceanic and Atmospheric Administration's National Centers for Environmental Information website (https://www.ncdc.noaa.gov/stormevents/ftp.jsp). The definitions and examples of property damage caused by wildfires can be found at https://www.nws.noaa.gov/directives/sym/pd01016005curr.pdf.

In this study, I put forward and verify two possible explanations: one is that people underestimate the risks of wildfires because some indirect risks caused by wildfire-induced disamenities are ignored; the second is that although people realize the wildfire risks, they are facing a tradeoff between enjoyment of natural resources and increased wildfire risks.

Environmental factors have a significant impact on quality of life, making them key considerations when making residential location decisions. Wildfires produce immediate damage and losses, which might be regarded as an environmental hazard, lowering purchasers' willingness to pay for houses with higher wildfire risks. Previous research has found that after wildfires, homebuyers' perceptions of wildfire risks increase, and thus wildfires have a negative effect on property prices (Loomis, 2004; Mueller et al., 2009; Stetler et al., 2010; Athukorala et al., 2016; Mueller et al., 2018). The distance to the wildfire, the view of the burning area, and the frequency of the fire all determine its impact. Living near flames and having a view of burned regions was linked to higher wildfire risks and had a strong and lasting negative influence on property prices, according to Stetler et al. (2010), but they found no evidence of the impact of unseen burned areas. In contrast, analyses by Loomis (2004) showed that even if the forest fire is invisible from individual residences in the unburned area, the level of amenities in the neighboring unburned area decreased. Athukorala et al. (2016) found a negative relationship between house values and distance to a wildfire risk area, probability due to residents' discounted chance of being affected by wildfires, adequate property insurance, and self-insured actions, or a trade-off between living close to green space and living with higher wildfire risks. Furthermore, when repeated wildfires occur, the second fire's property value reduction (about 23%) is found to be greater than the first (approximately 10%) (Mueller et al., 2009). The impacts of wildfires on the house price also vary significantly across the house price distribution, with the magnitudes of the effects for the 75th quantile being much larger than the 25th quantile (Mueller and Loomis, 2014).

In addition to the immediate negative consequences, wildfires are likely to affect homebuyer behavior through their effects on local environmental attributes. As a natural disturbance that occurs in most terrestrial ecosystems, wildfires can affect the air, soil, water, fauna, flora, fuels, recreation opportunities, cultural resources, and archeology (Sandberg et al., 2002; Neary et al., 2005; Venn and Calkin, 2011; Doerr and Santín, 2016). Thus, wildfires also have an indirect impact on human health and safety, economic development, and people's enjoyment of environmental goods and services. Air pollution is one of the most severe environmental consequences of wildfires. The burning of biomass and soil-based organic matter can generate a large amount of haze and smoke, which contains complex components including greenhouse gases, photochemically reactive compounds, sulfur dioxide (SO2), particulate matter (PM), and liquids (Neary et al., 2005; Viswanathan et al., 2006; Urbanski et al., 2008). Despite the complexity of smoke's components, particulate matter is the most significant pollutant (Sandberg et al., 2002; Stone et al., 2019). Particles from wildfire smoke are typically small (with a size range near the wavelength of visible light (0.4~0.7 μm)), and fine particles ($PM_{2.5}$), which have aerodynamic diameters of less than or equal to 2.5 μm (Stone et al., 2019; U.S. EPA, 2020)[40], make up about 90% of total particle masses.

Wildfires, meanwhile, account for 10% to 20% of all primary PM emissions in the U.S. each year (U.S. EPA, 2020). Khawand (2015) and Burke et al. (2021) simulated or assessed the air pollution emitted by wildfires, indicating that they contribute roughly 15% and up to 25% of $PM_{2.5}$ in the U.S., respectively. Furthermore, PM is one of the six primary pollutants defined by the US Environmental Protection Agency (EPA). Air pollution can have a variety of negative physical and mental health implications, as well as lead to limited visibility and increased danger of traffic accidents and crime, all of which can have a negative impact on the real estate market (Linden and Rockoff, 2008; Hyslop, 2009; Brook et al., 2010; Lavín et al. 2011; Hoek et al., 2013; Dawson et al., 2014; Bakian et al.,2015; Khawand, 2015; Lin et al., 2016; Huang and Lanz, 2017; H. Zhang et al., 2017; X. Zhang et al.; 2017; Gładka et al., 2018; Lu et al., 2018; Braithwaite et al., 2019; Burkhardt et al., 2019; Fontenla et al., 2019; Stone et al., 2019; U.S. EPA, 2019; Burkhardt et al., 2020). Moreover, wildfires-specific $PM_{2.5}$ was more toxic than equal doses from other sources (or ambient $PM_{2.5}$) and was associated with a higher respiratory effect than non-wildfire $PM_{2.5}$ (Kochi et al., 2010; Dittrich and McCallum, 2020; Aguilera et al., 2021). Owing to the EPA's stringent air quality regulations in recent decades, the U.S. has experienced a downward trend in ambient $PM_{2.5}$ concentrations (1990-2014) and the number of extreme $PM_{2.5}$ days (2000-2009) (H. Zhang et al., 2017; U.S. EPA, 2020). Considering the opposite, increasing trend in wildfires, I conclude that wildfires are becoming more relatively important in the control of $PM_{2.5}$, and wildfire management and education should remain a high priority.

Wildfire is commonly thought of as an environmental nuisance, but in fact, wildfire can

[40] Sandberg et al. (2002) summarized that 90% of all smoke particles are PM_{10}, and 90% of PM_{10} is $PM_{2.5}$.

generate both positive and negative effects on the ecosystems, except for air quality (Venn and Calkin, 2011). On the downside, wildfire-induced changes in soil structure can result in decreased soil production, greater vulnerability to postfire runoff and erosion, postfire floods, and impaired water quality, among other things. Wildfires, on the positive side, boost the availability of nutrients for plant development in the short term, reduce the likelihood of epidemic insect and disease infestations, add to the novelty of the burned environment, and reduce the risk of future wildfires, etc. (Neary et al., 2005; Venn and Calkin, 2011).

These wildfire-caused ecological changes may also affect the real estate market. When a family chooses a home, they examine the surroundings, the neighborhood, and the house's features to maximize utility. However, while some households may assess the potential wildfire risk throughout the decision-making process, they may only see the direct risk of wildfires and ascribe the indirect risks to other environmental factors such as air pollution, water pollution, flooding, and so on. The primary objective of the present paper is to assess both the direct and indirect effects of wildfires on the property market, as indicated by changes in home prices, to acquire a more comprehensive understanding of wildfire costs. Also, I explore whether the tradeoff between living near nature and wildfire risks exists.

This research is based on the Hedonic Pricing Method (HPM) and applies the instrumental variable method to distinguish between the direct and indirect effects of wildfires on house prices. I focus on the most significant potential channel, air pollution (specifically, $PM_{2.5}$), based on ecological evidence that wildfire is a significant producer of particulate matter, and the major component of wildfire smoke particles is $PM_{2.5}$, as well as existing economic evidence of PM's effects on the housing market. The property market data is obtained from Zillow's Assessor and Real Estate Database (ZTRAX), and I focus on the transactions of repeat-sale homes between 2010 and 2018. The ZTRAX database is an indispensable resource for conducting this evaluation. This database provides us with a wealth of information on property transactions and assessments in the U.S. First, I can locate each property using coordinates, and the datasets cover transaction and assessment records for the entire country. These features enable us to link each property with all wildfire occurrences in the U.S. between 1992 and 2018, $PM_{2.5}$ levels, and a set of meteorological, demographic, and surrounding environment characteristics so that I can more precisely measure the impacts of environmental factors on the housing markets nationwide. Second, ZTRAX transaction records cover a long period, which provides an opportunity to construct and use a repeat-sale dataset. When I apply the hedonic price method, it is crucial to control for property characteristics, However, I am unable to include all the property characteristics because a significant portion is missing. Therefore, a repeat-sale approach controls for all the time-invariant property characteristics, thus mitigating omitted variable problems.

To address the potential endogeneity of air pollution, I adopt the instrumental variable method. A range of instrumental variables for air pollution has been proposed in previous studies. A typical instrumental variable is a policy intervention. Policy interventions, such as the U.S. Clean Air Act's (1970) nonattainment status designation and the Chinese River Huai policy, were used by Chay and Greenstone (2005) and Huang and Lanz (2017) as sources of exogenous variability in air pollution. Another instrumental variable that has been used is the externality associated with transboundary air pollution (Bayer et al., 2009, Luechinger, 2010; Khawand, 2015; Zheng et al., 2014; Barwick et al., 2018; Yang and Zhang, 2018; Williams and Phaneuf, 2019; Zheng et al., 2019; Chen et al., 2021). These researchers constructed distant air pollution as an instrumental variable for local air pollution because air pollutants can travel a long distance by wind and distant air pollution emissions are unlikely to correlate with local economic activities. In a study from Indonesia, Tan-Soo (2018) considered the influence of forest fires on local air quality and used a similar source-receptor logic to build wind- and distance-based fire hotspots as an instrument for localized $PM_{2.5}$. Following previous studies, I employ the distant large-scale natural-caused wildfires as an instrumental variable. The instrumental variable estimate measures the impact of wildfire-caused $PM_{2.5}$ on house prices.

As a prelude to the full set of results, I find that local and distant wildfires result in a house price decline through $PM_{2.5}$ emissions. There are also significant price disparities between houses located upwind and downwind from wildfires. Moreover, house prices are positively associated with the distance from the nearest wildfire and the number of days since the most recent wildfire. I also distinguish between the effects of wildfires and living near green places. Households place a higher value on homes in locations with higher vegetation coverage, but they are also aware of the risks of living near a wildland-urban interface.

I contribute to the literature in the following ways. I provide a more comprehensive nationwide perspective of wildfire effects on house prices. While existing research has extensively studied the wildfire-induced ecological impacts and the economic impacts of wildfires, the mechanisms by which wildfires affect well-being have not been fully explored. In addition, while wildfires and wildfire-induced air pollution can have an impact on the housing market, the current research tends to focus on one or the other, not fully distinguishing between direct and indirect wildfire impacts through air pollution. Further, I control for the correlation between wildfires and living near green spaces by including a variable that measures vegetation coverage and the distance between the houses and the wildland-urban interface. This study can be used by policymakers to better understand the costs of wildfire better and improve decision-making about wildfire education, wildfire suppression, post-fire restoration, and air quality management.

This study uses a national-wide dataset and creates more accurate measurements. Most wildfire research focused on one or several specific wildfire events or areas as wildfires are typically localized disamenity. Air pollution caused by wildfires, on the other hand, can travel a great distance and affect far broader areas than wildfires. As a result, the spillover effects of wildfire-induced air pollution are more fully taken into account in this study. I also linked each property to each wildfire that occurred in the U.S. contiguous states during the study period to better measure wildfire effects, considering the distance and wind patterns. In addition, the measurements for air pollution and meteorological variables are accurate to the nearby about 1 km and 14 km of each house.

This study also explores whether wind patterns and wildfire causes affect the estimation of wildfire effects. In addition to considering total wildfire occurrences, I also divide wildfires into wildfires that occurred in upwind and downwind directions. While a majority of wildfires are caused by human activity, current research often treats wildfire incidence as exogenously determined. From 1992 to 2018, around 77.5 percent of wildfire incidents in the contiguous U.S. were triggered by humans, according to the Fire Program Analysis fire-occurrence database (FPA FOD) (Short, 2021)[41]. In this paper, I control the causes of wildfires by adding the ratio of natural-caused to all-cause wildfires as well as focusing on natural-caused wildfires.

The rest of this paper is organized as follows. In the next section, I present the theoretical framework. Section three provides details on the data and methodology. Section four presents and discusses the empirical results, and the last section concludes.

[41] Natural causes account for 14.4% of wildfires, and the causes of the remaining wildfires are missing, not specified or undetermined.

2. Theoretical Framework

2.1 Hedonic Price Model

I apply the hedonic price model to estimate the marginal willingness-to-pay (MWTP) for wildfires and wildfire-induced $PM_{2.5}$. A residential house h can be described by a vector of characteristics, $Q = (q_1, q_2, \dots, q_n)$, which may include the house's structural, neighborhood, and local environmental attributes (including wildfire events and air quality). The price of a house h can be described as

$$P_h = P(q_1, q_2, \dots, q_n) \text{ (1).}$$

The partial derivative of P_h with respect to the *j*th characteristic q_j, $\partial P_h / \partial q_j$, gives the marginal implicit price of the *j*th characteristic. In a competitive market, the housing market equilibrium arises from the interactions of buyers and sellers. The marginal implicit price of the *j*th characteristic equals individuals' MWTP for the *j*th characteristic.

When a wildfire occurs, it can change the community's amenities (such as property and infrastructure damages from wildfire), altering buyers' willingness to pay and the market price of neighboring properties. For example, the charred landscape left by wildfires may take a long time to recover, making life less pleasant and potentially more difficult for residents. In addition, the expectation of the consequences of future wildfires is another component of the direct impact of wildfires. Households may be able to forecast future wildfires based on present information, thus making more informed decisions and thus locating to open areas cleared by flames or locating to lower risk areas as dictated by wind direction.

Wildfires affect many ecosystems, so local environmental characteristics such as air quality, water quality, and flood threats may change simultaneously. However, home buyers may only consider the direct consequences of wildfires in their decision-making process while attributing the wildfire-induced effects to other environmental factors. For example, if both recent wildfires ($Wildfires$) and air pollution (PM) are important environmental factors that influence the households' decisions and thus the house prices, then:

$$P_h = P(q_1, q_2, \dots, Wildfires, \ PM, q_n) \text{ (2).}$$

The partial derivative of P_h with respect to wildfires, $\partial P_h/\partial Wildfires$, only gives the direct impact of recent wildfires on house prices and fails to capture the indirect impact of wildfires through air pollution because the indirect impact is incorporated into the impact of air pollution, $\partial P_h/\partial PM$. As a result, if the assumption that air quality also has an impact on house prices holds, just examining the direct impact of wildfires, $P_h/\partial Wildfires$, will lead to a biased estimate when investigating the total effect of wildfire occurrence on house prices.

Assume that air pollution can be written as a function of wildfires, other sources that contribute to air pollution (S), and meteorological factors (X) that influence the formation and transport of air pollutants, $PM = f(Wildfires, S, X)$, so the indirect impact of wildfires through air pollution can be derived as $(\partial P_h/\partial PM) * (\partial PM/\partial Wildfire)$. Furthermore, note that wildfires can affect house prices through a variety of channels, such as water pollution, floods, reduced future risks of wildfires, and more recreation opportunities. In this case, if these factors are not considered in equation 2, $P_h/\partial Wildfires$ should include the direct impact of wildfires and indirect impacts through these channels. If households also consider these factors when purchasing a property, the sum of the direct impact, $P_h/\partial Wildfires$, and the indirect impact through air pollution $(\partial P_h/\partial PM) * (\partial PM/\partial Wildfire)$ is just a fraction of the total impact of wildfires. To distinguish the direct and indirect impacts of wildfires and explore the existence of air pollution and other channels by which wildfires influence house prices, I apply a mediation analysis approach.

Last, I can draw three hypotheses regarding the impacts of wildfires on house prices. Air pollution is one of several potential pathways by which wildfire affects house prices, but the overall effect is complex. Therefore, the expected sign of $\partial P_h/\partial Wildfires$ is ambiguous. Second, wildfires are a major source of air pollution, which can negatively affect the quality of life in nearby areas. As a result, I anticipate that wildfire-caused air pollution will have a negative indirect impact on house prices. Third, as the total impact includes both direct and indirect impacts, it is also ambiguous. With this background, I propose the following hypotheses:

Hypothesis 1: The direct impact of wildfires on house prices (or a combination of direct and indirect impacts via other routes other than air pollution) is ambiguous.

Hypothesis 2: The indirect impact of wildfire via degrading air quality on house prices is negative.

Hypothesis 3: The total impact of wildfires on housing prices is ambiguous.

2.2 Mediation Analysis

As discussed above, wildfires can affect house prices directly and indirectly through air pollution and other environmental factors. As a result, these wildfire-caused environmental factors might be thought of as mediating variables or mechanisms that influence house prices. **Figure 3. 1** depicts the relationship, which is represented by equations 3, 4, and 5. The direct impact of wildfires, c', and the indirect impact of wildfires, ab, sum to yield the total effect of wildfires on health, c. Since there may be more than one mediating variable, the total mediation effect equals the sum of all the mediation effects of each path, and the total effect equals the sum of the total mediation effect and the direct effect. Given that I only discuss the air pollution channel in this paper, if the mediation effects of other channels exist, c' should include the direct impact and the indirect impacts of other channels.

$$Y = cT + dX_1 + e_1 \quad (3)$$
$$M = aT + d'X_2 + e_2 \quad (4)$$
$$Y = c'T + bM + d''X_1 + e_3 \quad (5)$$

The most well-known and widely used method to conduct mediation analysis is the causal step approach in the classic paper of Baron and Kenny (1986). However, this approach is found to have a series of problems[42] (MacKinnon et al., 2000; Zhao et al., 2010; Wen and Ye, 2014; Hayes, 2018). Recent studies developed a causal step approach and put forward some new procedures and tests to improve the mediation analysis (MacKinnon et al., 2000; Zhao et al., 2010; Imai et al., 2011; Wen and Ye, 2014). One important recommendation is that the establishment of the mediation effect should rely on the significance of the indirect effect, ab, only, instead of the separate test of a and b, the indirect effect is examined directly by the testing H_0: $ab = 0$ (Zhao et al., 2010). Another important development is the identification of the causal mediation effect.

---

[42] (1) the inference about the indirect effect is based on the outcomes of hypothesis tests on a and b, rather than the estimate of the indirect effect ab. The separate test of a and b are of the low power; (2) this approach involves several null hypothesis tests, but only one inferential test of the indirect effect is needed; (3) whether or not the mediation effect exists is just a qualitative claim, which is not based on a quantification of the indirect effect and does not carry information on the uncertainty that can be reflected by a confidence interval; and (4) the total effect, c, does not need to be statistically significantly different from zero because of the potential suppression (inconsistent mediation or competitive mediation), the lower power of the test of total effects, and the existence of subpopulations in which the total effects have different signs, etc.

