Migration in the Smoke: How Wildfire Pollution Drives Household Relocation and Firm Outcomes

Christopher Knittel* Jasdeep Mandia†

March 20, 2025

Working Paper

Abstract

Wildfires in the U.S. are increasing in frequency and severity, with California experiencing the most significant impacts. While much research focuses on direct damage and health effects, this study examines how wildfire smoke influences household migration and firm performance. Using temporal variation in smoke exposure from the 2018 California wildfires, we analyze migration and economic outcomes using $PM_{2.5}$ data, demographic data and mobility at the household level, and employment, revenue, and survival data at the firm level from 2013 to 2019. A difference-in-differences approach shows that the wildfires caused a net out-migration of 68,000 households, with older adults and families with children being the most affected. Property values declined by 4.4%, amounting to a $20 billion loss. The residential sorting model estimates a mean marginal willingness to pay (MWTP) of -2.2% of property value to avoid a one $\mu g/m^3$ increase in smoke exposure, with older adults exhibiting the highest MWTP. We find that the 2018 wildfire smoke led to a 2.9% drop in employment, a 4.3% decline in sales, and a 6.8% reduction in operating firms, disproportionately affecting small businesses. We plan to examine how smoke-induced migration disrupts labor markets and firm dynamics, as household relocation reduces labor supply, shifts consumer demand, and forces firms to adapt. To understand these mechanisms, we propose a model linking smoke-induced migration to firm performance and location decisions.

*Massachusetts Institute of Technology. Email: knittel@mit.edu

†Center for Energy and Environmental Policy Research (CEEPR), Massachusetts Institute of Technology. Email: jmandia@mit.edu

We are grateful to the seminar participants in the Enviro Lunch at MIT and UMass Amherst for their helpful comments and discussions.

1 Introduction

Climate change has amplified the frequency and intensity of wildfires across the United States, with California experiencing the brunt of these changes Brown et al. (2023); Jolly et al. (2015); Parks et al. (2018); Westerling et al. (2006). According to the National Interagency Fire Center, the number of acres burned annually by wildfires in the U.S. has grown by approximately 4% per year from 1983 to 2022. In 2020 alone, over 58,000 wildfires scorched more than 10 million acres of land (nif). The immediate destruction of landscapes is only the tip of the iceberg; wildfires also impose profound social and economic costs, including property damage, firefighting expenses, and adverse health effects from polluted air. In California alone, the estimated capital damage costs from fires are USD 27.7 billion, while health-related costs reach USD 32.2 billion Wang et al. (2021). Existing research shows that wildfire smoke increases hospital admissions Heft-Neal et al. (2023); Kochi et al. (2016); Moeltner et al. (2013), lowers birth weights Jones and Berrens (2021), raises the risk of preterm births Heft-Neal et al. (2022), and reduces labor market earnings Borgschulte et al. (2024).

While much research has focused on the direct effects of wildfires, such as capital damage and health impacts, fewer studies explore the long-term consequences of persistent smoke exposure, particularly human adaptation and migration. Knittel and Krebs (2024) use county-level migration data to study the causal relationship between wildfire smoke and migration in California. Wang (2024) further examines this issue, finding that wildfire smoke exposure significantly increases out-migration, with higher-income households and younger individuals being more likely to relocate, while lower-income households exhibit greater inertia due to financial constraints.

Wildfires and their smoke have significant direct and indirect impacts on firms. Directly, firms face challenges such as infrastructure damage, utility disruptions, and rising insurance costs. Indirectly, they suffer revenue losses, labor shortages, and reduced productivity as workers relocate to avoid prolonged smoke exposure. While much of the existing literature focuses on the immediate destruction caused by wildfires, less attention has been paid to the lingering effects of smoke, which impose longer-term economic and social costs. This study shifts the focus to the indirect effects of wildfire smoke, exploring how it drives household migration and, in turn, reshapes local labor markets and firm outcomes. As households relocate, their decisions can profoundly alter the economic landscape firm dynamics.

The literature on firm-household dynamics underscores how household behavior and location influences firm location decisions and local economies. George and Waldfogel (2003) show that firms adapt to diverse household preferences, as seen in gentrifying neighborhoods where an influx of young, college-educated households leads to a rise in bars and restaurants and a decline in mom-and-pop stores. Ferreira and Wong (2020) and Qian and Tan (2021) further illustrate how household preferences and firm entry reshape labor

markets and housing demand, while Diamond and Gaubert (2022) highlight how spatial sorting exacerbates economic inequality. Almagro and Domínguez-Iino (2024) examine how households' preferences for amenities influence firm locations. Their study leverages variation in Amsterdam's tourist distribution as a demand shifter, revealing significant heterogeneity in residents' preferences for amenities and in how amenities respond to changes in demand composition. Together, these studies emphasize the interconnectedness of household behavior, firm decisions, and local economic outcomes, providing a foundation for understanding how environmental stressors like wildfire smoke can disrupt these dynamics.

This study addresses three key research questions. First, how does wildfire smoke influence household migration patterns? We examine the extent to which households relocate in response to persistent smoke exposure, exploring heterogeneity in migration responses across demographic groups (e.g., income, age, family composition). Second, what are the economic and welfare consequences of smoke-induced migration? Using a sorting model, we estimate the distributional welfare effects of smoke-induced migration, considering how households and firms adapt to changing environmental conditions and how these adaptations reshape local economies. We also explore heterogeneity in marginal willingness to pay (MWTP) to avoid smoke across demographic groups, asking: do demographic attributes, such as income, age, or education, determine differences in MWTP? Third, how do household migration decisions affect firm location choices and performance? We examine how smoke-induced migration affects firm performance and location decisions, highlighting the indirect link between environmental hazards and economic outcomes.

Using novel data on individual household characteristics and relocation from 2013 to 2019, we estimate the causal effect of wildfire smoke on migration. We analyze how migration rates vary across demographic groups—such as age, family composition, and socioeconomic status—to shed light on environmental justice implications. Using the novel firm-level data, we study the impact of wildfire smoke on the firms' performance. This includes information on the firm employment, revenue, and survival data.

The analysis in this paper can be divided into four parts. First, we conduct a Difference-in-Differences (DiD) analysis at the tract level to examine the impacts on household migration. Second, we conduct DiD analysis on the firm performance like employment, sales, firms operating status, We also do the analysis by the size and sector of the firm. Third, we estimate a residential sorting model, presenting results from a representative county to understand how households relocate in response to wildfire smoke. Finally, we propose a model incorporating firm sorting, which helps us understand the mechanism of the indirect effects of smoke-induced household migration on firms' location decisions and performance. This approach allows us to disentangle how household relocation influences firms' sorting behavior and, ultimately, their economic outcomes.

This paper contributes to the growing literature on the impact of wildfires on households, including

studies such as Wang et al. (2021), Moeltner et al. (2013), Kochi et al. (2016), Heft-Neal et al. (2022), Knittel and Krebs (2024), Ang (2024), and Wang (2024). While much of this literature focuses on health and economic outcomes, fewer studies examine migration patterns. Knittel and Krebs (2024) uses aggregate migration flow data at the county level, whereas we leverage individual-level household data, similar to Wang (2024). However, our study differs in two key ways: First, we explicitly focus on the effects of wildfire smoke rather than the wildfires themselves, and second, we employ a residential sorting model to estimate welfare. This approach allows us to quantify heterogeneity in the marginal willingness to pay (mWTP) to avoid smoke, incorporate sorting behavior, and generate counterfactual scenarios to assess welfare impacts.

On the firm side, we contribute to a smaller but emerging literature, including Borgschulte et al. (2024) and Luo (2023), as well as studies on the relationship between firm location and household sorting, such as George and Waldfogel (2003), Ferreira and Wong (2020), Qian and Tan (2021), Diamond and Gaubert (2022), and Almagro and Domínguez-Iino (2024). Our contributions are threefold: first, we use firm-level data to analyze performance; second, we propose a firm-sorting model to understand how firms adjust their location decisions in response to environmental changes; and third, we explore the mechanisms linking smoke-induced household migration to firm performance. By integrating household and firm responses, we provide a comprehensive understanding of how wildfire smoke disrupts local economic dynamics through both direct and indirect channels.

