Sheltering from Wildfire Smoke: Evidence from Smart Thermostat Data

Devika Chirimar[1], Yanjun (Penny) Liao[2], and Matthew Wibbenmeyer[2]

[1]Georgetown University

[2]Resources for the Future

January 8, 2025

Preliminary draft: Please do not circulate or cite

Abstract

This paper studies people's avoidance behavior to wildfire smoke using data from smart thermostat motion sensors. We examine two margins of avoidance behaviors by users, spending more time at home and traveling away from home, and find that exposure to particulate matter pollution from wildfire smoke significantly increases both. Moreover, we find substantial heterogeneity in responses by users who live in localities with different demographics and have different home values. In particular, those with more favorable socio-economic conditions consistently have a greater propensity to travel away from home at smoke exposure, indicating uneven capacity to undertake costly avoidance actions.

1 Introduction

Wildfires in the U.S. have been increasing in frequency and intensity in recent years (Iglesias et al., 2022). When wildland fires burn, they emit an assortment of particles and gases—including carbon dioxide, carbon monoxide, methane, volatile organic compounds, nitrous oxide, nitrogen oxides, and particulate matter. Several of these pollutants are known to have harmful effects to human health, especially those of particulate pollution, a primary component of wildfire smoke emissions. Fine particulate matter known as PM2.5—so called because it consists of particles smaller than 2.5 microns in diameter—is small enough to enter the bloodstream through the lungs and cause a variety of harmful effects throughout the body. Exposure to PM2.5 has been linked to increased cardiopulmonary mortality and morbidity, increased incidence of diabetes and dementia, and adverse birth outcomes.[1]. While wildfire smoke pollution differs in temporal and toxicity characteristics from other sources[2], there are increasing direct evidence on similar health consequences, including increases in all-cause mortality, respiratory hospitalizations and emergency department visits, cardiovascular hospitalizations, and workplace injury claims (Wen et al., 2023; Heft-Neal et al., 2023; Gould et al., 2024; Cabral and Dillender, 2024).

These empirical studies generally identify health damages net of any behavioral changes by the public to minimize harm. Such avoidance behavior has been documented in other contexts of environmental risks including air and water pollution (Zivin and Neidell, 2009; Moretti and Neidell, 2011; Zivin et al., 2011). Efforts to avoid wildfire smoke are likely to be significant and widespread due to both the salience and broad geographic extent of the exposure. As avoidance behavior is itself costly, failing to take into account avoidance behavior may result in underestimation of the overall costs of environmental pollution.

One important way that individuals attempt to protect themselves from degraded air quality during periods of wildfire smoke is by increasing the time they spend indoors (often at home). However, staying indoors is only a partial defense against exposure to wildfire smoke, and the degree to which it is successful at reducing exposure depends on characteristics of the home (O'Dell et al., 2022; Krebs and Neidell, 2024). Therefore, in periods of extreme wildfire smoke, some households may choose an especially costly form of avoidance: temporarily leaving the affected area altogether.

In this study, we examine (1) whether households in California increase their time spent at home and (2) whether they are more likely to temporarily vacate their residences during periods of high pollution from wildfire smoke. Our analysis employs a unique high-frequency dataset of smart thermostat readings, which uses motion sensor(s) to detect the presence of the user(s) at home in 5-minute intervals. These data thus allow us to track individual users' time use behavior across normal times and smoke days and offer rich insights into their decision-making along the two margins of avoidance behavior. The sample spans 2017-19, which includes two of the most destructive wildfire seasons in Californian history with massive losses and widespread smoke exposure. We link the thermostat data to pollution exposure as measured by air pollution monitor readings and smoke plume extents. For the analysis, we estimate a fixed effects model of the effect of PM2.5 on user behaviors using

[1] See reviews from Feng et al. (2016); Thangavel et al. (2022); Shi et al. (2023)

[2] See Wegesser et al. (2009); Naeher et al. (2007); Aguilera et al. (2021).

both an ordinary least squares (OLS) approach and an instrumental variable (IV) framework designed to isolate pollution variation driven by wildfire smoke. Our model accounts for each user's distinct time-use pattern across day-of-week, the seasonality and trends in wildfire smoke, and growth of the user base over time.

We find that individuals engage in both types of avoidance behaviors: (1) increasing time spent indoors and (2) temporarily vacating their residences to avoid high pollution. We find that high PM2.5 levels (*i.e.* ≥ 35 $\mu g/m^3$) lead to a 14.3 percent increase in time spent at home during daytime on weekdays, relative to the mean time spent at home. The response is much weaker during evening hours, likely because there is less room for adjustment to begin with. We also find that one additional day of high PM2.5 in the previous seven days increases the probability of vacating residence by 14-29% of the mean probability of vacating residence. Notably, the IV estimates tend to be several times larger than the OLS estimates, suggesting that avoidance of pollution from wildfire smoke is stronger than that of pollution from all sources. Finally, we find that individuals avoid more hours of exposure by increasing time spent indoors compared to temporarily vacating their residences.