According to Imai et al (2011), “sequential ignorability” assumptions are needed to identify the causal mediation relationship, which can be written as equations (6) and (7). First, given a series of control variables, wildfire occurrence is independent of health outcomes and air pollution. Second, given a series of control variables and wildfire occurrence, air pollution is independent of health outcomes.

$$\{Y_i(m,t), M_i(t)\} \perp\!\!\!\perp T_i | X_i = x \quad (6)$$
$$Y_i(m,t) \perp\!\!\!\perp M_i(t) | T_i = t, X_i = x \quad (7)$$

If equations (6) and (7) hold, I can then estimate the average causal mediation effect through equations (4) and (5). However, in this study, equation (7) may not hold. To address this problem, Imai et al. (2011) suggested applying an instrumental variable approach. I will discuss this method in detail later in Section 3.2.

2.3 Factors influencing the Wildfire Exposures on House Prices

To measure wildfire effects more comprehensively, I consider the following perspectives: time, spatial, meteorological, and wildfire attributes, as shown in **Figure 3. 1 (a)**. These factors influence the wildfire effects on both house prices directly and air pollution transportation and thus house market indirectly. First, I consider measuring overall wildfire exposure. In the short run, wildfires can emit a large amount of smoke, destroy properties, lead to direct injuries and fatalities, etc., while the long-run wildfire-caused environmental changes can also affect local amenity levels and change people’s purchasing decisions. In addition to the time dimension, wildfire frequency, wildfire size (with at least 300 or 100 burned acres), wildfire cause, the distances between the house and wildfires, and wind patterns are important factors under consideration to measure overall wildfire exposure for each house. The details of constructing the overall wildfire exposure are discussed in Section 3.2.1. These factors influence the direct and indirect wildfire effects differently. For the direct effects, more frequent, larger-scale wildfires as well as wildfires that occurred closer to the property are more likely to lead to direct economic and health losses and thus reflect higher wildfire risks. In addition, the stronger the wind, the faster the wildfire expands. If houses are in the downwind location, the upwind wildfires are more likely to affect houses downwind. The rapid expansion of wildfires can also lead to increased levels of anxiety and thus reduce preferences for properties in downwind locations. In addition, wildfires caused by human activity and natural forces may be associated with different environmental attributes, community characteristics, etc. For example, lightning-caused wildfires may be accompanied by strong storms or drought events, which also influence preferences for living locations. Regarding the indirect impacts of wildfire via air pollution, more frequent or larger-scale wildfires can emit larger amounts of air pollution and influence property values. When wildfires occur closer to the property or upwind, the air pollution caused by wildfires is also more likely to influence property value. If the nearby wildland is ignited by lightning and there are strong storms, the indirect effects of wildfires via air pollution are expected to be stronger than wildfires caused by human activity which are less likely to be associated with wind pattern change.

In addition to the overall measurement of wildfire activity, extreme cases also make a difference in the household's decision. For example, when compared to a house that was exposed to three equal-size wildfires that occurred within 15 km over the last six months, a house that was exposed to a wildfire (same size) that occurred within 5 km and within 30 days of the transaction date may experience a more pronounced discount, even though these two properties had similar overall wildfire risks. Therefore, I also consider extreme factors: the number of days since the most recent wildfire and the distance between the house and the nearest wildfire. Therefore, I propose the following hypotheses:

Hypothesis 4-1: As the number of days since the wildfire increases, the effect of wildfire on house prices is reduced.

Hypothesis 4-2: As the distance between wildfire and house increases, the effect of wildfire on house prices is reduced.

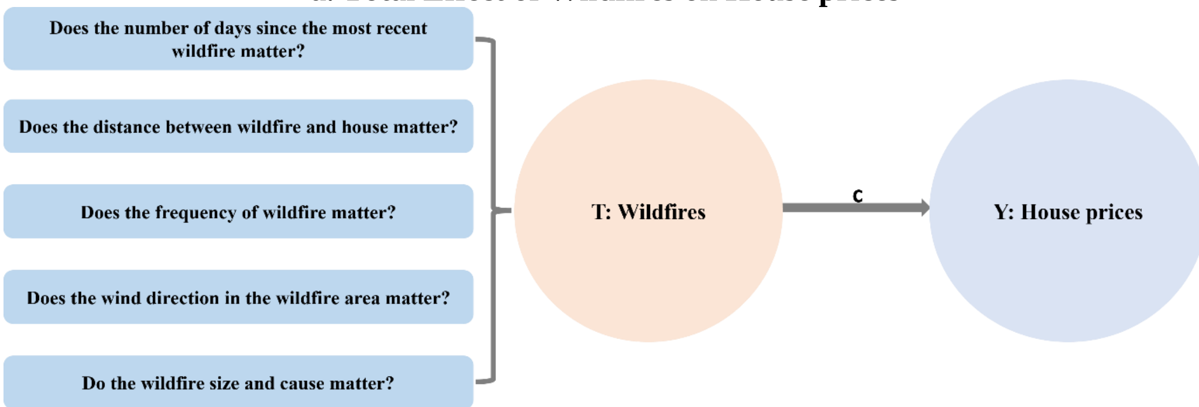
Hypothesis 4-3: As wildfire frequency/size increases, the magnitude of the wildfire effect increases.

Hypothesis 4-4: Houses located downwind of wildfires are more likely to be influenced by wildfires directly and indirectly through air pollution.

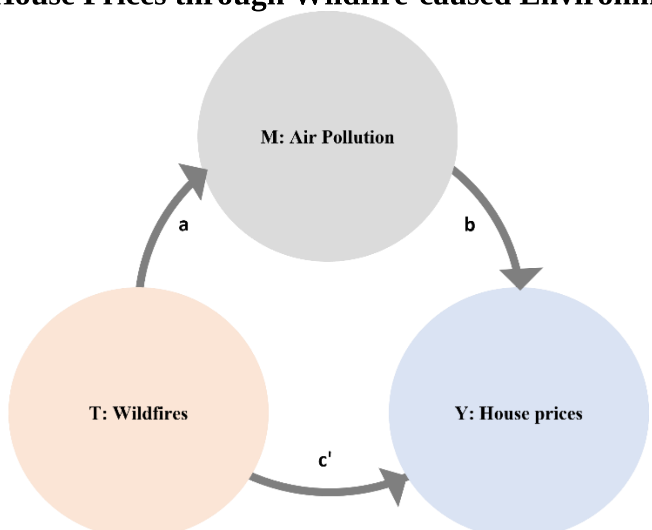
Hypothesis 4-5: Wildfires caused by human activity and natural forces affect house prices heterogeneously.

Figure 3. 1 Mediation Analysis

a. Total Effect of Wildfires on House prices



b. Wildfires Affect House Prices through Wildfire-caused Environmental (Dis)amenities



3. Data and Methods

3.1 Data

3.1.1 Property Data

This research is based on Zillow's Assessor and Real Estate Database (ZTRAX, version: 5 April 2021), which contains nationwide housing transaction and assessment records. I create a repeat sale dataset from 2010 to 2018 by merging the housing transaction and assessment datasets. I chose the years 2010 to 2018 because the nationwide housing market experienced a significant change due to the shock of the financial crises between 2007 and 2008, and the wildfire dataset I use provides data through 2018. I focus on the properties with repeat sales (properties with at least two transactions within the study period) because, on the one hand, many house characteristics are missing, and on the other hand, by using the parcel fixed effects, I can control all the time-constant house characteristics rather than control the house attribute variables, and thus mitigate the bias caused by omitted variables.

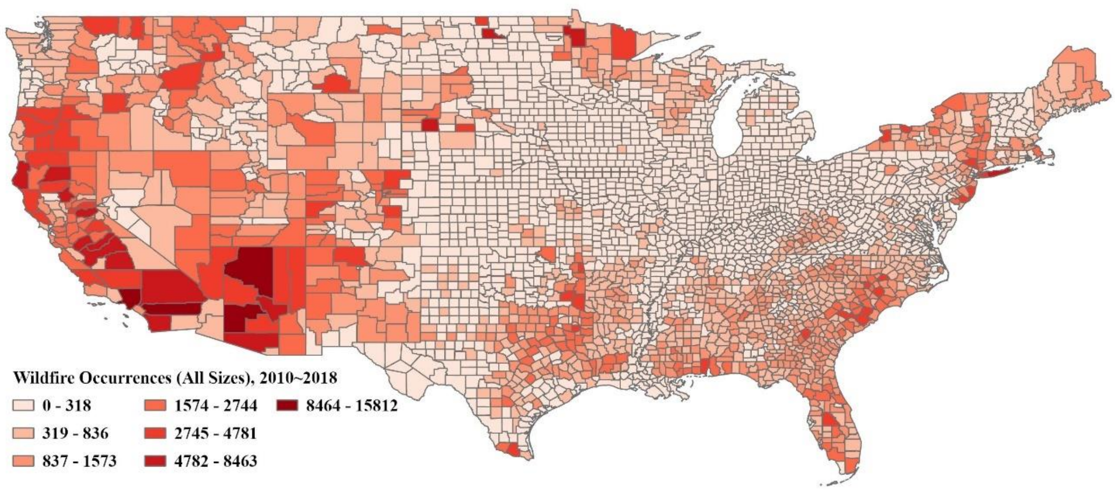
I discard transactions with missing sale prices and geographic location information and keep only deed transfers of single-family residential properties. I further eliminate transactions that are labeled as non-length arm's sales and intra-family transfers, as well as residences with frequent transactions (please see the appendix for the definition of frequent transactions), to reduce the bias created by transactions that do not reflect sales under typical conditions. Furthermore, because I use repeat sale data and want to control for parcel fixed effects, I only consider residences that have not had any remodeling, new building, or major rehabilitation between the first and last transaction dates between 2010 and 2018. Observations that report a construction year following the sale year are also dismissed. To minimize data entry errors, I trim the top and bottom 1% of the data using total rooms and lot size. I also eliminate the outliers with sale prices less than $10,000 and more than $10,000,000, as well as properties with more than 20 rooms and 100,000 square feet. The sale prices are adjusted using the monthly housing consumer price index from the U.S. Bureau of Labor and Statistics. The detailed data processing procedures and the figure of the sample distribution across the U.S. are presented in Appendix A.

3.1.2 Wildfire and Air Pollution Data

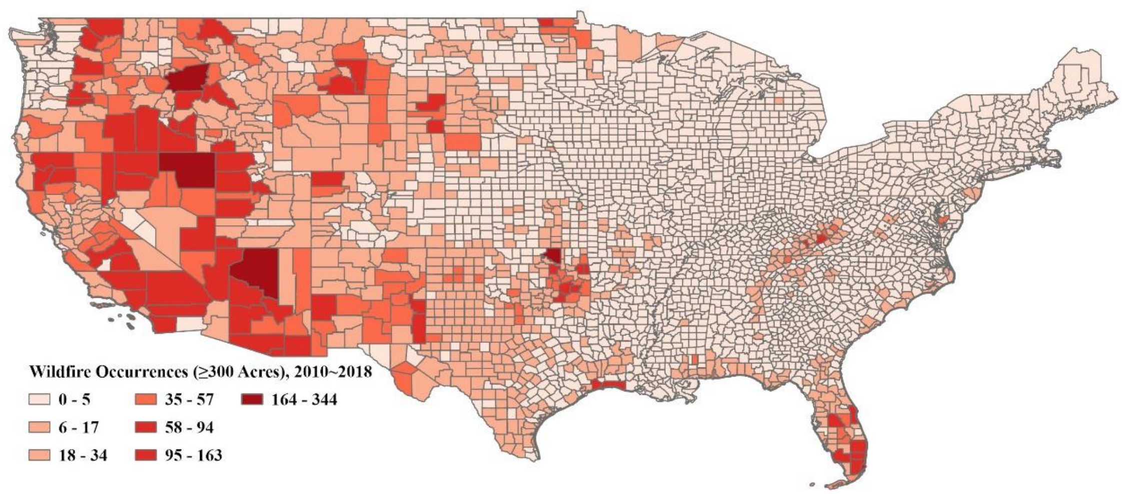
The wildfire data originates from the Fire Program Analysis fire-occurrence database (FPA FOD) (Short, 2021). This database includes 2,166,753 wildfire events from 1992 to 2018. After excluding the observations for Puerto Rico, Alaska, and Hawaii, I have 2,120,520 observations in total. I rely on the wildfire discovery date because most of the dates on which the wildfires were declared contained (or controlled) are missing. I assume that there has not been a wildfire if there is no record for a given county and specific date. The distributions of all-size wildfires and wildfires with at least 300 burned acres at the county level across the contiguous United States are shown in **Figure 3. 2**. The figures show that the wildfire incidences differ across the country, with more frequent and intense flames in the western and southern states, which is also consistent with the fact that these regions usually experience more extreme heat and drought events. The measurements of wildfires are discussed in Section 3.2 in detail.

I employ fine particulate matter ($PM_{2.5}$) to measure wildfire-induced air pollution in this article. The first reason is that particulate matter, particularly $PM_{2.5}$, is the primary pollutant of concern from wildfire smoke; $PM_{2.5}$ accounts for over 90% of the total wildfire smoke particle mass released by wildfires (Stone et al., 2019). Second, according to the U.S. EPA Integrated Science Assessment (ISA) for Particulate Matter Report (2019), causal relationships between health effects and $PM_{2.5}$ are relatively more likely to exist among the various size fractions of PM, which means that the level of $PM_{2.5}$ is more likely to influence health status and thus influence household purchase decisions.

Figure 3. 2 Distributions of Wildfires Occurrences (2010~2018)
(a) Total Wildfires Occurrences for All Wildfires Size Classes



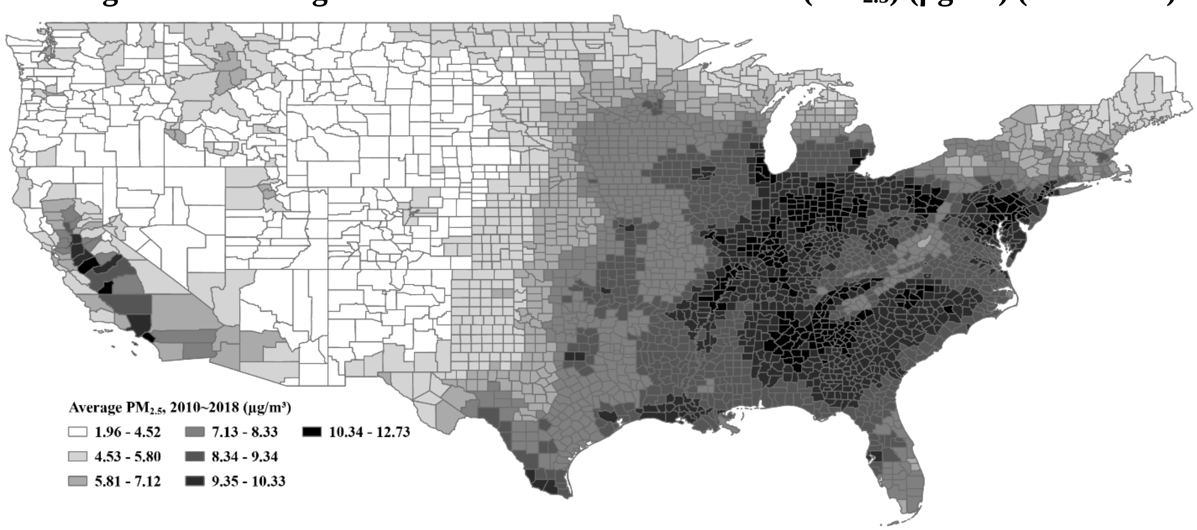
(b) Total Wildfires Occurrences for Wildfires with At Least 300 Burned Acres



Source: Authors' illustration. **Data:** Fire Program Analysis fire-occurrence database (FPA FOD) (Short, 2021).

The monthly total concentration estimates of ground-level $PM_{2.5}$ data in the United States originate from the Atmospheric Composition Analysis Group (ACAG)[43]. I spatial join each property's coordinate to the shapefiles of monthly $PM_{2.5}$ and then extract the values of the local $PM_{2.5}$ level. The monthly $PM_{2.5}$ dataset provided by ACAG is gridded at the finest solution (0.01°× 0.01°), allowing us to calculate the monthly $PM_{2.5}$ level for the property's surrounding area (about 1.11km × 1.11km). I utilize the mean level of $PM_{2.5}$ at the census tract level for properties for which I cannot extract a value from the shapefile directly. If the zonal mean at the census tract level is still unavailable, I use the county-level zonal mean of $PM_{2.5}$[44]. **Figure 3. 3** shows the distributions of mean $PM_{2.5}$ at the county level across the contiguous United States. Generally, the distribution is consistent with the U.S. EPA ISA report (2019, 2020) that the eastern areas of the country suffer higher but a more uniform level of $PM_{2.5}$ than western areas, whereas California has a significantly higher level of $PM_{2.5}$ than the surrounding states.

Figure 3. 3 Average Ground-level Particulate Matter ($PM_{2.5}$) (μg/m³) (2010-2018)



Source: Authors' illustration. **Data:** Atmospheric Composition Analysis Group (ACAG).

3.1.3 Other Data

I also control for weather, demographics, and surrounding environmental factors. The monthly data of precipitation, pressure, humidity, and temperature can be obtained from The

[43] Surface $PM_{2.5}$ dataset (North American Regional Estimates (V4.NA.03)) from Atmospheric Composition Analysis Group.

[44] About 0.0228% (902/3,948,371) observations that I cannot extract a value from the shapefiles directly.