The paper is structured as follows. Section 2 describes the data sources, including household-level migration data, firm-level performance metrics, and environmental data on wildfire smoke exposure. Section 3 outlines the empirical estimation strategy, detailing the difference-in-differences (DiD) approach used to identify the causal effects of smoke on household migration and firm outcomes. Section 4 presents the residential sorting model, which estimates household relocation patterns in response to smoke exposure and quantifies heterogeneity in willingness to pay to avoid smoke. Section 5 discusses the results, highlighting the economic and welfare implications of smoke exposure for households and firms. Section 6 introduces the firm sorting model (without the results), examining how firms adjust their location decisions and performance in response to smoke-induced household migration. Finally, Section 7 concludes by summarizing the key findings, discussing policy implications, and suggesting directions for future research.

2 Data

Our analysis combines multiple datasets to provide a comprehensive understanding of the impacts of wildfire smoke. For wildfire smoke exposure, we use $PM_{2.5}$ data developed by prior studies Childs et al. (2022), calculating the annual mean and median smoke levels for each census tract. To analyze household demo-

graphics and locations, we utilize consumer data from DataAxle, focusing on residents of California from 2013 to 2019. This dataset comprises approximately 100 million observations and includes demographic attributes such as the age of the household head, presence of children, marital status, income, and home value. Crucially, it allows us to track household relocations within the United States.

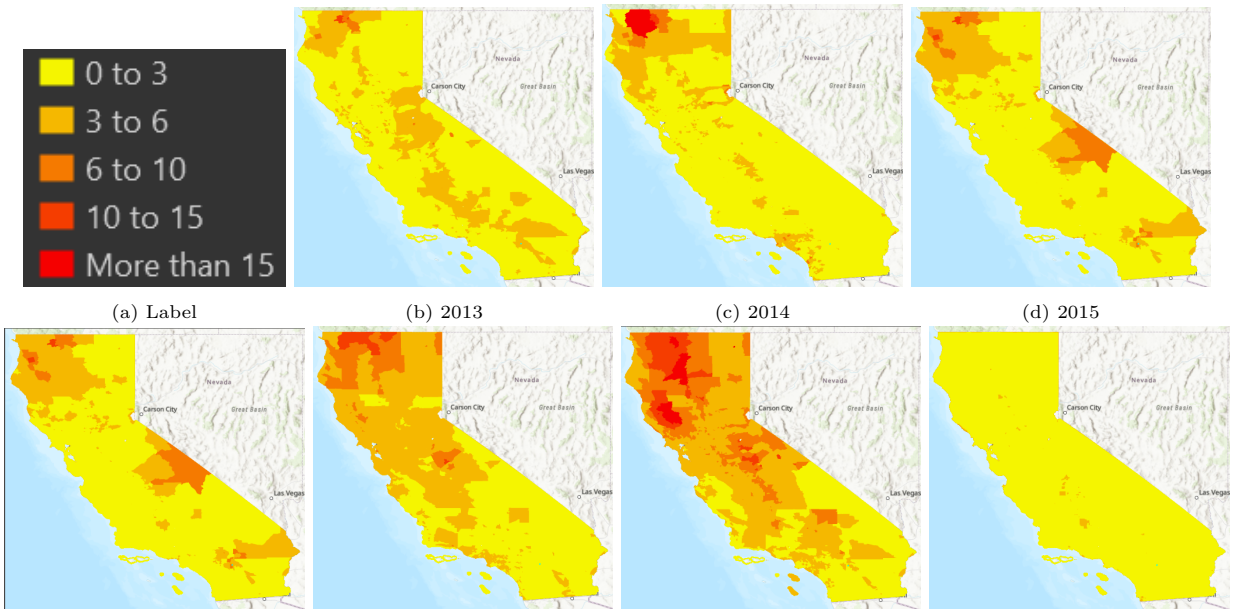
In addition to these primary datasets, we will incorporate two additional sources: (1) wildfire data from the California Department of Forestry and Fire Protection (CAL FIRE), which provides detailed information on wildfire incidents, and (2) property characteristics and transactions data, which will enhance our understanding of housing market dynamics and relocation patterns.

Currently, the data is aggregated at the census tract level to identify migration trends and analyze demographic patterns. Preliminary analyses answer the question of who moves, while future work will investigate where these households relocate. By incorporating individual-level analyses, we aim to provide a more nuanced understanding of migration and its broader implications.

Figure 1 illustrates the annual median smoke levels at the census tract level in California between 2013 and 2019, showing stable pre-period levels (2013–2016) and a marked increase in 2018. Our identification strategy relies on this temporal variation in smoke exposure caused by the significant wildfires of 2018. For the difference-in-differences (DiD) analysis, census tracts are classified as "treated" if their median smoke levels exceeded 6 $\mu g/m^3$ in 2018. Figure 2 highlights the temporal variation in treatment intensity, with a substantial increase in affected tracts in 2018, representing 12% of all tracts.

Figure 3 from Wang (2024) illustrates the increasing frequency of wildfires and mega-fires in California over the past three decades. Based on data from CAL FIRE, the graph highlights a notable rise in the number and severity of fires since 1990. Figure 4 from Knittel and Krebs (2024) presents the average annual wildfire smoke concentration ($PM_{2.5}$) categorized by quintiles of the overall distribution from 2006 to 2019. The figure shows how wildfire smoke exposure varies across regions, with higher quintiles experiencing significantly greater smoke concentrations.

Figure 5 presents the spatial distribution of household attributes at the tract level for 2017. The top two maps show the total household count per tract and the average age of residents, with urban areas tending to have a lower mean age. The bottom two maps illustrate the percentage of households with children and the percentage of owner-occupied households, highlighting variations in household composition and homeownership across different regions. Figure 6 displays tract-level firm attributes for 2017. The top two maps show the total number of firms per tract and the average number of employees per firm, indicating firm size. Most tracts comprise smaller firms, with an average of fewer than 28 employees. The bottom two maps compare firm counts in the education and manufacturing sectors. While both sectors have similar median and mean firm counts, education-related firms are more evenly distributed, whereas manufacturing



(a) Label (b) 2013 (c) 2014 (d) 2015

(e) 2016 (f) 2017 (g) 2018 (h) 2019

Figure 1: Annual median smoke at census tract level: 2013 to 2019

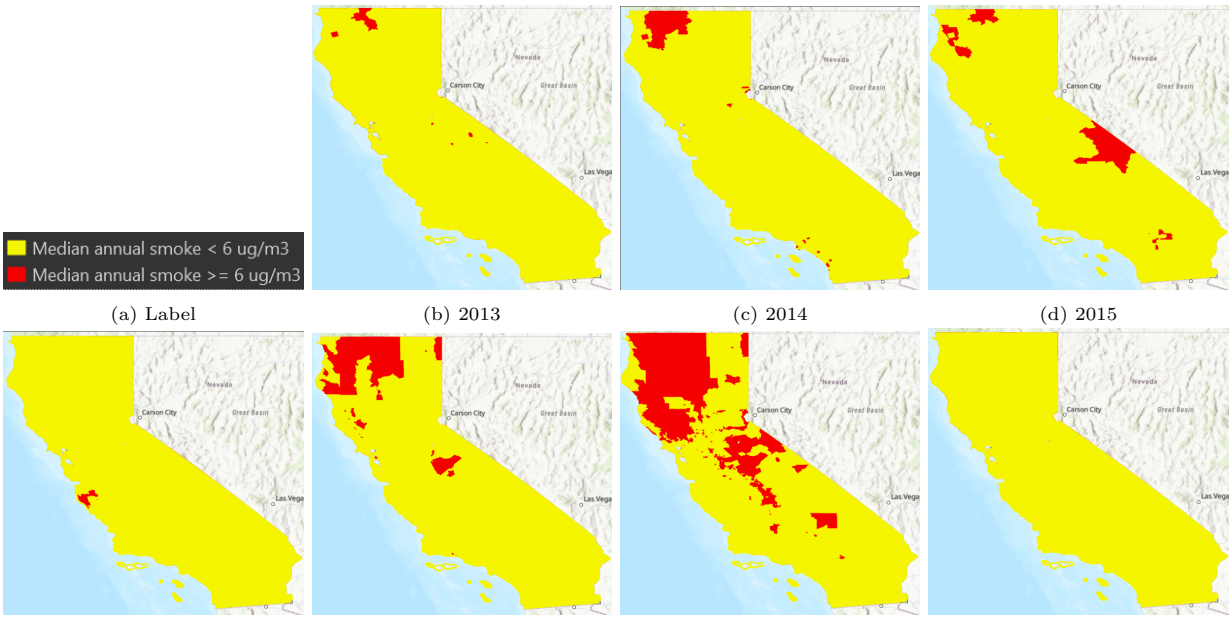
Note: These figures show the annual median smoke levels at the census tract level in California from 2013 to 2019, measured in $\mu g/m^3$. The progression of smoke levels is visually represented by a gradient of colors, ranging from yellow (lower levels) to orange and red (higher levels). While very tracts experienced smoke levels exceeding 6 $\mu g/m^3$ between 2013 and 2016, the situation changed dramatically in 2017 and 2018. In these years, particularly in 2018, more than one-third of the geographic area recorded smoke levels above 6 $\mu g/m^3$. By contrast, smoke levels in 2019 were markedly lower.

firms are more concentrated, with distinct clusters appearing in certain regions.

The reliability of consumer data has driven its increasing use in recent research. For instance, Christensen and Timmins (2023) employed DataAxle (formerly InfoUSA) consumer data, which offers detailed insights into household demographics and migration patterns. Similarly, Infutor data, another prominent consumer dataset, has been extensively utilized in studies such as Diamond et al. (2019), Diamond et al. (2020), Elliott and Wang (2023), and Wang (2024). A similar level of reliability is likely to hold for Data Axle firm data as well. Figure 7 from Cao et al. (2024) evaluates the reliability of Data Axle firm data by comparing county-level store counts for NAICS code 452 with corresponding store counts from the U.S. Census Bureau's County Business Patterns (CBP) data in 2019. The results indicate a high degree of consistency between Data Axle and CBP, reinforcing the credibility of Data Axle as a reliable source for firm-level analysis.

3 Empirical Estimation

First, we conduct a Difference-in-Differences (DiD) analysis at the tract level to examine the impacts of wildfire smoke on household migration and firm performance. The 2018 wildfires serve as an exogenous source of variation in smoke exposure, with treated tracts defined as those with a median smoke level



(a) Label (b) 2013 (c) 2014 (d) 2015

(e) 2016 (f) 2017 (g) 2018 (h) 2019

Figure 2: Treatment dummy variation at census tract level: 2013 to 2019

Note: These figures display the treatment dummy at the census tract level in California from 2013 to 2019. Treatment is assigned a value of one if the annual median smoke level exceeds 6 $\mu g/m^3$, represented in red. While the smoke levels in 2018 are used to define the treatment group across years, these maps illustrate the variation in the smoke over time. Only a small number of tracts experienced high smoke levels in earlier years (2013 to 2016), and these are excluded from our analysis. In 2018, there was a significant increase in the treatment dummy, with 12% of the tracts being affected.

exceeding 6 $\mu g/m^3$ in 2018.

The DiD specification is given by the following equation:

$$Y_{it} = \beta_0 + \beta_1 Treatment_i + \beta_2 Post_t + \beta_3 (Treatment_i \cdot Post_t) + \gamma_{city,t} + \epsilon_{it}$$

Here, β_3 is our coefficient of interest, capturing the causal effect of wildfire smoke on the outcomes of interest. On the household side, the dependent variables include the count of households (overall and disaggregated by demographic attributes such as age and marital status), as well as property value, wealth, and income. On the firm side, the dependent variables include employment, sales, and firms' operating status, further disaggregated by firm size and sector. This approach allows us to estimate the causal effects of wildfire smoke on both household and firm outcomes, providing a comprehensive understanding of its economic impacts.

4 Residential Sorting Model

4.1 Why Use a Model?

We employ a residential sorting model to address several key questions. Non-marginal changes in amenities, such as wildfire smoke, require a structural approach to understand their impacts. The model allows us to analyze how externalities, like smoke, affect housing markets, consumer choices, and firm performance. It also helps us uncover the mechanisms driving household and firm behavior. Furthermore, the model enables counterfactual analysis with sorting and price adjustments, providing insights into welfare implications.

4.2 Model Framework

We develop a static residential sorting model based on Bayer et al. (2007) to analyze household location decisions. Each household, denoted by i, derives utility U_{ht}^i from choosing housing type h in time period t. The utility function is specified as:

$$\max_{ht} U_{ht}^i = \alpha_X^i X_{ht} + \alpha_F^i F_{ht} + \alpha_S^i S_{ht} + \alpha_p^i p_{ht} + \xi_{ht} + \epsilon_{ht}^i \tag{1}$$

The utility depends on several observed and unobserved components. The observed components include X_{ht}, which represents the characteristics of the property and neighborhood; p_{ht}, the price of property h at time t; F_{ht}, amenities provided by firms in the area; and S_{ht}, the level of smoke exposure at property h during or after the event. The unobserved components include ξ_{ht}, which captures unobserved utility from the property type and neighborhood at time t, and ϵ_{ht}^i, an idiosyncratic taste shock for household i at time t.

4.3 Heterogeneity in Preferences

Household preferences are allowed to vary with demographic characteristics, denoted by z_t^i. The preference parameters α_j^i for each choice characteristic j (where $j \in \{X, p, F, S\}$) are modeled as:

$$\alpha_j^i = \alpha_{0j} + \sum_{k=1}^{K} \alpha_{kj} z_{kt}^i \tag{2}$$

Here, α_{0j} represents the mean preference for characteristic j, while α_{kj} captures the heterogeneity in preferences based on household characteristics z_{kt}^i. For example, households with senior citizens may place a higher value on cleaner air compared to other households. The utility function can then be decomposed into mean utility (δ_{ht}) and deviations from the mean utility (λ_{ht}^i):

$$U_{ht}^{i} = \underbrace{\delta_{ht}}_{\text{mean utility}} + \underbrace{\lambda_{ht}^{i} + \epsilon_{ht}^{i}}_{\text{deviation from mean utility}} \tag{3}$$

The idiosyncratic preference shocks, ϵ_{ht}^{i}, are assumed to be independently and identically distributed (iid) and drawn from a Type 1 Extreme Value (T1EV) distribution.

4.4 Econometric Implementation

The econometric implementation of the model combines Equations (2) and (3) to estimate the parameters of interest. The probability that household i chooses housing type h at time t is given by:

$$P_{ht}^{i} = \frac{\exp\left(\delta_{ht} + \lambda_{ht}^{i}\right)}{\sum_{k} \exp\left(\delta_{kt} + \lambda_{kt}^{i}\right)} \tag{4}$$

The likelihood function for the model is constructed as:

$$L = \sum_{T} \sum_{I} \sum_{H} \left(I_{ht}^{i} \ln(P_{ht}^{i}) \right) \tag{5}$$

Estimation proceeds in two steps. In the first step, maximum likelihood estimation (MLE) and contraction mapping are used to estimate the heterogeneous parameters in λ_{ht}^{i} and the mean indirect utilities (δ_{ht}). In the second step, the mean indirect utilities are regressed on the observed characteristics of the property types:

$$\delta_{ht} = \alpha_{0X} X_{ht} + \alpha_{0p} p_{ht} + \alpha_{0F} F_{ht} + \alpha_{0S} S_{ht} + \xi_{ht} \tag{6}$$

This step estimates the mean preference parameters (α_{0j}) while addressing endogeneity concerns through property type fixed effects and instrumental variables (IVs).