National Aeronautics and Space Administration Phase 2 of the North American Land Data Assimilation System[45]. The monthly local precipitation, pressure, humidity, and temperature are extracted from the shapefiles using a similar processing method to $PM_{2.5}$, and if the values are missing[46], I utilize the zonal means at the census tract or county level.

Neighborhood features include vegetation coverage, the distance between the wildland-urban interface and property, population density, home density, and the ratio of white people. Families that live near or in the WUI benefit from living close to the forest, but their houses are more vulnerable to wildfires. As a result, households are facing a tradeoff between living near nature and higher wildfire risks, which may also explain why more households are living in or near the WUI. To examine the existence of the tradeoff, I control for the proportion of vegetation at the census block level as well as the distances between each property and the nearest intermix or interface WUI to assess the benefits and dangers of living near woods, which, at the same time, can help mitigate the endogeneity problem of wildfires. The 1990-2010 wildland-urban interface of the contiguous United States geospatial data (2nd Edition) contains statistics on housing and population density at the census block level for 2000 and 2010, the ratio of vegetation coverage at the census block level for 2001 and 2011, and the WUI areas for 2000 and 2010. ArcGIS is used to spatial join the coordinates of properties with the shapefile, extract the values of housing and population density and proportion of vegetation, as well as compute the distances between residences and the nearest WUI. The race data at the census block group level are obtained from IPUMS National Historical Geographic Information System (Manson et al., 2021), which originates from the U.S. Census Bureau 2000 and 2010 census data. Next, I apply the interpolation and extrapolation methods to get observations between 2010 and 2018.

To create the wildfire measurements, I also calculate the distances between wildfire centroids and properties, as well as extract wind pattern data at wildfire centroids. Distances are calculated with ArcGIS. The wind data originates from the NLDAS-2. The wind direction at the wildfire centroid is determined by the monthly zonal and meridional wind speeds. Similarly, wind speeds are extracted directly from the shapefiles.

Finally, I obtained a repeat-sale dataset between 2010 and 2018, which covers 48

[45] The weather shapefiles provided by NASA are at the resolution level of $0.125° \times 0.125°$, allowing us to calculate the monthly weather condition of the surrounding area (about 13.88 km × 13.88 km) of the property.
[46] About 0.0002% (9/3,948,371) observations that I cannot extract a value from the shapefiles directly.

contiguous states and Washington, DC, 1,834 counties. There are 3,945,340 transaction records of 1,886,684 houses[47]. About 41.33%, 28.66%, 17.20%, and 12.82% of the sample is distributed in the South, West, Midwest, and Northeast region, respectively. That is, most of the observations are in areas with more frequent wildfire occurrences. In the next section, I introduce the variables in more detail. **Table 3. 1** presents the definitions of variables used in the analysis.

Table 3. 1 List of Variables in the Empirical Analysis

Dependent Variables		
House Price	Natural logarithm of the adjusted sale price of the house	$ln(P_{hym})$
Explanatory/Control Variables		
Air Pollution	Average ground-level Particulate Matter ($PM_{2.5}$) (μg/m³)	PM_{hmy}
Distance to Wildfire	The distance between the property and the nearest wildfire (km)	$Distance_{hym}$
Local Wildfires	The weighted sum of local wildfires	$Local_{hym}$
	The weighted sum of local upwind wildfires	$Local\_Up_{hmy}$
	The weighted sum of local downwind wildfires	$Local\_Down_{hmy}$
	The number of days since the most recent wildfire (in hundred)	$Days_{hym}$
	The ratio of local wildfires caused by natural causes	$\boldsymbol{W_{hym}}$
Meteorology	Average temperature (F)	
	Average precipitation (in hundred) (kg/m²)	
	Average specific humidity (kg/kg)	
	Average surface pressure (kPa)	
Neighborhood	Housing density by census block level (housing units / km²)	
	Population density by census block level (1000 persons / km²)	
	The ratio of the white people by census block group level	
	Wildland vegetation ratio by census block level	
	Distance between house and intermix/interface WUI (km)	
Instrumental Variables		
Instruments	Distant natural-caused upwind wildfires	$DistFire\_Up_{hmy}$

[47] Since there are some singleton observations, the estimation uses 3,943,418 transaction records and covers 1,885,744 properties.

3.2 Empirical Methodology

3.2.1 Direct and Indirect Impacts of Local Wildfires

In this section, I present the empirical methods used to estimate the impacts of wildfires and wildfire-induced air pollution on house prices based on the hedonic price model. Note that I only keep those properties that sold at least twice; thus, I employ a repeat sales framework here. I first estimate the total impacts of wildfires, as presented in equation (8):

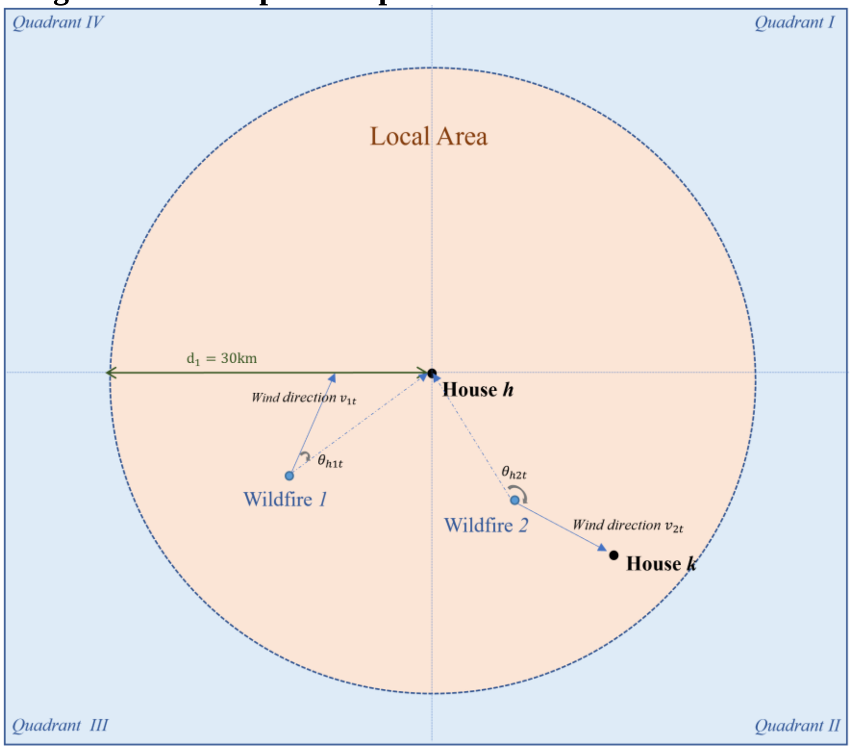
$$ln(P_{hym}) = \boldsymbol{\alpha_f Fire_{hym}} + \boldsymbol{\alpha_w W_{hym}} + \tau_c \times \sigma_y + \eta_m + \partial_h + u_{hym} \quad (8).$$

$ln(P_{hym})$ is the natural logarithm of the adjusted sale price of house h that sold in month m year y in county c. As presented in **Figure 3. 1**, when creating the wildfire measurements, I take into account the following factors: wildfire frequency, wildfire size, wildfire causes, wind pattern, the timing of wildfire, and the distance between the house and the wildfire. Thus, $Fire_{hym}$ is a vector of local wildfire measurements representing the overall wildfire exposures, including the weighted sum of wildfire occurrences, $Local_{hym}$ (I also divide the local wildfires into two categories: upwind wildfires, $Local\_Up_{hmy}$, and downwind wildfires, $Local\_Down_{hmy}$), as well as two additional considerations about extreme wildfire exposures including the number of days since the most recent wildfire that occurred within 80 kilometers of the house since 1992, $Days_{hym}$, and the distance between the property and the nearest wildfire that happened over 5 years before the transaction month m, $Distance_{hym}$[48]. **Figure 3. 4** provides examples to illustrate the upwind and downwind wildfires and the upwind and downwind areas. Furthermore, most wildfires, unlike other natural disasters such as hurricanes, are sparked by human activity;

[48] If no wildfires happened in areas within 80 kilometers of the house since 1992, then I assume $Days_{hym}$=10,000. About 10.2% houses (9.5% observations) are assumed to be 10,000. I choose 1992 as the start year because the wildfire information is only available since 1992 in the dataset I use. I choose 80 km because I define the distant area as the area that is more than 80/100/120 km away from the house. I also create a variable that consider the wildfires occurred in nearby 30 km and 38.6% houses (37.4% observations) never experience wildfires since 1992. The estimates are slightly greater than that of the main analysis, which is reasonable, since the more distant wildfires should have a smaller impact than nearby wildfires. The results are available upon request.
For $Distance_{hym}$, I only consider the recent 5-year wildfires for the distance between the houses and the nearest wildfire, because if the wildfire occurred a long time ago is very unlikely to influence current housing market, although it may occur very close to the houses. Further, large wildfires may also have reduced the local vegetation coverage and thus making reducing future wildfire risks in that area. I also examined the robustness using the recent 10-year wildfires, but the coefficient is not significant. The results are available upon request.

therefore, they may not be wholly exogenous. As a result, I include the ratio of wildfires caused by natural causes as an additional control variable.

Figure 3. 4 Examples of Upwind and Downwind Directions



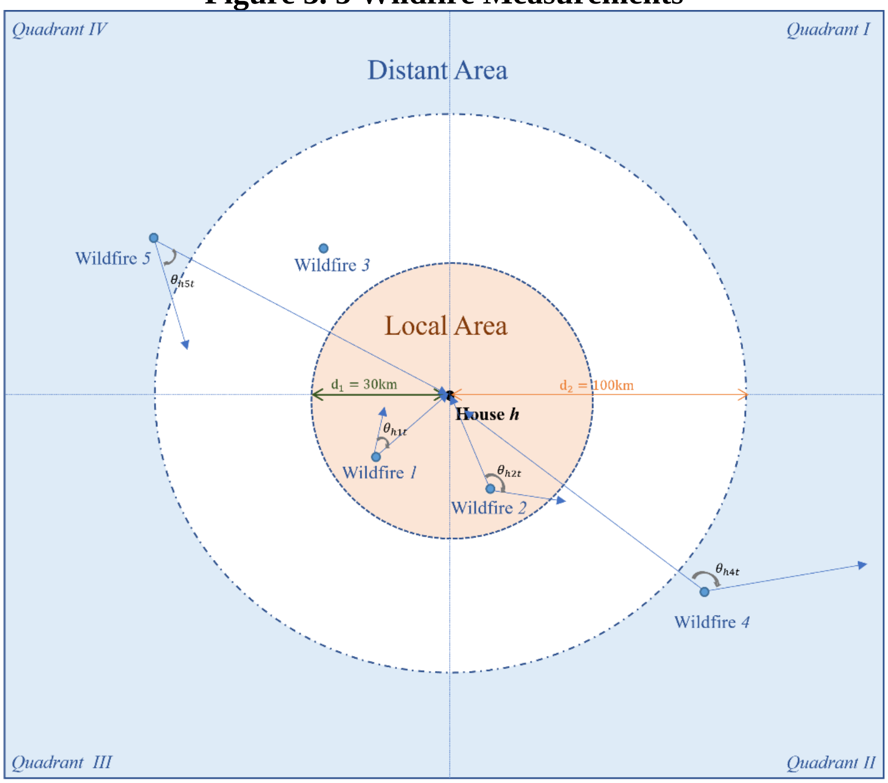
Note: For wildfire *1*, house *h* is located in the downwind areas, and wildfire *1* is an upwind wildfire for house *h*. For wildfire *2*, house *h* is in the upwind direction, and house *k* is located in the downwind area. For house *k*, wildfire *2* is upwind. House *h* is more likely to be affected by wildfire *1* than wildfire *2*.

To assess local effects, in the main analysis, I only consider wildfires that have burned at least 300 acres[49] within 30 kilometers of the property *h* over *t* months before the transaction month *m* year *y*, denoted by J_{ht}. The weighted sum of wildfire occurrences, $Local_{hym}$, is defined in equation (9). For each wildfire event ($j \in J_{hm}$), its impact is discounted by the distance (km) between the house and the wildfire, d_{hj}. The weighted upwind and downwind wildfire occurrences are defined in equations (10) and (11), respectively. If the angle between the wind vector at the wildfire centroid and the vector from the wildfire centroid to the property (θ_{hjt}) is less than 90

[49] Wildfires are coded to different sizes based on the number of acres within the final fire perimeter (A=greater than 0 but less than or equal to 0.25 acres, B=0. 26-9.9 acres, C=10.0-99.9 acres, D=100-299 acres, E=300 to 999 acres, F=1,000 to 4999 acres, and G=5,000+ acres). In the main analysis, I focus on wildfires of size E, F, and G.

degrees, I consider wildfires to be upwind, and the property is located in a downwind direction. Wildfires, on the other hand, burn in the downwind direction of the property. Because homebuyers typically make their housing purchase decision before the transaction date, changes in local environmental amenity levels should influence their decision-making process before the transaction date; I consider the wildfire events that occurred during the 12 months ($t = 12$) before the transaction month m for the short-term effect. **Figure 3. 5** shows examples of constructing wildfire measurements.

Figure 3. 5 Wildfire Measurements



Note: (1) To construct local wildfire measurements for house h sold in month m year y, I only consider wildfires that occurred within 30 kilometers of the house h over *t* months before the transaction month *m* year *y*, such as wildfire *1* and wildfire *2*.
(2) To construct distant wildfire measurements for house h sold in month m year y, I only consider wildfires that occurred more than 100 kilometers from house *h* over *t* months before the transaction month *m* year *y*, such as wildfire *4* and wildfire *5*.
(3) If the angle between the wind vector at wildfire centroid and the vector from wildfire *j*'s centroid to house h (such as θ_{h1t} and θ_{h5t}) is less than 90 degrees, I define wildfires to be upwind. Otherwise, if the angle (such as θ_{h2t} and θ_{h4t}) is greater than 90 degrees, wildfire *j* is considered to burn downwind.

$$Local_{hmy} = \sum_{-t}^{-1}\left(\sum_{j \in J_{ht}(d_{hj}<30\ km)} Wildfire_{hjt} * \frac{1}{d_{hj}}\right) \quad (9)$$

$$Local\_Up_{hmy} = \sum_{-t}^{-1}\left(\sum_{j \in J_{ht}(d_{hj}<30\ km)} Wildfire_{hjt} * I(\theta_{hjt} < 90) * \frac{1}{d_{hj}}\right) \quad (10)$$

$$Local\_Down_{hmy} = \sum_{-t}^{-1}\left(\sum_{j \in J_{ht}(d_{hj}<30\ km)} Wildfire_{hjt} * I(\theta_{hjt} \geq 90) * \frac{1}{d_{hj}}\right) \quad (11)$$

$$where\ \ Wildfire_{hjt} = \begin{cases} 1, & if\ burnt\ acres \geq 300 \\ 0, & otherwise \end{cases}$$

$\boldsymbol{W_{hym}}$ denotes a vector of time- and location-variant factors including the ratio of natural-caused wildfires[50], average temperature, precipitation, humidity, and pressure over the study period ($t = 12$) of the surrounding area (about 13.88 km × 13.88 km) of the property, housing density, population density, and ratio of vegetation by census block level in year *y*, ratio of white people by census block-group level in year *y*, and distance between house and intermix/interface WUI in year *y*. $\tau_c \times \sigma_y$ denotes the county-by-year fixed effects, which control the unobserved constant factors in each year of each county. η_m denotes the month-of-year fixed effects, which control the monthly variations over the annual cycle. ∂_h denotes the property fixed effects. The standard errors are clustered at the property level.

The next step is to test whether local wildfires have a significant impact on local air quality, as shown in Equation (12). If wildfires can significantly influence local air pollution levels, I then consider whether air pollution, specifically $PM_{2.5}$, can be regarded as the channel through which wildfires influence house prices. PM_{hmy} denotes the average $PM_{2.5}$ level of the house's surrounding area (about 1.11km × 1.11km) for the previous *t* (=12) months before transaction month *m* year *y*. I use the weighted sum of wildfire occurrences, $Local_{hym}$ (or $Local\_Up_{hmy}$ and/or $Local\_Down_{hmy}$), to measure wildfires. The ratio of nature-caused wildfires is controlled as well. The covariates $\boldsymbol{W_{hym}}$ are the same as those in equation (8). County-by-year fixed effects, $\tau_c \times \sigma_y$, month-of-year fixed effects, η_m, and property fixed effects, ∂_h, are also included. The

[50] Ratio of natural-caused wildfires $= \frac{\text{Weighted sum of wildfire occurrences}}{\text{Weighted sum of natural}-\text{caused wildfire occurrences}}$

standard errors are clustered at the property level.

$$PM_{hmy} = \gamma_0 + \gamma_{\mathrm{f}} Local_{hmy} + \boldsymbol{\gamma_{\mathrm{W}} W_{hmy}} + \tau_c \times \sigma_y + \eta_m + \partial_h + e_{hmy} \quad (12)$$

Last, I test whether wildfires influence house prices through $PM_{2.5}$. I add $PM_{2.5}$ as an additional variable in equation (8), as shown in equation (13). If $PM_{2.5}$ has a significant impact on house prices (β_p) as well as wildfire exposure ($\boldsymbol{\alpha_f}$ in equation (8) and $\boldsymbol{\beta_f}$ in equation (13)), then there is clear evidence that wildfires indirectly influence house prices through $PM_{2.5}$.

$$ln\left(P_{hmy}\right) = \boldsymbol{\beta_f Fire_{hym}} + \beta_p PM_{hym} + \boldsymbol{\beta_w W_{hmy}} + \tau_c \times \sigma_y + \eta_m + \partial_h + v_{hmy} \quad (13)$$

3.2.2 Addressing Endogeneity of $PM_{2.5}$

When evaluating the impacts of $PM_{2.5}$ as per equation (13), one concern is that $PM_{2.5}$ may be linked with unobserved factors in the error term (i.e., violate the second assumption of "sequential ignorability"), although I control several factors and fixed effects. For example, I cannot fully control the sectoral makeup of the economy, crime rate, school district boundaries, etc. The variation in these neighborhood attributes may influence consumer choices and home prices.