4.5 Challenges in Estimation

The second stage of estimation faces two primary challenges. First, smoke exposure (S_{ht}) is likely correlated with unobserved utility (ξ_{ht}). To address this, we use temporal variation in smoke exposure, which is assumed to be quasi-random. Second, property prices (p_{ht}) may also be correlated with ξ_{ht}. To mitigate this, we employ donut instrumental variables (IVs) that exploit spatial variation in prices while excluding nearby properties that may share unobserved characteristics.

5 Results

5.1 Household Migration and Property Values

Our DiD analysis reveals significant impacts of the 2018 wildfire smoke on household migration and property values. The smoke caused approximately 68,000 households to relocate, with older households and families with children being the most sensitive to smoke exposure. Disaggregating by age groups reveals that households with the oldest heads of family are the most sensitive group, causing 20,000 households to move out. Households with children move out more as well. We also observe a capitalization effect, with property values declining by 4.5%, amounting to a total loss of USD 20 billion, consistent with findings from Wang et al. (2021).

Table 1 presents the regression results for the count of households by age group. The coefficient of interest, β_3, shows a significant decline in household counts across all age groups, with the largest effect observed for households aged 65 and above. Table 2 further disaggregates the results by household status, showing significant declines for homeowners and households with children. Finally, Table 3 highlights the impact on property values, wealth, and income, with significant declines in home values and wealth.

Table 1: Regression Results: Count of Households by Age

	Count of HH				
	All	Less than 35	35 to 50	50 to 64	More than 65
Coef.	-70.11**	-6.16	-21.86*	-23.14	-18.95***
	(31.57)	(7.72)	(11.70)	(12.27)	(6.0345)
Observations	54,110	54,110	54,110	54,110	54,110
Adj R-sq	0.86	0.89	0.82	0.80	0.92
Perc.	100%	28%	30%	21%	20%

Table 2: Regression Results: Households by Status

	Count of HH		Log count of HH	
	Owner	HH with child	Owner	HH with child
Coef.	-72.99***	-31.48***	-0.0169**	-0.0195**
	(8.78)	(4.61)	(0.0083)	(0.0080)
Observations	54,110	54,110	54,110	54,110
Adj R-sq	0.98	0.95	0.99	0.98

Figure 8 presents the pre-intervention parallel trends for the household data used in the difference-in-differences (DiD) regression. The graphs display trends in the total household count, household distribution

by age group, and key economic indicators such as mean income, mean wealth, and mean home value. The consistency of these trends before the intervention supports the validity of the DiD approach.

Table 3: Regression Results: Household Property Value and Wealth

	Absolute measure ('000)			Log measure		
	Income	Home value	Wealth	Income	Home value	Wealth
Coef.	-0.60	-10.35*	-16.32*	-0.0153	-0.0423***	-0.0104*
	(0.80)	(5.67)	(8.99)	(0.0107)	(0.0162)	(0.0060)
Observations	54,110	54,110	54,110	54,110	46,227	54,110
Adj R-sq	0.96	0.92	0.98	0.96	0.89	0.97

5.2 Firm Performance

The 2018 wildfire smoke also had significant adverse effects on firms. Total employment, total sales, and the number of operating firms declined by 2.9%, 4.3%, and 6.8%, respectively. Smaller firms, defined as those with fewer than 9 employees or less than USD 10 million in revenue, were disproportionately affected.

Table 4 presents the regression results for firm employment and sales. The coefficient of interest, β_3, indicates significant declines in total employment and sales, as well as the count of operating firms. Table 5 and Table 6 further disaggregate these results by firm size and sector, respectively. Smaller firms and sectors such as education, finance, and retail experienced the largest declines.

Table 4: Regression Results: Firm Employment and Sales

	Log Employment		Log Sales		Count of Firms
	Total	Mean	Total	Mean	
Coef	-0.0287***	0.0023	-0.0428***	-0.0165	-0.0683***
	(0.0090)	(0.0075)	(0.0129)	(0.0116)	(0.0062)
Observations	16,038	16,038	16,038	16,038	16,038
Adjusted R-squared	0.98	0.91	0.96	0.88	0.98

Table 5: Regression Results: Firm Performance by Size

Firm Size	Coefficient
Small (less than 9 employees)	-0.0944***
Medium (9-50 employees)	-0.0507***
Large (more than 50 employees)	-0.0165

Table 6: Regression Results: Firm Performance by Sector

Sector	Coef	Sector	Coef
Accommodation/Food	-0.0208**	Manufacturing	-0.0037
Construction	-0.0399***	Other Services	-0.0098
Education	-0.0693***	Prof/Sci/Tech	-0.0763***
Finance	-0.0692***	Real Estate	-0.0290**
Health	0.0166	Retail	-0.0366***
Management/Admin	-0.0944***	Wholesale	-0.0507***

5.3 Household side Residential Sorting Model Estimation Details

The model is estimated using data from San Joaquin County for the years 2016 and 2019. Property types are defined at the census tract level, resulting in 115 distinct property types. The dataset includes over 500,000 households, with demographic characteristics such as age groups (4 categories) and the presence of children (2 categories). Smoke exposure is measured using mean $PM_{2.5}$ values. The estimation involves 75 taste parameters and 7,580 mean utilities.

5.4 Household side Residential Sorting Mode Results

The results of the residential sorting model are presented in Table 7, which shows the mean marginal willingness to pay (MWTP) to avoid smoke exposure. The mean MWTP is estimated to be -2.8%, indicating that households are willing to pay 2.8% of their property value to avoid an increase of 1 microgram per cubic meter ($\mu g/m^3$) of $PM_{2.5}$. This translates to approximately \$7,000 for the average property. There is significant heterogeneity across demographic groups, with MWTP ranging from -11% for households with seniors to 3% for younger households.

Table 7: Model Results: Mean MWTP (Log Price)

All		By Age		By Children	
All	-0.0282	Less than 35	0.031	0	-0.017
		35 to 50	-0.006	1	-0.056
		50 to 65	-0.029		
		More than 65	-0.116		

The results highlight the varying sensitivity of households to smoke exposure. Older households and those with children are more likely to pay a premium to avoid smoke, while younger households are less affected. These findings underscore the importance of considering demographic heterogeneity in policy design and welfare analysis.

6 Firm Sorting Model

6.1 Why Focus on Firms?

Wildfires and their smoke significantly impact firms' performance, both directly and indirectly. Direct impacts include infrastructure damage, utility interruptions, hindered operations, and higher insurance premiums. Indirectly, firms face revenue losses, reduced employee availability, and lower productivity due to the effects of smoke and wildfires on households. However, much of the existing literature focuses on the "wildfire" itself—the immediate destruction—rather than the "smoke," which represents the lingering, pervasive, and often overlooked consequences. This study shifts the focus to the smoke, examining how its indirect effects ripple through households and firms, shaping long-term economic and social outcomes.

6.2 Model Framework

We develop a firm sorting model to analyze how smoke-induced household migration affects firms' location decisions and performance. Each firm, denoted by f, chooses to locate at location j in time period t or not operate at all. The profit function for firm f is specified as:

$$\max_{jt} \pi_{jt}^{f} = \sum_{O} \bar{\alpha}_{o}^{f} C_{ojt} + \sum_{u} \bar{\alpha}_{u}^{f} C_{ujt} + \sum_{S} \bar{\alpha}_{S}^{f} C_{sjt} + \bar{\alpha}_{p}^{f} \bar{p}_{jt} A_{f} + \bar{\alpha}_{L} L_{jt} + \bar{\xi}_{jt} + \bar{\epsilon}_{jt}^{f} \tag{7}$$

The profit depends on several observed and unobserved components. The observed components include C_{ojt}, which represents the count of households with observed demographics o within a radius d miles of location j at time t; C_{sjt}, the count of firms in sector s within a radius d miles of location j at time t; $\bar{p}_{jt}$, the price per square foot of location j at time t; L_{jt}, the labor availability at location j and time t; and A_f, the square footage area required by firm f. The unobserved components include C_{ujt}, which captures the count of households with unobserved demographics u within a radius d miles of location j at time t; $\bar{\xi}_{jt}$, the unobserved desirability of location j at time t; and $\bar{\epsilon}_{jt}^{f}$, an idiosyncratic profit shock for firm f at time t.

6.3 Heterogeneity in Preferences

Firms' preferences for location characteristics are allowed to vary with their own characteristics, denoted by $\bar{z}_{t}^{f}$. The preference parameters $\bar{\alpha}_{o}^{f}$ for each observed demographic group o are modeled as:

$$\bar{\alpha}_{o}^{f} = \bar{\alpha}_{0o} + \sum_{k=1}^{K} \bar{\alpha}_{ko} \bar{z}_{kt}^{f} \tag{8}$$

Here, $\bar{\alpha}_{0o}$ represents the mean preference for demographic group o, while $\bar{\alpha}_{ko}$ captures the heterogeneity

in preferences based on firm characteristics $\bar{z}_{kt}^{f}$. For example, firms in the retail sector may place a higher value on locations with a high density of households. The profit function can then be decomposed into mean profit ($\bar{\delta}_{jt}$) and deviations from the mean profit ($\bar{\lambda}_{jt}^{f}$):

$$\pi_{jt}^{f} = \underbrace{\bar{\delta}_{jt}}_{\text{mean profit}} + \underbrace{\bar{\lambda}_{jt}^{f} + \bar{\epsilon}_{jt}^{f}}_{\text{deviation from mean profit}} \tag{9}$$

The idiosyncratic profit shocks, $\bar{\epsilon}_{jt}^{f}$, are assumed to be independently and identically distributed (iid) and drawn from a Type 1 Extreme Value (T1EV) distribution.

6.4 Econometric Implementation

The econometric implementation of the model combines Equations (8) and (9) to estimate the parameters of interest. The probability that firm f chooses location j at time t is given by:

$$P_{jt}^{f} = \frac{\exp\left(\bar{\delta}_{jt} + \bar{\lambda}_{jt}^{f}\right)}{\sum_{k} \exp\left(\bar{\delta}_{kt} + \bar{\lambda}_{kt}^{f}\right)} \tag{10}$$

The likelihood function for the model is constructed as:

$$L = \sum_{T} \sum_{F} \sum_{J} \left(I_{jt}^{f} \ln(P_{jt}^{f}) \right) \tag{11}$$

Estimation proceeds in two steps. In the first step, maximum likelihood estimation (MLE) and contraction mapping are used to estimate the heterogeneous parameters in $\bar{\lambda}_{jt}^{f}$ and the mean indirect profits ($\bar{\delta}_{jt}$). In the second step, the mean indirect profits are regressed on the observed characteristics of the locations:

$$\bar{\delta}_{jt} = \bar{\alpha}_{0o} C_{ojt} + \bar{\alpha}_{0u} C_{ujt} + \bar{\alpha}_{0S} C_{sjt} + \bar{\alpha}_{0p} \bar{p}_{jt} + \bar{\alpha}_{0L} L_{jt} + \bar{\xi}_{jt} \tag{12}$$

This step estimates the mean preference parameters ($\bar{\alpha}_{0j}$) while addressing endogeneity concerns through location-fixed effects and instrumental variables (IVs).

6.5 Estimation

6.5.1 Two Approaches to Estimation

The model can be estimated in two ways. The first approach involves solving the household and firm sorting markets separately, as described in the previous sections. The second approach involves solving the household and firm sorting markets jointly, which allows us to account for the interdependence between household and

firm location decisions.

6.5.2 Joint Estimation of Household and Firm Sorting

In the joint estimation approach, the likelihood function combines the probabilities of household and firm choices:

$$L = \sum_{T} \left(\sum_{I} \sum_{H} \left(I_{ht}^{i} \ln(P_{ht}^{i}) \right) + \sum_{F} \sum_{J} \left(I_{jt}^{f} \ln(P_{jt}^{f}) \right) \right) \tag{13}$$

Here, P_{ht}^{i} and P_{jt}^{f} are the probabilities of household i choosing housing type h and firm f choosing location j, respectively. These probabilities are defined as:

$$P_{ht}^{i} = \frac{\exp\left(\delta_{ht} + \lambda_{ht}^{i}\right)}{\sum_{k} \exp\left(\delta_{kt} + \lambda_{kt}^{i}\right)} \tag{14}$$

$$P_{jt}^{f} = \frac{\exp\left(\bar{\delta}_{jt} + \bar{\lambda}_{jt}^{f}\right)}{\sum_{k} \exp\left(\bar{\delta}_{kt} + \bar{\lambda}_{kt}^{f}\right)} \tag{15}$$

The joint likelihood function captures the interdependence between household and firm decisions, allowing us to estimate the parameters of both models simultaneously.

6.5.3 Challenges in Joint Estimation

A key challenge in joint estimation is the potential correlation between household preferences (α_{s}^{i}) and firm preferences (α_{o}^{f}). For example, young households may prefer to live near cafes, while cafes may prefer to locate in areas with a high density of young households. This interdependence creates endogeneity, as the location decisions of households and firms are mutually reinforcing.

6.5.4 Potential Solution: Joint Distribution

To address this challenge, we model the joint distribution of household and firm preference parameters. Specifically, we assume that the preference parameters α_{s}^{i} and α_{o}^{f} follow a bivariate normal distribution:

$$\begin{pmatrix} \alpha_{s}^{i} \\ \alpha_{o}^{f} \end{pmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} \sigma_{oo} & \sigma_{os} \\ \sigma_{os} & \sigma_{ss} \end{bmatrix}\right) \tag{16}$$

Here, σ_{oo} and σ_{ss} are the variances of α_{o}^{f} and α_{s}^{i}, respectively, and σ_{os} is the covariance between the two. By estimating the joint distribution, we can account for the interdependence between household and firm preferences, addressing the endogeneity issue.

6.5.5 Advantages of Joint Estimation

The joint estimation approach provides several advantages. First, it allows us to capture the feedback loop between household and firm location decisions, providing a more accurate representation of the sorting process. Second, it enables us to estimate the covariance between household and firm preferences, which can shed light on the mechanisms driving location choices. Finally, it improves the efficiency of the estimates by leveraging the interdependence between the two markets.

7 Conclusions

Wildfires in the U.S., particularly in California, are becoming more frequent and severe, with significant social and economic consequences. While much of the existing literature focuses on the direct impacts of wildfires, such as property damage and health effects, this study shifts the focus to the indirect effects of wildfire smoke on household migration and firm performance. Using detailed data on smoke exposure, household demographics, and firm-level outcomes, we analyze how smoke influences location decisions and economic activity.

Our findings highlight the importance of considering both households and firms when assessing the impacts of wildfire smoke. Households, particularly older adults and families with children are highly sensitive to smoke exposure, leading to significant migration patterns and changes in property values. On the firm side, smoke exposure reduces employment, sales, and the number of operating businesses, with small firms being the most vulnerable. These results underscore the pervasive and interconnected nature of wildfire smoke's economic impacts.

A key contribution of this study is the development of a firm sorting model, which complements our residential sorting model for households. By examining how smoke-induced household migration affects firm location and performance, we provide a more comprehensive understanding of the feedback loop between households and firms. The joint estimation of these models allows us to address endogeneity concerns and capture the interdependence between household and firm decisions.

In conclusion, this study demonstrates that wildfire smoke has far-reaching consequences beyond the immediate destruction caused by fires. By focusing on the indirect effects of smoke, we provide new insights into the long-term economic and social costs of wildfires. These findings emphasize the need for policies that address both the direct and indirect impacts of wildfires, particularly for vulnerable populations and small businesses.

References

National interagency fire center. total wildland fires and acres (1983-2022). https://www.nifc.gov/fire-information/statistics/wildfires. Accessed: 2024-01-09.