Another way to test the indirect or mediation effect of wildfires is by making use of the exclusion restriction assumption of the instrumental variable (Aguilera et al., 2021). Imai et al. (2011) also suggested that the instrumental variable method can be used to address the endogeneity issue of the mediator. Therefore, I need to find at least one excluded exogenous variable as the instrumental variable (Wooldridge, 2010). Previous studies explored a variety of potential instrumental variables for air pollution. For example, Bayer et al. (2009), Zheng et al. (2014), Barwick et al. (2018), Yang and Zhang (2018), and Chen et al. (2021) considered the transboundary spillover effect of air pollution and used distant air pollutants to create a source of local air pollutants. Considering wildfires are an important source of air pollution, especially $PM_{2.5}$, Tan-Soo (2018) constructed a wind- and distance-based forest fire hotspots instrumental variable for $PM_{2.5}$.

Following these studies, I also create a variable, $DistPM_{cmy}$, to measure distant air pollution and study the spillover effects of air pollution. $DistPM_{cmy}$ is defined as equation (14) and the example of the construction of the distant all-source $PM_{2.5}$ can be found in Appendix B. $DistPM_{cmy}$ denotes the average monthly imported $PM_{2.5}$ from counties ($j \in J(d_{hj} \geq 100km)$) that are at least 100 kilometers away from county c where the property h is located over the previous t month from the transaction month m year y[51]. Through regression analysis, I find a significant positive relationship between distant $PM_{2.5}$ and local $PM_{2.5}$ (results are available upon request). **Figure 3. 6** depicts the spatial distribution of the $DistPM_{cmy}$ from 2010 to 2018, which is the average of $DistPM_{cmy}$ for house h located in state s. This figure highlights the states that suffered more $PM_{2.5}$ from distant counties over the 2010~2018 period. Influenced by the wind direction, the distribution of distant $PM_{2.5}$ shows some differences with the local $PM_{2.5}$, but in general, the eastern and inland areas suffered more imported all-source air pollution. Given that wildfires are one of the major sources of $PM_{2.5}$ and there is significant evidence of transboundary $PM_{2.5}$ spillover effects, I use distant wildfires as an instrumental variable to address the endogeneity problem as well as examine whether air pollution can be regarded as a channel through which distant wildfires influence house prices.

$$DistPM_{cmy} = \frac{1}{t} * (\sum_{-t}^{-1} \sum_{j \in J(d_{hj} \geq 100km)} PM_{jt} * \mathrm{I}\,(\theta_{cjt} > 90) * \frac{1}{d_{cj}}) \quad (14)$$

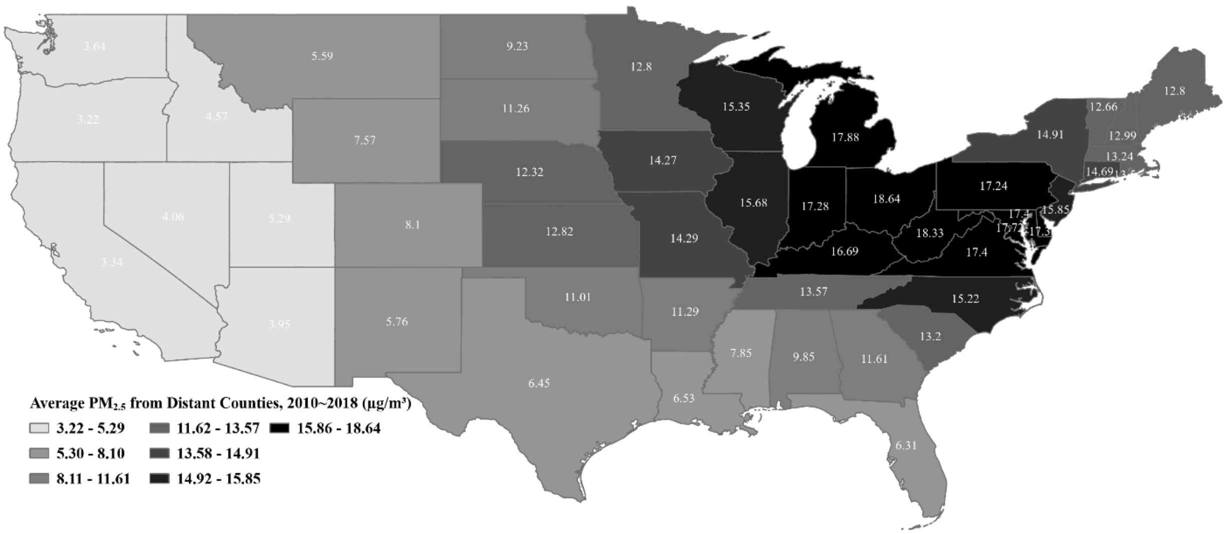
Based on Tan-Soo (2018), I create distant large-scale natural-caused upwind wildfires, $DistFire\_Up_{hmy}$, as an instrumental variable. However, a difference is that I only consider the natural-caused wildfires to further increase confidence in the exogeneity of the instrumental variable. This instrumental variable is obtained using the equation (15). I also create distant downwind large-scale natural-caused wildfires, to examine the robustness, which is defined in equation (16). Only natural-caused wildfires that burned at least 1,000 acres and occurred at least 100 kilometers away from the property during the t-month (t=12) period before the transaction month m year y are considered, which are denoted by S_{st}. Similarly, the distance and wind weights

---

[51] Similarly, I only consider the air pollutants imported from upwind counties. That is, the angle between the vector from the centroid of county j to the centroid of county i and the wind direction vector in county j in month m year y, θ_{cjt}, is less than 90 degrees. Besides, the imported $PM_{2.5}$ is weighted by the reciprocal of geographic distance (km) between county i and j, d_{hj}, as illustrated by equation (9). The farther the county is located, the smaller the transboundary effect of distant $PM_{2.5}$ is expected on local $PM_{2.5}$.

are taken into consideration. Because of the exclusion restriction assumption of the instrumental variable, which requires the instrumental variable not to influence house prices through channels other than $PM_{2.5}$, this method can be used to estimate the indirect effect of wildfires through the channel of $PM_{2.5}$. Aguilera et al. (2021) also used this method to estimate the effects of wildfire-caused $PM_{2.5}$ on respiratory admissions in southern California.

Figure 3. 6 Distribution of Average Distant $PM_{2.5}$ (μg/m³) (2010-2018)



$$DistFire\_Up_{hmy} = \sum_{-t}^{-1}\left(\sum_{s\in S_{st}(d_{hs}\geq 100\ km)} Natural\_Wildfire_{hst} * I(\theta_{hst} > 90) * \frac{1}{d_{hs}}\right) \quad (15)$$

$$DistFire\_Down_{hmy} = \sum_{-t}^{-1}\left(\sum_{s\in S_{st}(d_{hs}\geq 100\ km)} Natural\_Wildfire_{hst} * I(\theta_{hst} \leq 90) * \frac{1}{d_{hs}}\right) \quad (16)$$

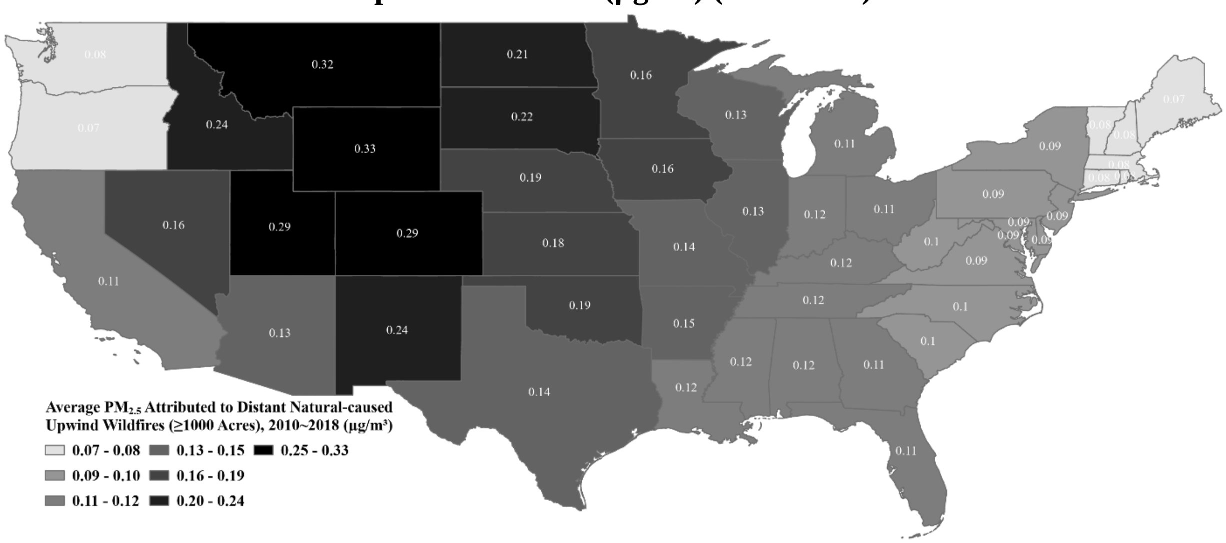
$$where\ \ Natural\_Wildfire_{hst} = \begin{cases} 1, & if\ burnt\ acres \geq 1000 \\ 0, & otherwise \end{cases}$$

Figure 3. 7 depicts the spatial distribution of the predicted value of $PM_{2.5}$ variation attributed to distant natural-caused upwind wildfires from 2010 to 2018 for house *h* located in state *s*. This predicted value can be obtained by equation (17). The distribution presents a different picture to that of the local $PM_{2.5}$ and wildfire occurrence. First, because the western and southern regions experienced significantly more frequent and intense wildfires than the eastern and northern

regions and I apply the distance discount when I create distant natural-caused upwind wildfires, it is reasonable to find that the western and southern regions suffer more the $PM_{2.5}$ attributed to distant wildfires. Second, perhaps because the sea wind clears the air pollutants in coastal states as compared to inland states, the coastal areas in the west and south suffer less the $PM_{2.5}$ attributed to distant large-scale natural-caused upwind wildfires relative to inland places.

$$PM_{hmy} = \gamma_0{}' + \gamma_{\mathrm{f1}}{}'Local\_Up_{hmy} + \gamma_{\mathrm{f2}}{}'Local\_Down_{hmy} + \gamma_{\mathrm{d1}}{}'DistFire\_Up_{hmy} + \gamma_{\mathrm{d2}}{}'DistFire\_Down_{hmy} + \boldsymbol{\gamma}_{\mathbf{W}}{}'\boldsymbol{W}_{\boldsymbol{hmy}} + \tau_c \times \sigma_y + \eta_m + \partial_h + \varepsilon_{hmy} \quad (17)$$

Figure 3. 7 Distribution of Average Predicted Local $PM_{2.5}$ Due to Distant Natural-caused Upwind Wildfires (μg/m³) (2010-2018)



I then can apply two-step regression to get the IV estimate, $\widehat{\beta_p}'$, as presented in equations (18) and (19), which is the local average treatment effect of $PM_{2.5}$ on house prices. In the first stage, I regress the endogenous variable, PM_{hmy}, on the instrumental variable ($DistFire\_Up_{hmy}$) and all the other exogenous variables from the previous models as well as county-by-year fixed effects, month-of-year fixed effects, and house fixed effects to obtain the predicted local $PM_{2.5}$. In the second stage, I substitute $PM_{2.5}$ with the predicted value obtained in the first stage. The exogenous variation of $PM_{2.5}$ in the second stage is attributed to the exogenous distant upwind natural-caused wildfires. As a result, the IV estimate measures how $PM_{2.5}$ is attributable to distant wildfires on local house prices, i.e., the indirect impact of distant wildfires. If the components of

$PM_{2.5}$ do not have significant changes, then I can assume that the indirect effects of local wildfires and distant wildfires are the same. That is, a one-unit increase in $PM_{2.5}$ attributed to wildfires (distant or local) leads to a $\widehat{\beta_p'}$ percentage change in house prices. Combing equation (18), shows that if the overall exposure to local upwind (downwind) wildfire increases by one unit, then the house prices will increase by a $\widehat{\beta_p'} * \widehat{\theta_u}$ ($\widehat{\beta_p'} * \widehat{\theta_d}$) percentage point. In addition, if the overall exposure to distant upwind wildfire increases by one unit, then the house prices will increase by a $\widehat{\beta_p'} * \widehat{\theta_{IV}}$ percentage point.

$$
\begin{aligned}
PM_{hmy} = \theta_0 + \theta_u Local\_Up_{it} + \theta_d Local\_Down_{it} + \theta_{day} Days_{it} + \theta_{dist} Distance_{it} \\
+ \theta_{IV} DistFire\_Up_{hmy} + \boldsymbol{\theta_w} \cdot \boldsymbol{W_{hmy}} + \tau_c \times \sigma_y + \eta_m + \partial_h + \omega_{hmy} \quad (18)
\end{aligned}
$$

$$
\begin{aligned}
ln(P_{hmy}) = \beta_u{}' \; Local\_Up_{it} + \beta_d{}' \; Local\_Down_{it} + \beta_{day}{}' \; Days_{it} + \beta_{dist}{}' \; Distance_{it} \\
+ \beta_p{}' \; \widehat{PM_{hym}} + \boldsymbol{\beta_w}' \; \boldsymbol{W_{hmy}} + \tau_c \times \sigma_y + \eta_m + \partial_h + v_{hmy} \quad (19)
\end{aligned}
$$

Next, I discuss the validity of IV. The valid IV should satisfy two conditions: relevance and exogeneity. First, the air quality in other counties should have an impact on local air quality. Previous research has found evidence of a significant transboundary spillover effect of air pollution, which is mostly driven by wind and is related to distance (Bayer et al., 2009; Banzhaf and Chupp, 2010; Luechinger, 2010; Khawand, 2015; Zheng et al., 2014; Barwick et al., 2018; Yang and Zhang, 2018; Chen and Ye, 2019; Williams and Phaneuf, 2019; Zheng et al., 2019; Chen et al., 2021). Second, distant wildfires can produce a large amount of $PM_{2.5}$, which can be carried by the wind and affect local air quality as well. The t-statistic of the instrumental variable obtained in the first-stage regression (as presented in Section 4) also shows that the instrumental variable is a strong predictor of local $PM_{2.5}$.

Second, the instrumental variable should not directly influence house prices or influence house prices through channels other than $PM_{2.5}$. To reduce the possibility that $PM_{2.5}$ in surrounding counties is correlated with variables affecting property prices in the focal county, I create a 100-kilometer buffer zone and only focus on counties outside the buffer zones. Other factors used to construct instrumental variables include exogenous wind direction, geographic distance, and natural-caused wildfires, further ensuring the exogeneity of the instrumental variables. To improve confidence in the instrumental variable, robustness examinations are performed. In addition, although $DistPM_{cmy}$ cannot be used to get the indirect impact of wildfires, I use $DistFire\_Up_{hmy}$ and $DistPM_{cmy}$ as instrumental variables to conduct the overidentification test. In section 4, I present the robustness examinations and overidentification test results, which suggest that I cannot reject the hypothesis that the instrumental factors are valid and that the results are robust.

Moreover, there may also be a concern that other air pollutants emitted by wildfires can be carried by wind to the local area and thus affect house prices. Air pollutants may influence people's decisions because they can lead to limited visibility and adverse health outcomes. As I discussed in Section 1, wildfire smoke has complex components including greenhouse gases, photochemically reactive compounds, sulfur dioxide (SO2), particulate matter (PM), and liquids, and PM is the major component of wildfire smoke. These particles tend to be very small (with a size range near the wavelength of visible light (0.4-0.7 μm)), and about 90% of total particle masses consist of $PM_{2.5}$ (Stone et al., 2019). Other components are not as visible as $PM_{2.5}$, and thus are less likely to affect purchase decisions. In terms of health concerns, three pollutants (particulate matter, ozone, and carbon monoxide) may pose health threats during wildfire events (Stone et al., 2019). First, $PM_{10\text{-}2.5}$ (PM_{10} is comprised of $PM_{2.5}$ and $PM_{10\text{-}2.5}$) is not a major concern, because about 90% of total particle masses consist of $PM_{2.5}$ (Stone et al., 2019). Second, carbon monoxide dilutes rapidly, so it is rarely a concern unless people are in very close proximity to wildfires (Stone et al., 2019). As a result, it is improbable that carbon monoxide travels to the focal county and influences local health. Third, ozone is not directly emitted from a wildfire, but forms in the plume as wildfire smoke moves downwind (Stone et al., 2019), so ozone can be another channel through which distant wildfires influence local health. However, ozone is not the major component of wildfire smoke. While ozone can affect health, it is invisible and thus less likely to be recognized by people. Thus, ozone can be another potential channel, for the reasons outlined above I assume that the impact of this channel is minor. I will also conduct an empirical study to verify this assumption in the future.

3.2.3 Long-term Effects and Spillover Effect of Wildfires

In the main analysis, I study the wildfires that occurred in the previous twelve months leading up to the transaction month. In this section, longer-run impacts are explored: 3-year and 5-year periods. Instead of using wildfires over the previous 1-year period, I sum up the weighted wildfires over the recent 3 and 5 years. An alternative specification is using wildfires that occurred in each year over the previous 3-year or 5-year period. Also, upwind and downwind wildfires are created similarly.

The instrumental variable strategy enables us to analyze the spillover effects of distant wildfires. If I assume that the marginal effect of $PM_{2.5}$ attributed to distant wildfires and that attributed to local wildfires on house prices are the same, then I can compare the indirect impacts of local wildfires and spillover effects of distant wildfires by comparing the magnitudes of effects of distant and local wildfires on $PM_{2.5}$, as discussed in Section 3.2.2.