M. Almagro and T. Domínguez-Iino. Location sorting and endogenous amenities: Evidence from amsterdam. Technical report, National Bureau of Economic Research, 2024.

Q. Q. A. Ang. Paradise lost: Population growth and wildfire mitigation. *Working paper*, 2024.

P. Bayer, F. Ferreira, and R. McMillan. A unified framework for measuring preferences for schools and neighborhoods. *Journal of political economy*, 115(4):588–638, 2007.

M. Borgschulte, D. Molitor, and E. Y. Zou. Air pollution and the labor market: Evidence from wildfire smoke. *Review of Economics and Statistics*, 106(6):1558–1575, 2024.

P. T. Brown, H. Hanley, A. Mahesh, C. Reed, S. J. Strenfel, S. J. Davis, A. K. Kochanski, and C. B. Clements. Climate warming increases extreme daily wildfire growth risk in california. *Nature*, 621(7980): 760–766, 2023.

Y. Cao, J. A. Chevalier, J. Handbury, H. Parsley, and K. R. Williams. Distributional impacts of the changing retail landscape. Technical report, Cowles Foundation for Research in Economics, Yale University, 2024.

M. L. Childs, J. Li, J. Wen, S. Heft-Neal, A. Driscoll, S. Wang, C. F. Gould, M. Qiu, J. Burney, and M. Burke. Daily local-level estimates of ambient wildfire smoke pm2. 5 for the contiguous us. *Environmental Science & Technology*, 56(19):13607–13621, 2022.

P. Christensen and C. Timmins. The damages and distortions from discrimination in the rental housing market. *The Quarterly Journal of Economics*, 138(4):2505–2557, 2023.

R. Diamond and C. Gaubert. Spatial sorting and inequality. *Annual Review of Economics*, 14(1):795–819, 2022.

R. Diamond, T. McQuade, and F. Qian. The effects of rent control expansion on tenants, landlords, and inequality: Evidence from san francisco. *American Economic Review*, 109(9):3365–3394, 2019.

R. Diamond, A. Guren, and R. Tan. The effect of foreclosures on homeowners, tenants, and landlords. Technical report, National Bureau of Economic Research, 2020.

J. R. Elliott and Z. Wang. Managed retreat: A nationwide study of the local, racially segmented resettlement of homeowners from rising flood risks. *Environmental Research Letters*, 18(6):064050, 2023.

F. V. Ferreira and M. Wong. Estimating preferences for neighborhood amenities under imperfect information. Technical report, National Bureau of Economic Research, 2020.

L. George and J. Waldfogel. Who affects whom in daily newspaper markets? *Journal of Political Economy*, 111(4):765–784, 2003.

S. Heft-Neal, A. Driscoll, W. Yang, G. Shaw, and M. Burke. Associations between wildfire smoke exposure during pregnancy and risk of preterm birth in california. *Environmental Research*, 203:111872, 2022.

S. Heft-Neal, C. F. Gould, M. L. Childs, M. V. Kiang, K. C. Nadeau, M. Duggan, E. Bendavid, and M. Burke. Emergency department visits respond nonlinearly to wildfire smoke. *Proceedings of the National Academy of Sciences*, 120(39):e2302409120, 2023.

W. M. Jolly, M. A. Cochrane, P. H. Freeborn, Z. A. Holden, T. J. Brown, G. J. Williamson, and D. M. Bowman. Climate-induced variations in global wildfire danger from 1979 to 2013. *Nature communications*, 6(1):7537, 2015.

B. A. Jones and R. P. Berrens. Prescribed burns, smoke exposure, and infant health. *Contemporary Economic Policy*, 39(2):292–309, 2021.

C. Knittel and B. Krebs. Wildfire smoke causes out-migration from california. *Working paper*, 2024.

I. Kochi, P. A. Champ, J. B. Loomis, and G. H. Donovan. Valuing morbidity effects of wildfire smoke exposure from the 2007 southern california wildfires. *Journal of Forest Economics*, 25:29–54, 2016.

T. Luo. Labor market impacts of destructive california wildfires. *Monthly Labor Review*, 2023.

K. Moeltner, M.-K. Kim, E. Zhu, and W. Yang. Wildfire smoke and health impacts: A closer look at fire attributes and their marginal effects. *Journal of Environmental Economics and Management*, 66(3): 476–496, 2013.

S. A. Parks, L. M. Holsinger, M. H. Panunto, W. M. Jolly, S. Z. Dobrowski, and G. K. Dillon. High-severity fire: evaluating its key drivers and mapping its probability across western us forests. *Environmental research letters*, 13(4):044037, 2018.

F. Qian and R. Tan. The effects of high-skilled firm entry on incumbent residents. *Stanford Institute for Economic Policy Research (SIEPR) Working Paper*, pages 21–039, 2021.

D. Wang, D. Guan, S. Zhu, M. M. Kinnon, G. Geng, Q. Zhang, H. Zheng, T. Lei, S. Shao, P. Gong, et al. Economic footprint of california wildfires in 2018. *Nature Sustainability*, 4(3):252–260, 2021.

V. Wang. Migration as climate adaptation: Evidence from california wildfires. *Working paper*, 2024.

A. L. Westerling, H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam. Warming and earlier spring increase western us forest wildfire activity. *science*, 313(5789):940–943, 2006.

A Appendix Figures

A.1 Wildfire Trends

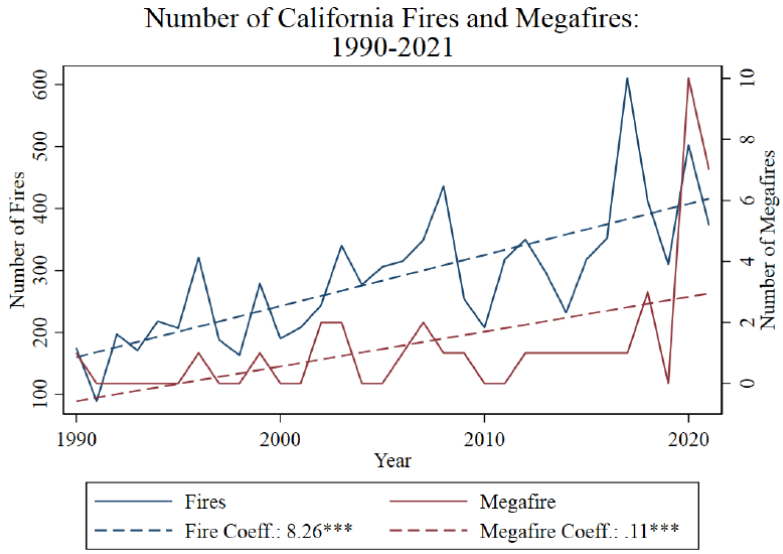


Figure 3: Count of California fires and megafires in the last three decades

Source: Wang (2024)
Note: The original data source is CAL FIRE. The graph shows an increase in the instances of both fires and mega-fires in California since 1990.