3.2.4 Robustness Check

First, in the main analysis, I focus on the wildfires that have burned at least 300 acres. In this section, I study the wildfires that have burned at least 100 acres as well. Second, I consider alternative ways to construct the instrumental variable. In the main analysis, I define faraway wildfires as those located at least 100 kilometers away from a given property. To examine the robustness, I utilize various buffer zone radius sizes, including 80 and 120 kilometers. Besides, I also consider applying the distant wildfires in both upwind and downwind directions and the distant wildfires in the downwind direction as another two sets of instrumental variables.

Third, previous studies on the effects of wildfires and wildfire smoke typically assume that wildfires are exogenous, ignoring the potential endogeneity of wildfire-caused air pollution. I control for the ratio of natural-caused wildfires in the previous section. In this section, rather than focusing on local all-cause wildfires, for comparison, I examine how natural-caused wildfires that can be regarded as totally exogenous affect house prices[52].

4. Results

First, I present in **Table 3. 2** the major findings on how wildfires that occurred within a year of the transaction month, directly and indirectly, affect house prices. Second, I discuss the long-term consequences of wildfires as well as compare the local wildfire effects and spillover effects of distant wildfires, which are presented in **Table 3. 3** and **Table 3. 4**, respectively. Third, I estimate the change in house prices on average due to wildfires, as shown in **Table 3. 5**. Finally, I perform a series of robustness tests, as shown in **Table 3. 6**, **Table 3. 7**, and **Table 3. 8**.

4.1 Main results

I begin by showing how wildfires affect house prices in general, as presented in **Table 3.**

[52] If the natural causes of wildfires cannot influence local house market directly, I can assume that the natural-caused wildfires are fully exogeneous,

2. The completed estimation results can be found in **Table 3. 9**, **Table 3. 10**, and **Table 3. 11** of Appendix C. In the main analysis, I focus on the wildfires that burned at least 300 acres and occurred less than 30 kilometers away from the house during the 1 year before the transaction month. As seen in Column 1 of Panel A of **Table 3. 2**, the total effect of wildfires is not significant, which is somewhat unexpected. Thus, I divide wildfires into upwind and downwind wildfires. From column 2, I find that upwind wildfires have a significant negative impact, whereas downwind wildfires are significantly and positively associated with house prices at a 1% significance level.

In Panel A, there are another two measurements of wildfires: the number of days since the most recent wildfire since 1992 and the distance from the house to the nearest wildfire that occurred within the last five years, both of which have positive and significant impacts on house prices. These coefficients indicate that the longer the property's adjacent areas stay free of wildfires and the farther the nearest recent wildfire is, the higher the property's sale price. The results on the distance also confirm the assumption of the distance discount when I create the wildfire measurements. Also, since wildfires are more likely to happen in areas with higher vegetation coverage and areas in or near the WUI, I control these two factors. Both the vegetation ratio of the census block in which the property is located and the distance between the property and WUI is significantly and positively associated with house prices at a 1% significance level. This finding implies that the households more highly value homes located in the areas with more vegetation coverage, but they are also aware of the risks of living near WUI.

Second, I investigate how local wildfires influence local $PM_{2.5}$ levels, as presented in Panel B. Here, I show that wildfires, both upwind and downwind, have significantly degraded the local air quality. Furthermore, wildfires burning upwind have a larger impact on $PM_{2.5}$ levels than wildfires burning downwind. People who live in places with more vegetation and are closer to the WUI, on the other hand, are less affected by air pollution.

Third, as shown in Panel C, $PM_{2.5}$ has a negative and significant impact on house prices and wildfires negatively influence house prices via $PM_{2.5}$. After controlling $PM_{2.5}$, I also find that the magnitudes of wildfire effects change. However, due to the potential endogeneity issue of air pollution, the estimates of $PM_{2.5}$ may lead to inconsistent estimates of indirect effects. As a result, I use distant upwind large-scale natural-caused wildfires as an instrumental variable. According to the first-stage estimation results, the instrumental variable is a strong predictor of local $PM_{2.5}$. The detailed first-stage estimation results can be found in **Table 3. 12** of Appendix C. To increase confidence in the IV, I conduct an overidentification test by adding distant $PM_{2.5}$ as a second IV. The Hansen J statistic indicates that I cannot reject that these IVs are valid (The results are available upon request). In addition, I find significant evidence of the endogeneity of local $PM_{2.5}$.

After addressing endogeneity, the estimates of $PM_{2.5}$ increase substantially in magnitude, although they only measure the effects of $PM_{2.5}$ attributed to distant wildfires. Also, the magnitudes of indirect effects increase substantially. Similarly, the impact of overall exposure to wildfires is still insignificant, and there are positive upwind wildfire impacts and negative downwind wildfire impacts, both of which are significant at the 1% significance level. In Section 3.2.2, I discussed that the distant wildfires, i.e., the instrumental variable, should primarily influence house prices through $PM_{2.5}$ and that the effects of other channels are likely minor. However, for local wildfires, there may be other channels by which wildfires affect house markets, and wildfires influence the market directly. If I assume that local wildfire-caused $PM_{2.5}$ is the only channel via which local wildfires indirectly affect house prices, then the estimates of wildfires (all/upwind/downwind wildfires) reflect the direct impact of wildfires on house prices. Otherwise, these estimates suggest mixed effects of wildfires' direct and indirect impacts through various pathways other than $PM_{2.5}$.

For the opposite signs of the upwind and downwind wildfire estimates, particularly the positive effects of downwind wildfires, I put forward the following possible explanations. First, there could be substitution effects because downwind locations are riskier and are more likely to be affected by smoke. As a result, people tend to prefer houses that are less affected by upwind wildfires. Second, there is the externality of other air pollutants caused by local wildfires. While I focus on the impact of local wildfire-caused $PM_{2.5}$, wildfire smoke also contains other pollutants. These pollutants have an immediate impact on human welfare as well as people’s perceptions of wildfire severity. Third, there are still other pathways by which wildfires can indirectly affect house prices. According to the literature, wildfires also affect ecosystems, and these effects can be good or harmful. As a result, the coefficients of upwind or downwind wildfires might represent the direct impact of wildfires or mixed effects that include both direct impact and other indirect impacts. Also, note that though the coefficient or marginal effect of downwind wildfires is positive, when the distance to the nearest wildfire and the days since the most recent wildfire are taken into account, the net effect of downwind wildfires is still negative, which is illustrated in Section 4.2.

4.2 Additional Results

First, I investigate the long-term effects of wildfires. **Table 3. 3** presents the long-term effects of the total weighted wildfire occurrences during the last three or five years. The completed results are shown in **Table 3. 13**, **Table 3. 14**, **Table 3. 15**, and **Table 3. 16** of Appendix C. I also estimated another specification, using wildfires that occurred in each year over the previous 3-year or 5-year period; these results are available upon request. Overall, the longer-term effects of upwind and downwind wildfires are similar to the shorter-term effects, but wildfires that occurred a long time ago tend to have a smaller impact on house prices than those that occurred more recently. Further, in contrast to the short-term impact of local wildfires on air quality, in the long term, downwind wildfires have an almost nearly equal negative effect on the air quality as upwind wildfires, implying that, in the long run, wildfires that occur nearby, regardless of wind direction, will eventually contribute to an increase in local air pollution. Also, after addressing the endogeneity problem, the magnitudes of $PM_{2.5}$ and indirect effects increase substantially.

Second, I compare the local and spillover effects of distant wildfires, as shown in **Table 3. 4**. I find that local upwind wildfires have a substantially greater impact on the local $PM_{2.5}$ level than downwind wildfires. The spillover effects of distant wildfires are substantially greater than the indirect effects of local wildfires. Also, comparing the house price changes due to the $PM_{2.5}$ caused by a one standard deviation of local and distant wildfires, I find that distant upwind natural-caused wildfires are associated with an approximate price drop of $573, while the local upwind and downwind wildfires are associated with approximate price drops of $25 and $13, respectively.

Last, I summarize the change in house prices on average due to wildfires with at least 300 burned acres. **Table 3. 5** shows the price difference between a house that has not been affected by recent wildfires and a house that has recently been affected by a local wildfire. To estimate the overall wildfire effects, I consider the price changes due to the changes in all the wildfire measurements (as discussed in Section 3.2). The estimate of each component of wildfire effects can be found in **Table 3. 17** of Appendix C. In **Table 3. 17**, I estimate the price changes due to one more wildfire occurring 30km, 20km, and 10km away from the house within one year, three years, and five years of the transaction month. Overall, the negative impacts of wildfires on house prices diminish as time goes by and as the distance increases, and there are significant price differentials between the houses that are affected by upwind wildfires and those affected by downwind wildfires. Furthermore, I find that the total effect of wildfires (i.e., the impacts of all wildfire measures: weighted upwind and downwind wildfires, indirect effects through $PM_{2.5}$, the number of days since the most recent wildfires, and distance to the nearest wildfire) is negative.

To give a more straightforward picture, I start with a baseline case that no wildfires have occurred within 80 kilometers of a certain *House A* in the last five years. Next, I consider a scenario in which a most recent upwind/downwind wildfire occurred 10/20/30 km away from the same house one year ago. To calculate the price differences between the baseline and alternative scenarios, I assume that a wildfire occurred within 80 km from *House A* five years ago for the baseline case, so the estimates should be the upper bound of price changes. For example, if the most recent wildfire occurred one year ago in an upwind direction and was only 10 kilometers away from *House A*, the sale price of *House A* will decline by $3,797. Similarly, if the wildfire occurred downwind, the sale price would drop by $2,572, which is less than the upwind wildfire but still substantial.

4.3 Robustness Check

In this section, I examine the robustness of the estimation results by varying the wildfire size I focus on and utilizing alternative instrumental variables, as well as consider whether the wildfire causes affect the conclusions.

First, instead of focusing on wildfires that have burned at least 300 acres, I study wildfires that have burned at least 100 acres, which occur more frequently, as shown in **Table 3. 6**. In Panel A, the effect of all local wildfires becomes positive and significant, as the positive impact of downwind wildfires is larger than the negative effects of upwind wildfires. This finding may indicate that, in comparison to the intensity of the wildfire, a high frequency of wildfires is more likely to instill people's perceptions of wildfire hazards, affecting families' preferences for upwind places. In line with my expectations, wildfires of a smaller magnitude have a smaller impact on air quality, but the effects of $PM_{2.5}$ on house prices are still significant. Thus, the indirect effects of wildfire through $PM_{2.5}$ on house prices are smaller in magnitude compared to $PM_{2.5}$ induced by larger-size wildfires.

Second, I consider alternative approaches to creating the instrumental variable. I also build buffer zones with a radius of 80 and 120 kilometers, in addition to a 100-kilometer radius. The results are presented in **Table 3. 7**. Despite the minor difference in the estimations, the results are robust. Again, all the instrumental variables are strong predictors of $PM_{2.5}$. As the distance increases, the prediction power (F-statistics) decreases. In addition, I also consider distant all-cause wildfires and downwind wildfires. Since they fail to pass the overidentification test, I do not present them in this paper; however, the complete results are available upon request.

Last, in the main analysis, I control for the ratio of natural-caused wildfires to account for potential omitted human components that may also affect the housing market. In this evaluation, I only focus on natural-caused wildfires. As shown in **Table 3. 8**, the indirect impact of upwind wildfires through air pollution is substantially greater than that of downwind wildfires. One possible explanation is that natural-caused wildfires are usually accompanied by storms, allowing wind patterns to play a more significant role in contaminant transport. This result may imply that natural-caused upwind wildfires primarily affect local house prices by releasing air pollutants, while natural-caused downwind wildfires are more likely to influence household decisions by altering perceptions of future wildfire risks in upwind and downwind areas. Based on these results, I further identify the sources of the opposite signs of the effects of the upwind and downwind wildfire that were discussed in Section 4.1: wind directions play a crucial role in causing the house price differences between upwind and downwind areas.

Table 3. 2 Direct and Indirect Effects of Wildfires on House Prices (Wildfires Over the recent 12 months, 0~30km, ⩾ 300 acres)

Panel A. Total Effects of Wildfires on House Prices		
Dependent variable: ln(Price)	**(1)**	**(2)**
All wildfires	0.0008	
	(0.17)	
Upwind wildfires		-0.0311***
		(-4.55)
Downwind wildfires		0.0326***
		(5.19)
Days since the most recent wildfires	0.0002***	0.0002***
	(4.85)	(4.83)
Distance to the nearest wildfires	0.0002***	0.0002***
	(6.63)	(6.65)
Vegetation ratio	0.4576***	0.4571***
	(45.98)	(45.94)
Distance to WUI	0.0068***	0.0068***
	(5.99)	(5.98)

Panel B. Effects of Wildfires on Air Pollution		
Dependent variable: $PM_{2.5}$	**(1)**	**(2)**
All wildfires	0.1602***	
	(13.81)	
Upwind wildfires		0.1847***
		(8.33)
Downwind wildfires		0.1357***
		(12.25)
Vegetation ratio	-0.1680***	-0.1676***
	(-11.91)	(-11.88)
Distance to WUI	-0.0405***	-0.0404***
	(-23.25)	(-23.24)

Table 3. 2 (cont'd)

Panel C. Indirect Effects of Wildfires on House Prices				
Dependent variable: ln(Price)	**(1) OLS**	**(2) IV**	**(3) OLS**	**(4) IV**
All wildfires	0.0016	0.0049		
	(0.37)	(1.11)		
Upwind wildfires			-0.0301***	-0.0263***
			(-4.44)	(-3.93)
Downwind wildfires			0.0333***	0.0362***
			(5.31)	(5.71)
PM	-0.0051***	-0.0247***	-0.0050***	-0.0249***
	(-8.41)	(-5.59)	(-8.40)	(-5.63)
Days since the most recent wildfires	0.0002***	0.0002***	0.0002***	0.0002***
	(4.79)	(4.54)	(4.77)	(4.52)
Distance to the nearest wildfires	0.0002***	0.0002***	0.0002***	0.0002***
	(6.77)	(7.28)	(6.80)	(7.31)
Vegetation ratio	0.4568***	0.4534***	0.4563***	0.4529***
	(45.89)	(45.41)	(45.85)	(45.37)
Distance to WUI	0.0066***	0.0058***	0.0066***	0.0057***
	(5.81)	(5.04)	(5.80)	(5.02)
Kleibergen-Paap rk Wald F statistics		2.2e+04		2.2e+04
Endogeneity test		20.221		20.686
(p-value)		[0.0000]		[0.0000]
Indirect Impact				
All wildfires	-0.08%	-0.23%		
Upwind wildfires			-0.09%	-0.32%
Downwind wildfires			-0.07%	-0.16%

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 3 Long-term Effects of Wildfires (⩾ 300 acres)

Panel A. Total Effects of Wildfires on House Prices		
Dependent variable: ln(Price)	**3 years**	**5 years**
	(1)	**(2)**
Upwind wildfires	-0.0290***	-0.0107***
	(-6.06)	(-2.93)
Downwind wildfires	0.0204***	0.0137***
	(5.13)	(3.83)
Days since the most recent wildfires	0.0002***	0.0003***
	(4.99)	(5.32)
Distance to the nearest wildfires	0.0002***	0.0001***
	(5.96)	(4.44)

Panel B. Effects of Wildfires on Air Pollution		
Dependent variable: $PM_{2.5}$	**(1)**	**(2)**
Upwind wildfires	0.1408***	0.1101***
	(13.22)	(10.47)
Downwind wildfires	0.1566***	0.1191***
	(16.45)	(12.94)

Panel C. Indirect Effects of Wildfires on House Prices				
	(1)	**(2)**	**(3)**	**(4)**
Dependent variable: ln(Price)	**OLS**	**IV**	**OLS**	**IV**
Upwind wildfires	-0.0282***	-0.0254***	-0.0100***	-0.0078***
	(-5.92)	(-5.33)	(-2.77)	(-2.15)
Downwind wildfires	0.0213***	0.0244***	0.0144***	0.0168***
	(5.32)	(5.96)	(3.99)	(4.52)
PM	-0.0051***	-0.0240***	-0.0051***	-0.0233***
	(-8.46)	(-5.49)	(-8.48)	(-5.34)
Days since the most recent wildfires	0.0002***	0.0002***	0.0002***	0.0002***
	(4.93)	(4.68)	(5.25)	(4.68)
Distance to the nearest wildfires	0.0002***	0.0002***	0.0001***	0.0002***
	(6.11)	(6.62)	(4.60)	(6.62)
Kleibergen-Paap rk Wald F statistics		2.2e+04		2.2e+04
Endogeneity test (p-value)		19.183 [0.0000]		17.886 [0.0000]
Indirect Impact				
Upwind wildfires	-0.07%	-0.31%	-0.06%	-0.26%
Downwind wildfires	-0.08%	-0.34%	-0.06%	-0.28%

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 4 Local and Spillover Effects of Wildfires

Panel A. Effects of Wildfires on Air Pollution		
Dependent variable: $PM_{2.5}$	**(1)**	**(2)**
Local Wildfires ($\geq$ 300 acres)	0.0949^{***} (10.92)	
Local Upwind Wildfires ($\geq$ 300 acres)		0.1265^{***} (7.93)
Local Downwind Wildfires ($\geq$ 300 acres)		0.0633^{***} (6.11)
Distant Upwind Wildfires (Natural-caused, $\geq$ 5000 acres)	2.3600^{***} (147.27)	2.3605^{***} (147.21)
Panel B. Indirect Effects of Wildfires on House Prices		
Channel of $PM_{2.5}$	**Using IV Estimates**	
	(1)	**(2)**
Local Wildfires	-0.24%	
Local Upwind Wildfires		-0.31%
Local Downwind Wildfires		-0.16%
Distant Upwind Wildfires	-5.88%	-5.88%
Panel C. Comparisons of House Price Changes ($) on Average Due to the Local and Spillover Effects		
A One SD Change of Local Wildfires ($\approx$0.066), Local Upwind Wildfires ($\approx$0.042), Local Downwind Wildfires ($\approx$0.043), and Distant Upwind Wildfires ($\approx$0.051) Mean Adjusted Sale Prices$\approx$$191089.4		
	(1)	**(2)**
Local Wildfires	-29.80	
Local Upwind Wildfires		-25.28
Local Downwind Wildfires		-12.95
Distant Upwind Wildfires	-572.69	-572.81

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 5 Total Effects of Wildfires Under Different Scenarios

Baseline: No wildfire happened less than 80 km away from House A over recent five years			
Scenarios: If the most recent wildfire happened upwind/downwind one year ago			
The wildfire happened *d* km away from House A	*d*=30km	*d*=20km	*d*=10km
Price change between baseline and different scenarios (dollars)			
Upwind	-2656.78	-3132.91	-3796.95
Downwind	-2248.49	-2520.47	-2572.06

Note: (1) I do not consider the spillover effects from distant wildfires. (2) I assume that there is a wildfire that happened 80 km away from *House A* five years ago for the baseline case to calculate the price change, so the estimate of the total effect of an upwind or downwind wildfire is just a lower bound.