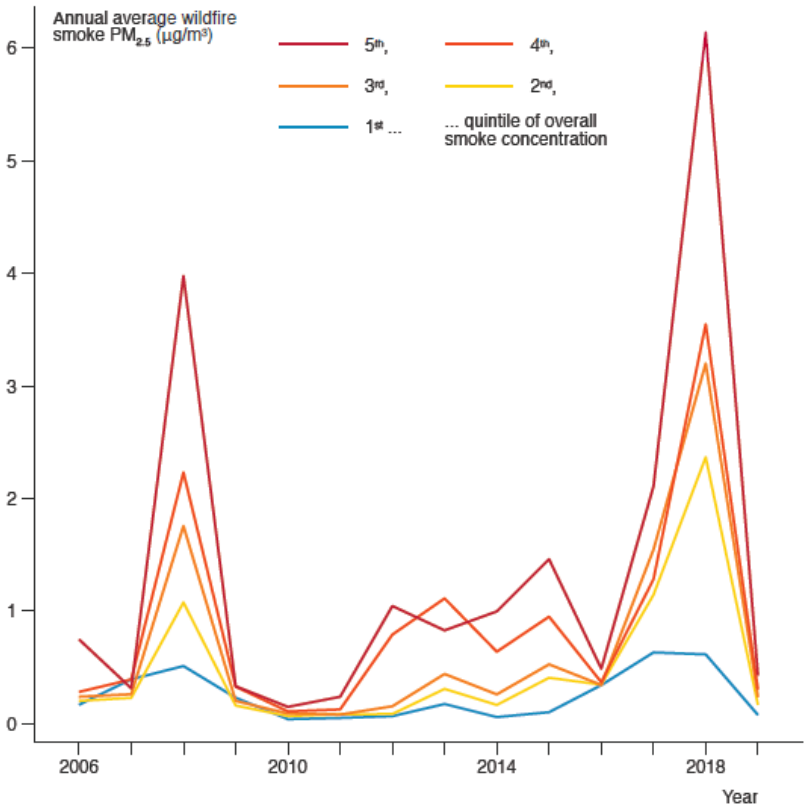


Figure 4: Average annual smoke by quintile of overall distribution

Source: Knittel and Krebs (2024)
Note: Annual average wildfire smoke $PM_{2.5}$ concentration in $\mu g/m^3$ for five quintiles of the 2006–2019 average smoke concentration.

A.2 Spatial Distribution

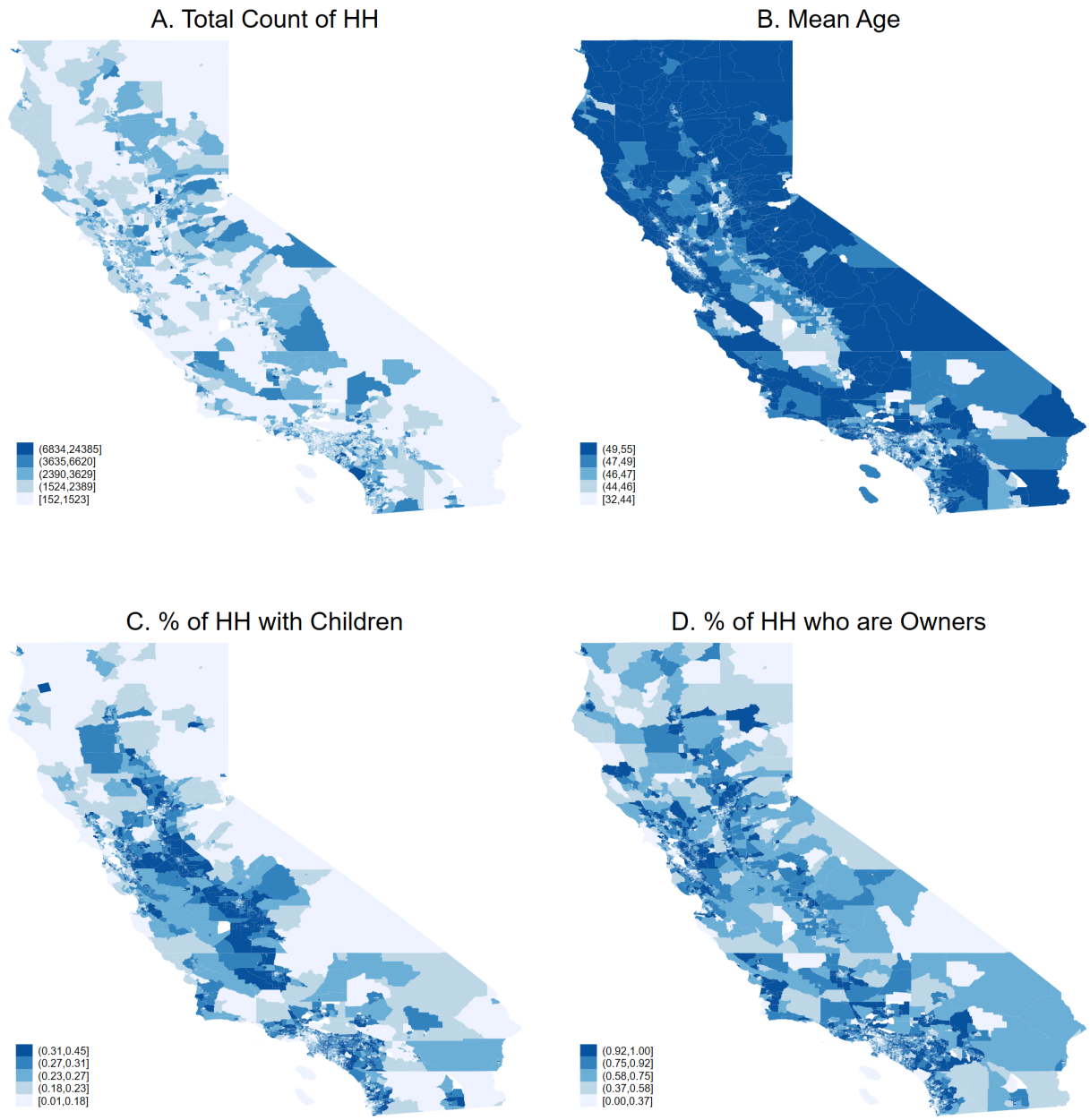


Figure 5: Tract level household attributes in 2017

Note: The set of maps shows tract-level household attributes for 2017. The top two maps depict the total household count per tract and the average age of residents. The mean age tends to be lower in and around urban areas. The bottom two maps show the percentage of households with children and the percentage of owner-occupied households.
Source of the data is DataAxle Consumer Data