Table 3. 6 Robustness Check – Different Wildfire Size (⩾ 100 acres)

Panel A. Total Effects of Wildfires on House Prices		
Dependent variable: ln(Price)	**(1)**	**(2)**
All wildfires	0.0055**	
	(2.05)	
Upwind wildfires		-0.0111***
		(-3.07)
Downwind wildfires		0.0217***
		(4.60)
Days since the most recent wildfires	-0.0001	-0.0001
	(-1.05)	(-1.06)
Distance to the nearest wildfires	0.0001***	0.0001***
	(3.87)	(3.88)
Vegetation ratio	0.4584***	0.4581***
	(46.06)	(46.03)
Distance to WUI	0.0068***	0.0068***
	(5.99)	(5.98)

Panel B. Effects of Wildfires on Air Pollution		
Dependent variable: $PM_{2.5}$	**(1)**	**(2)**
All wildfires	0.0686***	
	(9.90)	
Upwind wildfires		0.0943***
		(10.73)
Downwind wildfires		0.0437***
		(4.50)
Vegetation ratio	-0.1677***	-0.1673***
	(-11.88)	(-11.85)
Distance to WUI	-0.0407***	-0.0406***
	(-23.36)	(-23.36)

Table 3. 6 (cont'd)

Panel C. Indirect Effects of Wildfires on House Prices				
Dependent variable: ln(Price)	**(1) OLS**	**(2) IV**	**(3) OLS**	**(4) IV**
All wildfires	0.0059**	0.0072***		
	(2.18)	(2.59)		
Upwind wildfires			-0.0107***	-0.0090**
			(-2.94)	(-2.47)
Downwind wildfires			0.0219***	0.0228***
			(4.62)	(4.69)
PM	-0.0050***	-0.0223***	-0.0050***	-0.0223***
	(-8.34)	(-8.93)	(-8.31)	(-8.93)
Days since the most recent wildfires	-0.0001	-0.0001	-0.0001	-0.0001
	(-1.13)	(-1.42)	(-1.14)	(-1.42)
Distance to the nearest wildfires	0.0001***	0.0001***	0.0001***	0.0001***
	(3.98)	(4.35)	(3.99)	(4.37)
Vegetation ratio	0.4575***	0.4546***	0.4573***	0.4543***
	(45.97)	(45.62)	(45.94)	(45.60)
Distance to WUI	0.0066***	0.0059***	0.0066***	0.0059***
	(5.80)	(5.16)	(5.80)	(5.15)
Kleibergen-Paap rk Wald F statistics		2.2e+04		2.2e+04
Endogeneity test (p-value)		18.823 [0.0000]		18.855 [0.0000]

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 7 Robustness Check – Different Radius of Buffer Zone

Dependent Variable: ln(Price)	Main Analysis	Robustness Check	
	$d = 100\ km$	$d = 80\ km$	$d = 120\ km$
Upwind wildfires (≥ 300 acres)	-0.0263***	-0.0271***	-0.0262***
	(-3.93)	(-4.04)	(-3.92)
Downwind wildfires (≥ 300 acres)	0.0362***	0.0356***	0.0363***
	(5.71)	(5.62)	(5.71)
PM	-0.0249***	-0.0207***	-0.0252***
	(-5.63)	(-4.91)	(-5.49)
Days since the most recent wildfires	0.0002***	0.0002***	0.0002***
	(4.52)	(4.57)	(4.52)
Distance to the nearest wildfires	0.0002***	0.0002***	0.0002***
	(7.31)	(7.20)	(7.31)
Vegetation ratio	0.4529***	0.4536***	0.4528***
	(45.37)	(45.45)	(45.35)
Distance to WUI	0.0057***	0.0059***	0.0057***
	(5.02)	(5.17)	(5.00)
Kleibergen-Paap rk Wald F statistics	2.2e+04	2.3e+04	2.1e+04
Endogeneity test (p-value)	20.686 [0.0000]	14.175 [0.0000]	19.706 [0.0000]

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 8 Direct and Indirect Effects of Wildfires on House Prices
(Natural-caused wildfires, Over the recent 12 months)

Panel A. Total Effects of Wildfires on House Prices		
Dependent variable: ln(Price)	**(1)**	**(2)**
All wildfires (≥ 300 acres)	0.0487***	
	(3.69)	
Upwind wildfires (≥ 300 acres)		-0.0016
		(-0.07)
Downwind wildfires (≥ 300 acres)		0.0749***
		(4.32)
Days since the most recent wildfires	0.0002***	0.0002***
	(4.18)	(4.12)
Distance to the nearest wildfires	0.00005***	0.00005***
	(5.76)	(5.74)
Vegetation ratio	0.4587***	0.4587***
	(46.10)	(46.09)
Distance to WUI	0.0068***	0.0068***
	(6.01)	(6.02)

Panel B. Effects of Wildfires on Air Pollution		
Dependent variable: $PM_{2.5}$	**(1)**	**(2)**
All wildfires (≥ 300 acres)	0.3298***	
	(12.13)	
Upwind wildfires (≥ 300 acres)		0.7961***
		(20.06)
Downwind wildfires (≥ 300 acres)		0.0860***
		(3.87)
Vegetation ratio	-0.1700***	-0.1698***
	(-12.05)	(-12.03)
Distance to WUI	-0.0409***	-0.0409***
	(-23.50)	(-23.53)

Table 3. 8 (cont'd)

Panel C. Indirect Effects of Wildfires on House Prices				
Dependent variable: ln(Price)	**(1) OLS**	**(2) IV**	**(3) OLS**	**(4) IV**
All wildfires (≥ 300 acres)	0.0503***	0.0577***		
	(3.80)	(4.27)		
Upwind wildfires (≥ 300 acres)			0.0024	0.0201
			(0.11)	(0.90)
Downwind wildfires (≥ 300 acres)			0.0754***	0.0773***
			(4.34)	(4.41)
PM	-0.0051***	-0.0277***	-0.0051***	-0.0277***
	(-8.45)	(-6.25)	(-8.41)	(-6.25)
Days since the most recent wildfires	0.0002***	0.0001***	0.0002***	0.0001***
	(3.97)	(2.95)	(3.91)	(2.91)
Distance to the nearest wildfires	0.00005***	0.0001***	0.00005***	0.0001***
	(6.00)	(6.95)	(5.98)	(6.93)
Vegetation ratio	0.4579***	0.4540***	0.4579***	0.4540***
	(46.01)	(45.48)	(46.01)	(45.48)
Distance to WUI	0.0066***	0.0057***	0.0066***	0.0057***
	(5.83)	(4.95)	(5.84)	(4.95)
Kleibergen-Paap rk Wald F statistics		2.1e+04		2.1e+04
Endogeneity test (p-value)		26.712 [0.0000]		26.552 [0.0000]

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

5. Conclusions

With more frequent and intense wildfires occurring worldwide, understanding how wildfires affect the quality of life is more important than ever. Wildfires can cause property and infrastructure damage, as well as injuries and fatalities. Wildfires occur in most terrestrial ecosystems and are associated with a variety of ecological changes. Wildfire smoke is a major environmental hazard, as it can be carried by the wind and transported to distant regions. $PM_{2.5}$ is the most common component of wildfire smoke particles, and it poses health and social risks to people. As more people live near or in the WUI, and as increasingly frequent and intense extreme weather events are expected to be in the coming decades, wildfire-related losses are expected to increase. In this broader context, this paper investigates how wildfires, directly and indirectly, affect the pricing of residential homes in the United States as well as verifies that households are facing a tradeoff between amenities caused by living near nature and the increasing wildfire risks. Clarifying the mechanisms through which wildfires affect house prices can help policymakers enact related policies and initiatives, such as wildfire education, post-fire restoration priority, and post-fire air pollution control.

This paper is based on the hedonic pricing model and applies the instrumental variable method to explore the total effect of wildfires on house prices and distinguish the indirect effect of wildfires-caused $PM_{2.5}$. I use a national repeat-sale dataset from 2010 to 2018 and match each property with each wildfire event since 1992. This enables us to create wildfire measurements that account for wildfire frequency and severity, wind pattern, wildfire causes, the distance between wildfires and houses, and wildfire timing. Considering that the wildfire risks are extremely high near the WUI and in areas with high vegetation coverage and that people place a high value on the houses with better views of natural scenery, I control for the distance between the house and WUI as well as the neighborhood vegetation coverage. Moreover, I address the endogeneity problem of $PM_{2.5}$, explore the spillover effects and long-term effects of wildfires, and examine whether wildfire causes have an impact on the findings.

I find that the frequency and severity of the wildfires, the wind patterns, wildfire causes, the distances between wildfires and houses, and the timing of wildfires are important factors that influence the wildfire effects on house prices. Overall, wildfires have a significant negative impact on house prices. There is a significant price differential between houses located in upwind and downwind locations of the wildfires. I put forward three possible explanations: the substitution effect, externality, and the existence of other channels (other than air pollution) via which wildfires affect house prices. Future research is needed to further examine these three possible explanations. Meanwhile, $PM_{2.5}$ emitted by wildfires, both upwind and downwind, results in a drop in house prices. This negative impact is more significant after addressing the endogeneity of $PM_{2.5}$. Although the indirect effect of nearby wildfires via $PM_{2.5}$ is considerably smaller than that of the direct effect of wildfires, the indirect/spillover effects of distant wildfires via $PM_{2.5}$ are substantial; wildfire smoke can travel a long distance and influence a broader area. As a result, local or nearby wildfires mainly influence house prices via direct damages to the local amenities, whereas distant wildfires have a significant impact on the broader house market via exporting wildfire-caused $PM_{2.5}$. $PM_{2.5}$ may not be the only pathway that local wildfires affect the property market. Future research is needed to further explore other pathways such as water pollution and postfire floods. Also, this study distinguishes between the wildfire effects and living near nature. Although people prefer houses with more vegetation, they are still concerned about the wildfire risks. When I only consider wildfires generated by natural causes, I find that wind plays a more significant role in transporting wildfire-caused air pollutants to the downwind areas.

The results offer policymakers a more comprehensive view of how wildfires affect homebuyers' decisions. Wildfires affect house prices in multiple ways, which many homebuyers may be unaware of and thus attribute to other environmental factors. The implementation of wildfire prevention, education, and restoration programs may be improved by taking into consideration local meteorological factors (such as wind patterns), environmental characteristics (such as vegetation coverage and the number of households living near or in WUI), and mechanisms by which wildfires influence local environmental amenities such as wildfire-related air pollution, water pollution, floods, etc.

REFERENCES

Aguilera, R., Corringham, T., Gershunov, A., & Benmarhnia, T. (2021). Wildfire smoke impacts respiratory health more than fine particles from other soures: Observational evidence from Southern California. *Nature Communications*, *12*(1), 1–8.

Athukorala, W., Martin, W., Neelawala, P., Rajapaksa, D., & Wilson, C. (2016). Impact of wildfires and floods on property values: A before and after analysis. *The Singapore Economic Review*, *61*(01), 1640002.

Badger, S. G., & Foley, M. (2019, September 1). *Large-loss fires and explosions in the United States during 2018*. National Fire Protection Association. https://www.nfpa.org/News-and-Research/Publications-and-media/NFPA-Journal/2019/November-December-2019/Features/Large-Loss

Bakian, A. V., Huber, R. S., Coon, H., Gray, D., Wilson, P., McMahon, W. M., & Renshaw, P. F. (2015). Acute air pollution exposure and risk of suicide completion. *American Journal of Epidemiology*, *181*(5), 295–303.

Banzhaf, H. S., & Chupp, B. A. (2010). *Heterogeneous harm vs. spatial spillovers: Environmental federalism and US air pollution* (Working Paper No. 15666). National Bureau of Economic Research. https://doi.org/10.3386/w15666

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Barwick, P. J., Li, S., Rao, D., & Zahur, N. B. (2018). *The morbidity cost of air pollution: Evidence from consumer spending in China* (SSRN Scholarly Paper No. 2999068). https://doi.org/10.2139/ssrn.2999068

Bayer, P., Keohane, N., & Timmins, C. (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management*, *58*(1), 1–14. https://doi.org/10.1016/j.jeem.2008.08.004

Braithwaite, I., Zhang, S., Kirkbride, J. B., Osborn, D. P., & Hayes, J. F. (2019). 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*(12), 126002.

Brook, R. D., Rajagopalan, S., Pope, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., Holguin, F., Hong, Y., Luepker, R. V., Mittleman, M. A., Peters, A., Siscovick, D., Smith, S. C., Whitsel, L., Kaufman, J. D., & American Heart Association Council on Epidemiology and Prevention, Council on the Kidney in Cardiovascular Disease, and Council on Nutrition, Physical Activity and Metabolism. (2010). Particulate matter air pollution and cardiovascular disease: An update to the scientific statement from the American Heart Association. *Circulation*, *121*(21), 2331–2378. https://doi.org/10.1161/CIR.0b013e3181dbece1

Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., & Wara, M. (2021). The changing risk and burden of wildfire in the United States. *Proceedings of the National Academy of Sciences*, *118*(2), e2011048118. https://doi.org/10.1073/pnas.2011048118

Burkhardt, J., Bayham, J., Wilson, A., Berman, J. D., O'Dell, K., Ford, B., Fischer, E. V., & Pierce, J. R. (2020). The relationship between monthly air pollution and violent crime across the United States. *Journal of Environmental Economics and Policy*, *9*(2), 188–205.

Burkhardt, J., Bayham, J., Wilson, A., Carter, E., Berman, J. D., O'Dell, K., Ford, B., Fischer, E. V., & Pierce, J. R. (2019). The effect of pollution on crime: Evidence from data on particulate matter and ozone. *Journal of Environmental Economics and Management*, *98*, 102267.

Chay, K. Y., & Greenstone, M. (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy*, *113*(2), 376–424.

Chen, S., Chen, Y., Lei, Z., & Tan-Soo, J.-S. (2021). Chasing clean air: Pollution-induced travels in China. *Journal of the Association of Environmental and Resource Economists*, *8*(1), 59–89. https://doi.org/10.1086/711476

Chen, X., & Ye, J. (2019). When the wind blows: Spatial spillover effects of urban air pollution in China. *Journal of Environmental Planning and Management*, *62*(8), 1359–1376. https://doi.org/10.1080/09640568.2018.1496071

Dawson, J. P., Bloomer, B. J., Winner, D. A., & Weaver, C. P. (2014). Understanding the meteorological drivers of U.S. particulate matter concentrations in a changing climate. *Bulletin of the American Meteorological Society*, *95*(4), 521–532.

Dittrich, R., & McCallum, S. (2020). How to measure the economic health cost of wildfires–A systematic review of the literature for northern America. *International Journal of Wildland Fire*, *29*(11), 961–973.

Doerr, S. H., & Santín, C. (2016). Global trends in wildfire and its impacts: Perceptions versus realities in a changing world. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1696), 20150345.

Fontenla, M., Ben Goodwin, M., & Gonzalez, F. (2019). Pollution and the choice of where to work and live within Mexico City. *Latin American Economic Review*, *28*(1), 1–17.

Gładka, A., Rymaszewska, J., & Zatoński, T. (2018). Impact of air pollution on depression and suicide. *Int J Occup Med Environ Health*, *31*(6), 711–721.

Hammer, M. S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A. M., Hsu, N. C., Levy, R. C., Garay, M. J., Kalashnikova, O. V., Kahn, R. A., Brauer, M., Apte, J. S., Henze, D. K., Zhang, L., Zhang, Q., Ford, B., Pierce, J. R., & Martin, R. V. (2020). Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). *Environmental Science & Technology*, *54*(13), 7879–7890. https://doi.org/10.1021/acs.est.0c01764

Hayes, A. F. (2017). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford publications.

Hoek, G., Krishnan, R. M., Beelen, R., Peters, A., Ostro, B., Brunekreef, B., & Kaufman, J. D. (2013). Long-term air pollution exposure and cardio- respiratory mortality: A review. *Environmental Health*, *12*(1), 43. https://doi.org/10.1186/1476-069X-12-43

Huang, X., & Lanz, B. (2018). The value of air quality in Chinese cities: Evidence from labor and property market outcomes. *Environmental and Resource Economics*, *71*(4), 849–874.

Hyslop, N. P. (2009). Impaired visibility: The air pollution people see. *Atmospheric Environment*, *43*(1), 182–195.

Imai, K., Keele, L., Tingley, D., & Yamamoto, T. (2011). Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review*, *105*(4), 765–789.

Khawand, C. (2015). *Air quality, mortality, and perinatal health: Causal evidence from wildfires* (No. pkh318). Job Market Papers. https://ideas.repec.org/p/jmp/jm2015/pkh318.html

Kochi, I., Donovan, G. H., Champ, P. A., & Loomis, J. B. (2010). The economic cost of adverse health effects from wildfire-smoke exposure: A review. *International Journal of Wildland Fire*, *19*(7), 803–817.

Lavín, F. V., Dresdner, J., & Aguilar, R. (2011). The value of air quality and crime in Chile: A hedonic wage approach. *Environment and Development Economics*, *16*(3), 329–355.

Lin, G.-Z., Li, L., Song, Y.-F., Zhou, Y.-X., Shen, S.-Q., & Ou, C.-Q. (2016). The impact of ambient air pollution on suicide mortality: A case-crossover study in Guangzhou, China. *Environmental Health*, *15*(1), 1–8.

Linden, L., & Rockoff, J. E. (2008). Estimates of the impact of crime risk on property values from Megan's laws. *American Economic Review*, *98*(3), 1103–1127.

Loomis, J. (2004). Do nearby forest fires cause a reduction in residential property values? *Journal of Forest Economics*, *10*(3), 149–157.

Lu, J. G., Lee, J. J., Gino, F., & Galinsky, A. D. (2018). Polluted morality: Air pollution predicts criminal activity and unethical behavior. *Psychological Science*, *29*(3), 340–355.

Luechinger, S. (2010). Life satisfaction and transboundary air pollution. *Economics Letters*, *107*(1), 4–6. https://doi.org/10.1016/j.econlet.2009.07.007

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science*, *1*(4), 173–181.

Manson, S., Schroeder, J., Van Riper, D., Kugler, T., & Ruggles, S. (2021). *IPUMS National Historical Geographic Information System: Version 16.0* [Data set]. IPUMS. https://doi.org/10.18128/D050.V16.0

Mocko, D., & NASA/GSFC/HSL. (2012). *NLDAS Primary Forcing Data L4 Monthly 0.125 x 0.125 degree V002. 2012* [Data set]. Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5067/Z62LT6J96R4F

Mueller, J., Loomis, J., & González-Cabán, A. (2009). Do repeated wildfires change homebuyers' demand for homes in high-risk areas? A hedonic analysis of the short and long-term effects of repeated wildfires on house prices in Southern California. *The Journal of Real Estate Finance and Economics*, *38*(2), 155–172.

Mueller, J. M., Lima, R. E., Springer, A. E., & Schiefer, E. (2018). Using matching methods to estimate impacts of wildfire and postwildfire flooding on house prices. *Water Resources Research*, *54*(9), 6189–6201.

Mueller, J. M., & Loomis, J. B. (2014). Does the estimated impact of wildfires vary with the housing price distribution? A quantile regression approach. *Land Use Policy*, *41*, 121–127.

Neary, D. G., & Leonard, J. M. (2019). Physical Vulnerabilities from Wildfires: Flames, Floods, and Debris Flows. In *Natural Resources Management and Biological Sciences*. IntechOpen.

Neary, D. G., Ryan, K. C., & DeBano, L. F. (2005). Wildland fire in ecosystems: Effects of fire on soils and water. *Gen. Tech. Rep. RMRS-GTR-42-Vol. 4. Ogden, UT: US Department of Agriculture, Forest Service, Rocky Mountain Research Station. 250 p., 42*.

Radeloff, V. C., Helmers, D. P., Kramer, H. A., Mockrin, M. H., Alexandre, P. M., Bar Massada, A., Butsic, V., Hawbaker, T. J., Martinuzzi, S., Syphard, A. D., & others. (2017). *The 1990-2010 wildland-urban interface of the conterminous United States-geospatial data*.

Radeloff, V. C., Helmers, D. P., Kramer, H. A., Mockrin, M. H., Alexandre, P. M., Bar-Massada, A., Butsic, V., Hawbaker, T. J., Martinuzzi, S., Syphard, A. D., & others. (2018). Rapid growth of the US wildland-urban interface raises wildfire risk. *Proceedings of the National Academy of Sciences*, *115*(13), 3314–3319.

Sandberg, D. V. (2003). *Wildland fire in ecosystems: Effects of fire on air*. US Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Short, K. C. (2021). *Spatial wildfire occurrence data for the United States, 1992-2018 [FPA_FOD_20210617] (5th Edition)* [Data set]. Forest Service Research Data Archive. https://doi.org/10.2737/RDS-2013-0009.5

Stetler, K. M., Venn, T. J., & Calkin, D. E. (2010). The effects of wildfire and environmental amenities on property values in northwest Montana, USA. *Ecological Economics*, *69*(11), 2233–2243.

Stone, S. L., Anderko, L., Berger, M., Butler, C., Cascio, W., Clune, A., & others. (2019). Wildfire smoke: A guide for public health officials. *US Environmental Protection Agency*.

Tan-Soo, J.-S. (2018). Valuing Air Quality in Indonesia Using Households' Locational Choices. *Environmental and Resource Economics*, *71*(3), 755–776. https://doi.org/10.1007/s10640-017-0182-z

Urbanski, S. P., Hao, W. M., & Baker, S. (2008). Chemical composition of wildland fire emissions. *Developments in Environmental Science*, *8*, 79–107.

U.S. Bureau of Labor Statistics. (n.d.). *Chained CPI for All Urban Consumers, U.S. city average (C-CPI-U) (Series Id: SUUR0000SAH*） [Data set]. U.S. Bureau of Labor Statistics. Retrieved May 13, 2021, from https://www.bls.gov/data/home.htm

U.S. Environmental Protection Agency. (2019). *Integrated Science Assessment (ISA) for Particulate Matter (Final Report, Dec 2019)* (EPA/600/R-19/188). U.S. Environmental Protection Agency. https://cfpub.epa.gov/ncea/isa/recordisplay.cfm?deid=347534

U.S. Environmental Protection Agency. (2020). *Policy Assessment for the Review of the National Ambient Air Quality Standards for Particulate Matter*. U.S. Environmental Protection Agency. https://www.epa.gov/sites/production/files/2020-01/documents/final_policy_assessment_for_the_review_of_the_pm_naaqs_01-2020.pdf

Van Donkelaar, A., Martin, R. V., Li, C., & Burnett, R. T. (2019). Regional estimates of chemical composition of fine particulate matter using a combined geoscience-statistical method with information from satellites, models, and monitors. *Environmental Science & Technology*, *53*(5), 2595–2611.

Venn, T. J., & Calkin, D. E. (2011). Accommodating non-market values in evaluation of wildfire management in the United States: Challenges and opportunities. *International Journal of Wildland Fire*, *20*(3), 327–339.

Viswanathan, S., Eria, L., Diunugala, N., Johnson, J., & McClean, C. (2006). An analysis of effects of San Diego wildfire on ambient air quality. *Journal of the Air & Waste Management Association*, *56*(1), 56–67.

Wen, Z., & Ye, B. (2014). Analyses of mediating effects: The development of methods and models. *Advances in Psychological Science*, *22*(5), 731.

Williams, A. M., & Phaneuf, D. J. (2019). The morbidity costs of air pollution: Evidence from spending on chronic respiratory conditions. *Environmental and Resource Economics*, *74*(2), 571–603.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., & others. (2012). Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *Journal of Geophysical Research: Atmospheres*, *117*(D3). https://doi.org/10.1029/2011JD016048

Yang, J., & Zhang, B. (2018). Air pollution and healthcare expenditure: Implication for the benefit of air pollution control in China. *Environment International*, *120*, 443–455.

Zhang, H., Wang, Y., Park, T.-W., & Deng, Y. (2017). Quantifying the relationship between extreme air pollution events and extreme weather events. *Atmospheric Research*, *188*, 64–79.

Zhang, X., Zhang, X., & Chen, X. (2017). Happiness in the air: How does a dirty sky affect mental health and subjective well-being? *Journal of Environmental Economics and Management*, *85*, 81–94.

Zhao, X., Lynch Jr, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, *37*(2), 197–206.

Zheng, S., Cao, J., Kahn, M. E., & Sun, C. (2014). Real estate valuation and cross-boundary air pollution externalities: Evidence from Chinese cities. *The Journal of Real Estate Finance and Economics*, *48*(3), 398–414.

Zheng, S., Zhang, X., Sun, W., & Lin, C. (2019). Air pollution and elite college graduates' job location choice: Evidence from China. *The Annals of Regional Science*, *63*(2), 295–316.

Zillow Group, Inc. (2021). *Zillow's Transaction and Assessment Database (ZTRAX)* [Data set]. Zillow Group, Inc. http://zillow.com/ztrax

APPENDIX A: HOUSE DATA PROCESSING AND SAMPLE DISTRIBUTION

1. Processing procedures for Zillow's Assessor and Real Estate Database (ZTRAX)

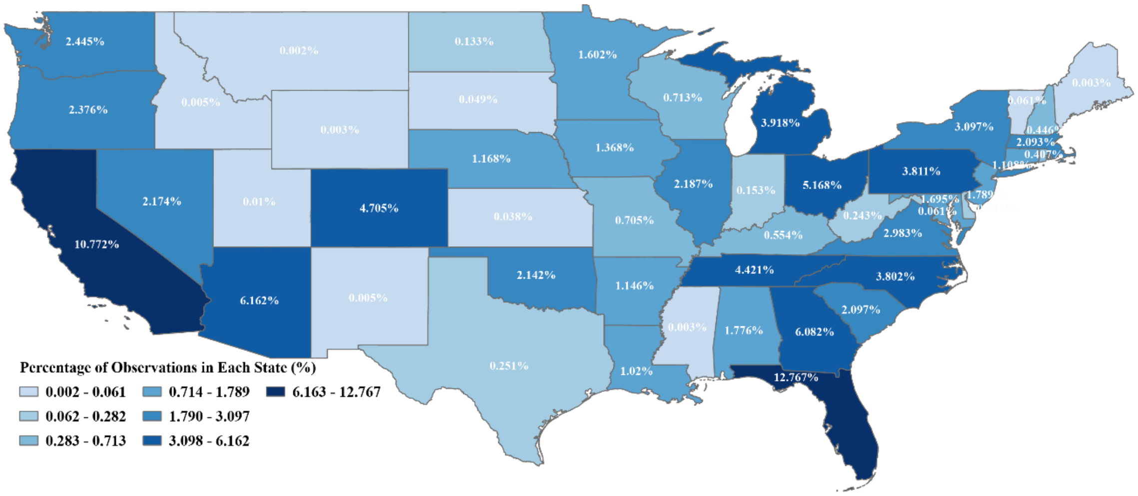
The detailed data cleaning procedures are presented below:

1) Transaction-Main table:
    i. Drop the transactions with missing sale prices.
    ii. Generate transaction date: use the DocumentDate. If DocumentDate is missing, then use SignatureDate. If DocumentDate and SignatureDate are missing, then use RecordingDate.
    iii. Keep the most recent record (with the greatest LoadID) for each TransID.
    iv. Keep the transactions that are deed transfers (DataClassStndCode is "D" or "H")
2) Transaction- PropertyInfo table:
    i. Drop the transactions with missing ImportParcelID.
    ii. Drop the transactions with missing all the geographic location information (street addresses and coordinates).
    iii. Keep the most recent record for each TransID and PropertySequenceNumber.
    iv. Drop transactions of multiple parcels.
3) Assessment-Main table:
    i. Keep the most recent record for each ImportPropertyID.
    ii. Keep the single-family residential homes (PropertyLandUseStndCode is "RR101").
4) Assessment-Building table:
    i. Drop RowID of multiple assessment records.
5) Merge the Main table and PropertyInfo table in the transaction database using FIPS and TransId.
6) Merge the Main table and Building table in the assessment database using FIPS and RowID.
7) Merge the transaction table and assessment table obtained in steps 5 and 6 using FIPS and ImportParcelID.
    i. Keep transactions that happened between 2010 and 2018.

ii. For the coordinates and street addresses of the properties, use the information in the PropertyInfo table of the transaction database. If any information is missing, use the information in the Building table of the assessment database.
iii. Drop outliers:
    - Drop if SalesPriceAmount is less than 10,000 or SalesPriceAmount is greater than 10,000,000.
    - Trim the top and bottom 1% of the data using TotalRooms and LotSizeSquareFeet (since there are too many missing values, I keep the observation with missing values). Drop houses with more than 20 rooms and houses with lot size greater than 100,000 feet.
iv. Drop the houses that were sold before the built year.
v. Drop if a house has inconsistent coordinates.
vi. Drop the houses flagged as the non-arm's length transaction (SalesPriceAmountStndCode is "NA").
vii. Drop the houses flagged as the intra-family transfer (IntraFamilyTransferFlag is "Y").
viii. Drop the houses with frequent transactions: houses with at least one transaction per year on average and houses that have two transactions within 366 days.
ix. Drop if the property has any remodeling, new construction, or major rehabilitation between the first transaction date and the last transaction date between 2010 and 2018.
x. The sale prices are adjusted using the monthly housing consumer price index from the U.S. Bureau of Labor and Statistics.
xi. I dropped the observations without coordinates.

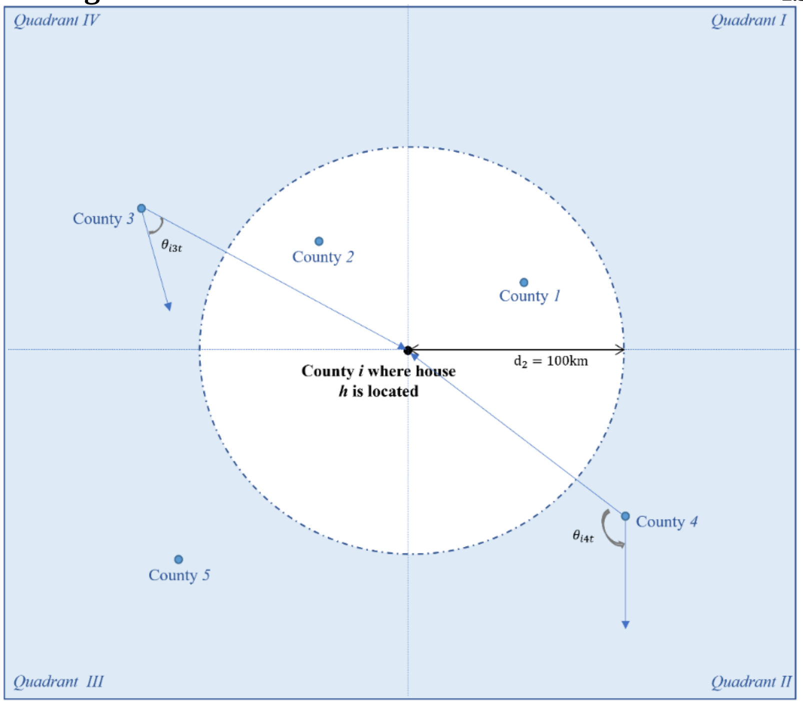
2. Sample Distribution

Figure 3. 8 Sample Distribution



APPENDIX B: INSTRUMENTAL VARIABLE CONSTRUCTION

Figure 3. 9 Construction of the Distant all-source $PM_{2.5}$



Note: (1) To construct distant all-source $PM_{2.5}$ for house h sold in month m year y, I only consider the counties in more than 100 kilometers of county i where house h is located, such as county *3* and county 4.

(2) If the angle between the wind vector at the county j's centroid and the vector from the county j's centroid to the county i's centroid (such as θ_{i3t}) is less than 90 degrees, I consider distant county j to be upwind. Otherwise, if the angle (such as θ_{i4t}) is greater than 90 degrees, county j is in a downwind direction.

APPENDIX C: ADDITIONAL ESTIMATION RESULTS

Table 3. 9 Total Effects of Wildfires
(Over the recent 12 months, 0~30km, ≥ 300 acres)

Dependent variable: ln(Price)	(1)	(2)	(3)	(4)
All wildfires	0.0008			
	(0.17)			
Upwind wildfires		-0.0311***	-0.0319***	
		(-4.55)	(-4.65)	
Downwind wildfires		0.0326***		0.0333***
		(5.19)		(5.31)
Ratio of natural-caused Wildfires	-0.0010	-0.0012	0.0000	-0.0022
	(-0.50)	(-0.60)	(0.01)	(-1.09)
Days since the most recent wildfires	0.0002***	0.0002***	0.0002***	0.0002***
	(4.85)	(4.83)	(4.87)	(4.82)
Distance to the nearest wildfires	0.0002***	0.0002***	0.0002***	0.0002***
	(6.63)	(6.65)	(6.47)	(6.81)
Humidity	-8.2037***	-8.2531***	-8.2199***	-8.2376***
	(-5.89)	(-5.93)	(-5.91)	(-5.92)
Temperature	0.0016***	0.0017***	0.0017***	0.0016***
	(3.15)	(3.18)	(3.19)	(3.14)
Pressure	0.0005	0.0007	0.0000	0.0011
	(0.05)	(0.07)	(0.00)	(0.11)
Precipitation	0.0077***	0.0078***	0.0075***	0.0080***
	(2.72)	(2.76)	(2.66)	(2.82)
Population density	-0.0274***	-0.0274***	-0.0274***	-0.0274***
	(-14.99)	(-14.99)	(-15.00)	(-14.98)
House density	-0.0001	-0.0001	-0.0001	-0.0001
	(-0.14)	(-0.13)	(-0.14)	(-0.14)
The ratio of white people	0.1661***	0.1659***	0.1659***	0.1661***
	(17.20)	(17.18)	(17.18)	(17.20)
Vegetation ratio	0.4576***	0.4571***	0.4574***	0.4574***
	(45.98)	(45.94)	(45.96)	(45.96)
Distance to WUI	0.0068***	0.0068***	0.0067***	0.0068***
	(5.99)	(5.98)	(5.95)	(6.02)
_cons	11.5829***	11.5655***	11.6263***	11.5242***
	(12.26)	(12.24)	(12.30)	(12.19)
Number of observations	3,943,418	3,943,418	3,943,418	3,943,418
Number of houses	1,885,744	1,885,744	1,885,744	1,885,744
Adjusted R2	0.8099	0.8099	0.8099	0.8099

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 10 The Effects of Wildfires on PM2.5
(Over the recent 12 months, 0~30km, ≥ 300 acres)