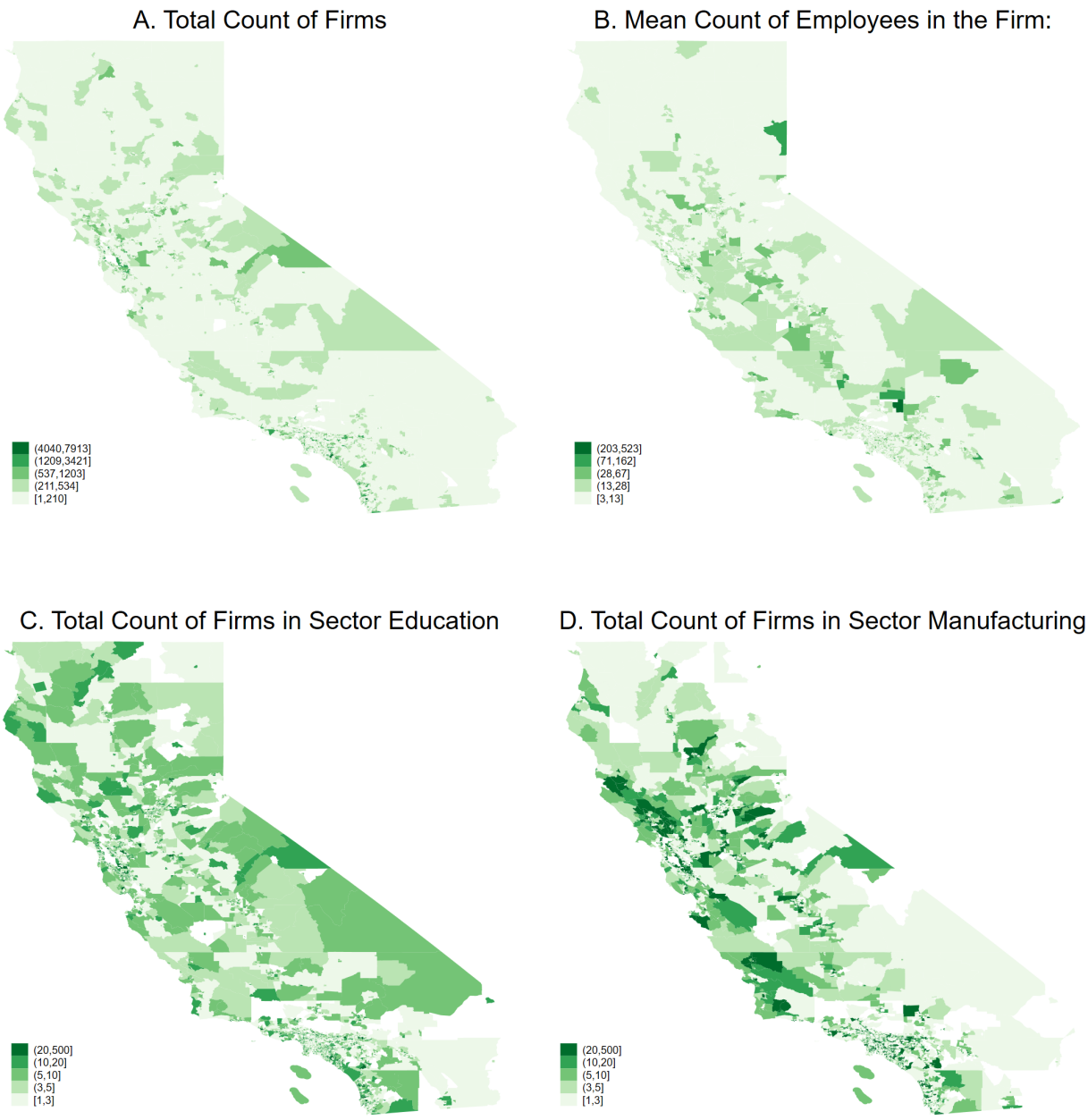


Figure 6: Tract level firm attributes in 2017

Note: The set of maps shows tract-level firm attributes for 2017. The top two maps depict the total number of firms per tract and the average number of employees per firm, which measures firm size. Most tracts are dominated by smaller firms, with an average of fewer than 28 employees. The bottom two maps compare firm counts in the education and manufacturing sectors. While the median and mean firm counts for these sectors are similar, the education sector is more evenly distributed, as indicated by consistent shades of green. In contrast, manufacturing firms are more clustered, with tracts exhibiting a mix of light and very dark green shades.
Source of the data is DataAxle Firm Data

A.3 Data Reliability

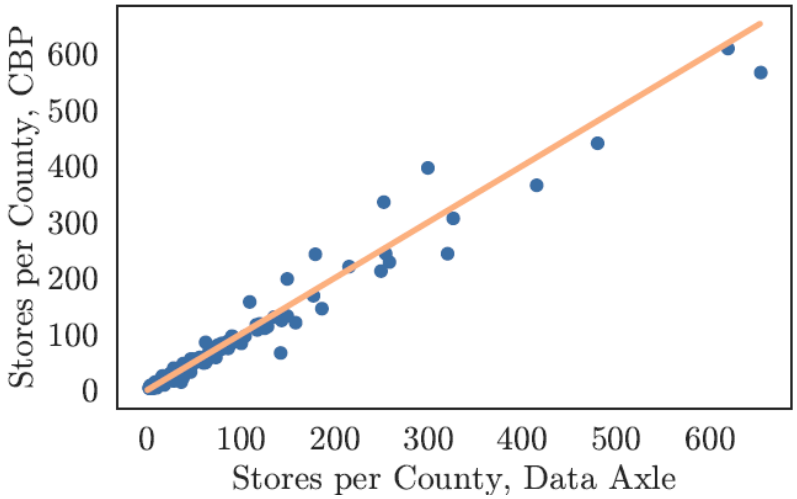


Figure 7: Count of stores from Data Axle and County Business Patterns (CBP) in 2019.

Source: Cao et al. (2024)
Note: To assess the reliability of the DataAxle firm data, Cao et al. (2024) compare county-level store counts derived from Data Axle for NAICS code 452 with the corresponding store counts from the U.S. Census Bureau's County Business Patterns (CBP) data. The results show that the county-level store counts from Data Axle are highly consistent with those from the CBP.

A.4 Parallel Trends

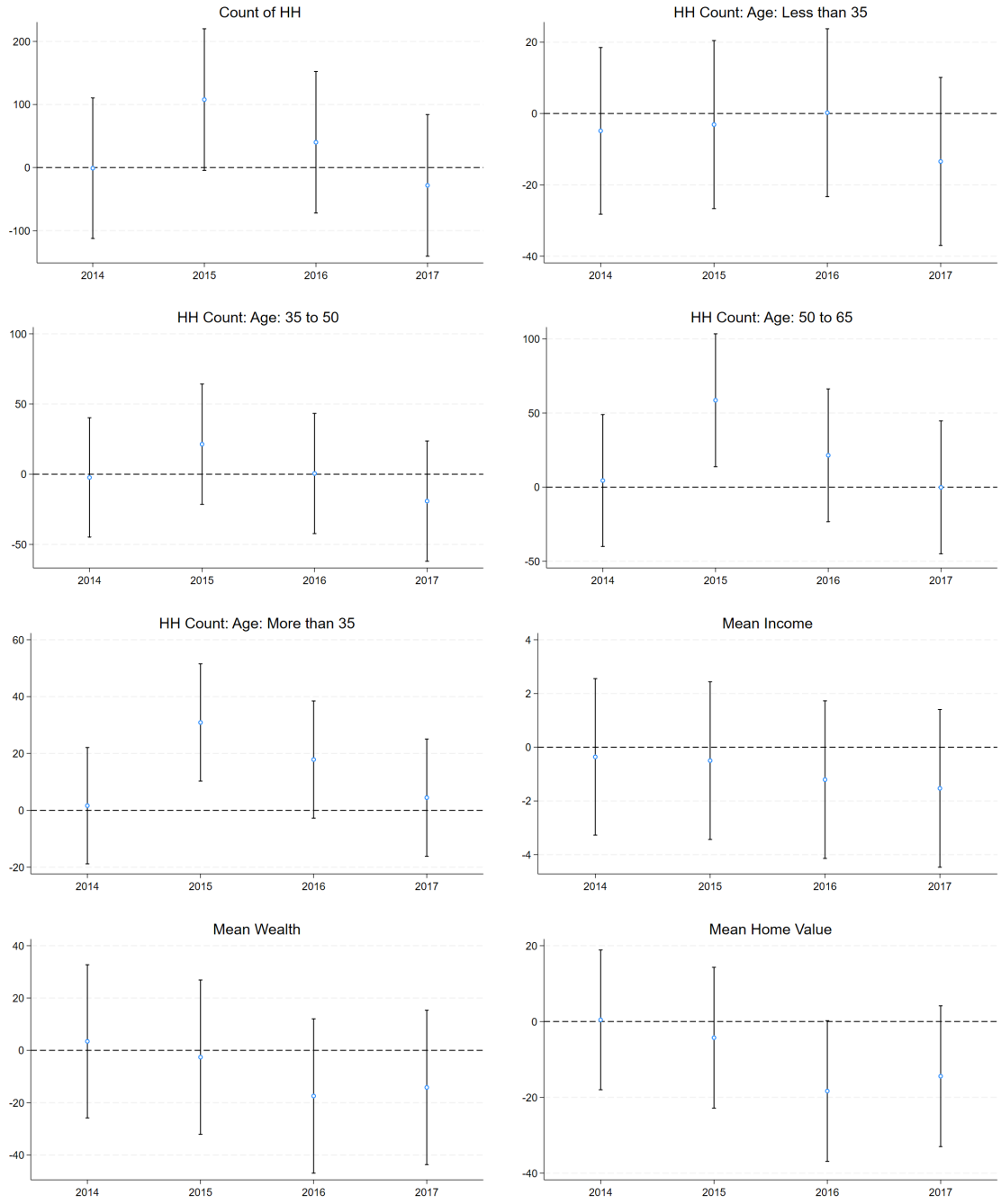


Figure 8: Pre-intervention parallel trends for the DiD regression using household data

Note: This set of graphs illustrates the pre-intervention parallel trends for the household data. It includes the total count of households, the count of households by attributes such as age group, and other census tract-level variables. These variables include mean income, mean wealth, and mean home value.