Dependent variable: $PM_{2.5}$	**(1)**	**(2)**	**(3)**	**(4)**
All wildfires	0.1602***			
	(13.81)			
Upwind wildfires		0.1847***	0.1817***	
		(8.33)	(8.29)	
Downwind wildfires		0.1357***		0.1316***
		(12.25)		(11.92)
Ratio of natural-caused Wildfires	0.0676***	0.0678***	0.0729***	0.0736***
	(28.18)	(28.12)	(30.79)	(31.96)
Humidity	-1.1e+02***	-1.1e+02***	-1.1e+02***	-1.1e+02***
	(-67.60)	(-67.58)	(-67.50)	(-67.65)
Temperature	0.0250***	0.0250***	0.0251***	0.0252***
	(40.90)	(40.87)	(40.94)	(41.13)
Pressure	-0.1137***	-0.1138***	-0.1162***	-0.1162***
	(-10.69)	(-10.70)	(-10.93)	(-10.92)
Precipitation	-0.2120***	-0.2121***	-0.2133***	-0.2131***
	(-71.61)	(-71.66)	(-72.09)	(-72.02)
Population density	0.0039***	0.0039***	0.0039***	0.0040***
	(4.98)	(4.98)	(4.98)	(5.01)
House density	0.0327***	0.0327***	0.0326***	0.0325***
	(12.60)	(12.61)	(12.58)	(12.55)
The ratio of white people	0.3041***	0.3043***	0.3043***	0.3033***
	(20.08)	(20.09)	(20.09)	(20.03)
Vegetation ratio	-0.1680***	-0.1676***	-0.1667***	-0.1691***
	(-11.91)	(-11.88)	(-11.82)	(-11.98)
Distance to WUI	-0.0405***	-0.0404***	-0.0406***	-0.0407***
	(-23.25)	(-23.24)	(-23.33)	(-23.40)
_cons	18.8241***	18.8369***	19.0750***	19.0670***
	(17.94)	(17.95)	(18.18)	(18.17)
Number of observations	3,943,418	3,943,418	3,943,418	3,943,418
Number of houses	1,885,744	1,885,744	1,885,744	1,885,744
Adjusted R2	0.8099	0.8099	0.8099	0.8099

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 11 Indirect Effects of Wildfires
(Over the recent 12 months, 0~30km, ≥ 300 acres)

Dependent variable: ln(Price)	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
All wildfires	0.0016	0.0049		
	(0.37)	(1.11)		
Upwind wildfires			-0.0301***	-0.0263***
			(-4.44)	(-3.93)
Downwind wildfires			0.0333***	0.0362***
			(5.31)	(5.71)
PM	-0.0051***	-0.0247***	-0.0050***	-0.0249***
	(-8.41)	(-5.59)	(-8.40)	(-5.63)
Ratio of natural-caused Wildfires	-0.0007	0.0007	-0.0009	0.0005
	(-0.33)	(0.33)	(-0.43)	(0.24)
Days since the most recent wildfires	0.0002***	0.0002***	0.0002***	0.0002***
	(4.79)	(4.54)	(4.77)	(4.52)
Distance to the nearest wildfires	0.0002***	0.0002***	0.0002***	0.0002***
	(6.77)	(7.28)	(6.80)	(7.31)
Humidity	-8.7634***	-10.9352***	-8.8114***	-11.0073***
	(-6.28)	(-7.44)	(-6.31)	(-7.49)
Temperature	0.0018***	0.0023***	0.0018***	0.0023***
	(3.39)	(4.25)	(3.42)	(4.30)
Pressure	-0.0001	-0.0022	0.0001	-0.0020
	(-0.01)	(-0.23)	(0.01)	(-0.21)
Precipitation	0.0066**	0.0024	0.0067**	0.0025
	(2.33)	(0.81)	(2.38)	(0.83)
Population density	-0.0272***	-0.0266***	-0.0273***	-0.0266***
	(-14.91)	(-14.57)	(-14.92)	(-14.58)
House density	-0.0001	0.0000	-0.00005	0.0000
	(-0.10)	(0.05)	(-0.10)	(0.06)
The ratio of white people	0.1677***	0.1737***	0.1675***	0.1735***
	(17.36)	(17.80)	(17.34)	(17.79)
Vegetation ratio	0.4568***	0.4534***	0.4563***	0.4529***
	(45.89)	(45.41)	(45.85)	(45.37)
Distance to WUI	0.0066***	0.0058***	0.0066***	0.0057***
	(5.81)	(5.04)	(5.80)	(5.02)
_cons	11.6747***		11.6571***	
	(12.36)		(12.34)	

Table 3. 11 (cont'd)

Kleibergen-Paap rk Wald F statistics		2.2e+04		2.2e+04
Endogeneity test		20.221		20.686
(p-value)		[0.0000]		[0.0000]
Number of observations	3,943,418	3,943,418	3,943,418	3,943,418
Number of houses	1,885,744	1,885,744	1,885,744	1,885,744

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 12 First-Stage Regression Results
(Over the recent 12 months, 0~30km, ≥ 300 acres)

Dependent variable: $PM_{2.5}$	$d = 80\ km$		$d = 100\ km$		$d = 120\ km$	
	(1)	(2)	(3)	(4)	(5)	(6)
Distant natural-caused	2.4377***	2.4381***	2.3600***	2.3605***	2.3013***	2.3018***
upwind wildfires	(151.21)	(151.14)	(147.27)	(147.21)	(143.99)	(143.93)
All wildfires	0.0867***		0.0949***		0.0982***	
	(10.29)		(10.92)		(11.15)	
Upwind wildfires		0.1179***		0.1265***		0.1298***
		(7.78)		(7.93)		(7.98)
Downwind wildfires		0.0555***		0.0633***		0.0665***
		(5.38)		(6.11)		(6.42)
Ratio of natural-caused	0.0453***	0.0455***	0.0489***	0.0491***	0.0502***	0.0504***
Wildfires	(19.80)	(19.77)	(21.26)	(21.21)	(21.77)	(21.72)
Days since the most	-0.0005***	-0.0005***	-0.0005***	-0.0005***	-0.0005***	-0.0005***
recent wildfires	(-10.56)	(-10.54)	(-10.43)	(-10.41)	(-10.40)	(-10.38)
Distance to the nearest	0.0007***	0.0007***	0.0007***	0.0007***	0.0007***	0.0007***
wildfires	(23.13)	(23.09)	(23.35)	(23.31)	(23.44)	(23.40)
Humidity	-73.9069***	-73.8518***	-76.2077***	-76.1512***	-78.1186***	-78.0619***
	(-44.11)	(-44.08)	(-45.54)	(-45.50)	(-46.75)	(-46.71)
Temperature	0.0119***	0.0119***	0.0128***	0.0128***	0.0136***	0.0136***
	(19.47)	(19.44)	(20.99)	(20.96)	(22.31)	(22.27)
Pressure	0.2766***	0.2765***	0.2615***	0.2614***	0.2505***	0.2504***
	(25.87)	(25.86)	(24.44)	(24.43)	(23.39)	(23.38)
Precipitation	-0.2488***	-0.2489***	-0.2477***	-0.2479***	-0.2470***	-0.2472***
	(-84.80)	(-84.83)	(-84.34)	(-84.38)	(-84.04)	(-84.08)
Population density	0.0320***	0.0320***	0.0320***	0.0320***	0.0320***	0.0320***
	(12.49)	(12.50)	(12.50)	(12.50)	(12.47)	(12.47)
House density	0.0041***	0.0041***	0.0041***	0.0041***	0.0041***	0.0041***
	(5.27)	(5.27)	(5.24)	(5.24)	(5.24)	(5.24)
The ratio of white people	0.2973***	0.2975***	0.2978***	0.2980***	0.2986***	0.2988***
	(19.69)	(19.70)	(19.72)	(19.73)	(19.77)	(19.78)
Vegetation ratio	-0.1717***	-0.1712***	-0.1716***	-0.1712***	-0.1733***	-0.1728***
	(-12.23)	(-12.19)	(-12.22)	(-12.18)	(-12.33)	(-12.30)
Distance to WUI	-0.0407***	-0.0407***	-0.0404***	-0.0404***	-0.0405***	-0.0404***
	(-23.47)	(-23.46)	(-23.32)	(-23.31)	(-23.34)	(-23.33)
_cons	-19.2655***	-19.2553***	-17.8058***	-17.7963***	-16.7560***	-16.7469***
	(-18.26)	(-18.25)	(-16.87)	(-16.86)	(-15.86)	(-15.85)

Table 3. 12 (cont'd)

F statistics	22863.36	22843.46	21689.23	21671.37	20732.15	20714.58
Number of observations	3,943,418	3,943,418	3,943,418	3,943,418	3,943,418	3,943,418
Number of houses	1,885,744	1,885,744	1,885,744	1,885,744	1,885,744	1,885,744
Adjusted R2	0.9596	0.9596	0.9595	0.9595	0.9594	0.9594

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 13 Total Effects of Wildfires
(Long-term effects, 0~30km, ≥ 300 acres)

Dependent variable: ln(Price)	3 years	5 years
	(1)	(2)
Upwind wildfires	-0.0290***	-0.0107***
	(-6.06)	(-2.93)
Downwind wildfires	0.0204***	0.0137***
	(5.13)	(3.83)
Ratio of natural-caused Wildfires	-0.0196***	-0.0377***
	(-10.34)	(-18.72)
Days since the most recent wildfires	0.0002***	0.0003***
	(4.99)	(5.32)
Distance to the nearest wildfires	0.0002***	0.0001***
	(5.96)	(4.44)
Humidity	-8.2864***	-8.1564***
	(-5.95)	(-5.86)
Temperature	0.0017***	0.0016***
	(3.20)	(3.13)
Pressure	0.0012	0.0012
	(0.13)	(0.13)
Precipitation	0.0079***	0.0082***
	(2.80)	(2.90)
Population density	-0.0275***	-0.0276***
	(-15.04)	(-15.07)
House density	-0.0001	-0.0001
	(-0.13)	(-0.13)
The ratio of white people	0.1664***	0.1667***
	(17.23)	(17.27)
Vegetation ratio	0.4556***	0.4531***
	(45.78)	(45.52)
Distance to WUI	0.0067***	0.0065***
	(5.93)	(5.78)
_cons	11.5123***	11.5153***
	(12.18)	(12.19)
Number of observations	3,943,418	3,943,418
Number of houses	1,885,744	1,885,744
Adjusted R2	0.8099	0.8099

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 14 The Effects of Wildfires on $PM_{2.5}$
(Long-term effects, 0~30km, ≥ 300 acres)

Dependent variable: $PM_{2.5}$	3 years	5 years
	(1)	(2)
Upwind wildfires	0.1408***	0.1101***
	(13.22)	(10.47)
Downwind wildfires	0.1566***	0.1191***
	(16.45)	(12.94)
Ratio of natural-caused Wildfires	-0.0300***	-0.0210***
	(-11.76)	(-7.49)
Humidity	-1.1e+02***	-1.1e+02***
	(-67.53)	(-68.05)
Temperature	0.0253***	0.0255***
	(41.29)	(41.69)
Pressure	-0.1214***	-0.1189***
	(-11.41)	(-11.17)
Precipitation	-0.2162***	-0.2163***
	(-73.10)	(-73.07)
Population density	0.0332***	0.0336***
	(12.82)	(13.00)
House density	0.0039***	0.0038***
	(4.89)	(4.82)
The ratio of white people	0.3038***	0.3008***
	(20.08)	(19.92)
Vegetation ratio	-0.1732***	-0.1719***
	(-12.27)	(-12.19)
Distance to WUI	-0.0403***	-0.0411***
	(-23.19)	(-23.62)
_cons	19.5735***	19.3143***
	(18.65)	(18.40)
Number of observations	3,943,418	3,943,418
Number of houses	1,885,744	1,885,744
Adjusted R2	0.9587	0.9587

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 15 Indirect Effects of Wildfires (Long-term effects, 0~30km, ≥ 300 acres)

Dependent variable: ln(Price)	3 years		5 years	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Upwind wildfires	-0.0282***	-0.0254***	-0.0100***	-0.0078**
	(-5.92)	(-5.33)	(-2.77)	(-2.15)
Downwind wildfires	0.0213***	0.0244***	0.0144***	0.0168***
	(5.32)	(5.96)	(3.99)	(4.52)
PM	-0.0051***	-0.0240***	-0.0051***	-0.0233***
	(-8.46)	(-5.49)	(-8.48)	(-5.34)
Ratio of natural-caused Wildfires	-0.0198***	-0.0203***	-0.0377***	-0.0379***
	(-10.42)	(-10.66)	(-18.76)	(-18.87)
Days since the most recent wildfires	0.0002***	0.0002***	0.0002***	0.0002***
	(4.93)	(4.68)	(5.25)	(5.00)
Distance to the nearest wildfires	0.0002***	0.0002***	0.0001***	0.0001***
	(6.11)	(6.62)	(4.60)	(5.13)
Humidity	-8.8485***	-10.9395***	-8.7245***	-10.7554***
	(-6.34)	(-7.45)	(-6.25)	(-7.32)
Temperature	0.0018***	0.0023***	0.0018***	0.0022***
	(3.45)	(4.28)	(3.38)	(4.19)
Pressure	0.0006	-0.0016	0.0007	-0.0014
	(0.07)	(-0.16)	(0.07)	(-0.15)
Precipitation	0.0068**	0.0027	0.0071**	0.0031
	(2.40)	(0.89)	(2.50)	(1.03)
Population density	-0.0273***	-0.0267***	-0.0274***	-0.0268***
	(-14.96)	(-14.63)	(-14.99)	(-14.66)
House density	-0.0000	0.0000	-0.0000	0.0000
	(-0.09)	(0.06)	(-0.09)	(0.04)
The ratio of white people	0.1680***	0.1737***	0.1683***	0.1738***
	(17.39)	(17.82)	(17.43)	(17.83)
Vegetation ratio	0.4547***	0.4513***	0.4522***	0.4490***
	(45.69)	(45.21)	(45.43)	(44.97)
Distance to WUI	0.0065***	0.0057***	0.0063***	0.0056***
	(5.74)	(5.01)	(5.59)	(4.86)
_cons	11.6084***		11.6102***	
	(12.29)		(12.29)	
Kleibergen-Paap rk Wald F statistics		2.2e+04		2.2e+04
Endogeneity test (p-value)		19.183 [0.0000]		17.886 [0.0000]
Number of observations	3,943,418	3,943,418	3,943,418	3,943,418
Number of houses	1,885,744	1,885,744	1,885,744	1,885,744

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 16 First-Stage Regression Results (Long-term effects, 0~30km, ≥ 300 acres)

Dependent variable: $PM_{2.5}$	3 years	5 years
	(1)	(2)
Distant natural-caused upwind wildfires	2.3767***	2.3803***
	(148.92)	(149.16)
Upwind wildfires	0.1279***	0.1099***
	(12.77)	(10.47)
Downwind wildfires	0.1420***	0.1194***
	(15.75)	(13.02)
Ratio of natural-caused Wildfires	-0.0284***	-0.0177***
	(-11.39)	(-6.44)
Days since the most recent wildfires	-0.0005***	-0.0006***
	(-10.87)	(-11.43)
Distance to the nearest wildfires	0.0007***	0.0008***
	(24.66)	(25.60)
Humidity	-75.7669***	-76.4071***
	(-45.26)	(-45.67)
Temperature	0.0128***	0.0131***
	(21.03)	(21.43)
Pressure	0.2589***	0.2620***
	(24.19)	(24.49)
Precipitation	-0.2508***	-0.2510***
	(-85.44)	(-85.48)
Population density	0.0326***	0.0331***
	(12.71)	(12.91)
House density	0.0040***	0.0039***
	(5.15)	(5.07)
The ratio of white people	0.2981***	0.2951***
	(19.76)	(19.60)
Vegetation ratio	-0.1764***	-0.1753***
	(-12.55)	(-12.49)
Distance to WUI	-0.0402***	-0.0409***
	(-23.19)	(-23.59)
_cons	-17.5641***	-17.8826***
	(-16.63)	(-16.94)
F statistics	22176.57	22250.16
Number of observations	3,943,418	3,943,418
Number of houses	1,885,744	1,885,744
Adjusted R2	0.9595	0.9595

Note: (1) *t* statistics in parentheses. (2) * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. 17 Change of House Prices ($) on Average due to Wildfires

Panel A. The impact of one more wildfire event			
The wildfire happened *d* km away	**(1)** ***d*=30km**	**(2)** ***d*=20km**	**(3)** ***d*=10km**
Change of weighted wildfires	1/30	1/20	1/10
Mean Adjusted Sale Prices≈$ 191089.4			
1. Local wildfire (over recent one year)			
Weighted wildfires			
Upwind wildfire	-167.52	-251.28	-502.57
Downwind wildfire	230.58	345.87	691.74
Weighted wildfires through $PM_{2.5}$			
Upwind wildfire	-20.38	-30.57	-61.15
Downwind wildfire	-10.19	-15.29	-30.57
2. Local wildfire (over recent three years)			
Weighted wildfires			
Upwind wildfire	-161.79	-242.68	-485.37
Downwind wildfire	155.42	233.13	466.26
Weighted wildfires through $PM_{2.5}$			
Upwind wildfire	-19.75	-29.62	-59.24
Downwind wildfire	-21.66	-32.49	-64.97
3. Local wildfire (over recent five years)			
Weighted wildfires			
Upwind wildfire	-49.68	-74.52	-149.05
Downwind wildfire	107.01	160.52	321.03
Weighted wildfires through $PM_{2.5}$			
Upwind wildfire	-16.56	-24.84	-49.68
Downwind wildfire	-17.84	-26.75	-53.51
Panel B. The impact of distance between the nearest wildfire and the house			
Distance increases every *d* km	***d*=30km**	***d*=20km**	***d*=10km**
House price change	1146.54	764.36	382.18
Panel C. The impact of the number of days since the most recent wildfire			
The number of days increases every *t* days	***t*=300 days**	***t*=200 days**	***t*=100 days**
House price change	114.65	76.44	38.22

Note: (1) The impacts of the distance between the nearest wildfire and house and the number of days since the most recent wildfire are obtained from the main analysis (Table 2).