We also find substantial heterogeneity in these behaviors based on aggregate demographics and individual home values. The response in time spent at home is stronger among users living in lower-income and less educated Census places, while the reverse is true for traveling away from home. Given that the latter is a much costlier but more effective margin of adjustment, these findings are indicative of unequal capacity of different demographic groups to undertake avoidance behavior. In addition, we find stronger responses along both margins among users living in higher-value homes relative to the median in their county. Such disparities in avoidance behavior could lead to uneven exposure to wildfire smoke, making marginalized populations more vulnerable.

These findings add to the literature on costly avoidance behaviors regarding environmental risks. For example, Zivin and Neidell (2009) show that air quality alerts result in changes in decreased outdoor activity; Moretti and Neidell (2011) find evidence that avoidance of poor air quality results in decreased health impacts. Zivin et al. (2011) find that bottle water sales increase in places with Safe Water Drinking Act violations. More specifically, past studies find responses to wildfire smoke include increased defensive expenditures (e.g. home air purifiers, Richardson et al., 2012), labor force exits and changes in hours worked (Borgschulte et al., 2022), and declines in otherwise welfare-enhancing activities, such as recreation (e.g Gellman et al., 2023). Existing estimates indicate that costs of such behavior may be large. Borgschulte et al. (2022) estimate the annual costs of labor supply reductions due to smoke at approximately $94 billion in 2018 dollars. For those planning to visit a federal campground, Gellman et al. (2023) value welfare losses of wildfire smoke in the week of their visit at over $100 per trip. Our findings on both margins of adjustment are consistent with these studies. Traveling away from home incurs significant monetary and welfare costs. While adjustments in home presence during workdays are less costly, these insidious changes can have important labor market implications if accumulated over time.

A small set of recent studies have empirically examined home presence and travel behaviors for avoiding wildfire smoke. Most of these studies use cell phone mobility data (e.g. Safegraph) to track how exposure to smoke affects time spent at home and out-of-county travel. For example, Burke et al. (2022) study how the percentage of users staying at home the entire day changes due to smoke exposure; when PM2.5 is above 50 μg/m^3, the number of individuals in the data spending their day at home increases by about 10% on average. Holloway and Rubin (2022) use Safegraph data to study out-of-county travel in response to smoke. They find that the share of sites visited outside users' home county increases by 0.28 percentage points on average during weeks when their home county was exposed to smoke, an increase of about 1.3 percent above baseline. Both Burke et al. (2022) and Holloway and Rubin (2022) find that avoidance adaptation is increasing in Census block group household income.

While cell phone tracking data has many useful characteristics, datasets made available to researchers are generally aggregate data, which come with some downsides. Safegraph provides data on the share of users observed to have stayed entirely at home on a given day at the Census block group level. It is not possible to observe from these data changes in time spent at home, or the times of day at which time spent at home is more likely to increase. Moreover, these data do not allow inclusion of user fixed effects. If users with particular patterns of home occupancy are more likely to live in areas affected by wildfire smoke, estimates based on aggregate data could be biased. For example, rural areas, which may be more likely to be affected by wildfire smoke, are also older on average,[3] and older residents may have different capacities and/or needs to modify their behavior in response to smoke. Safegraph also provides data on the number of visits to selected places of interest (POIs), and the counties that visitors are from. Holloway and Rubin (2022) use these data to measure changes in out-of-county travel in response to smoke, but because the data are not at the individual level, it is not possible to discern whether users overall number of visits to POIs changes when they travel out-of-county.

More broadly, our findings also connect to an emerging literature on the economic impacts of wildfire smoke, such as credit card and mortgage defaults, as well as property and rent values (An et al., 2023; Addoum et al., 2199; Lopez and Tzur-Ilan, 2023).

[3] https://www.census.gov/newsroom/blogs/random-samplings/2016/12/a_glance_at_the_age.html

Our findings have important policy implications for mitigating impacts of wildfire smoke. First, the observed heterogeneity in avoidance behaviors underscores the need for targeted outreach to demographic groups that are less likely to engage in protective actions, such as renters or individuals with less flexibility in their daily schedules. Second, our results highlight the value of high-frequency, individual-level data in designing and evaluating interventions aimed at reducing pollution exposure. Incorporating these insights into urban planning and disaster response can help communities better adapt to the growing risks associated with wildfire smoke.

2 Data

2.1 Smart-Thermostat

We measure time spent at home using smart-thermostat data from Ecobee, a major smart thermostat company. Ecobee collects anonymized data from users who voluntarily participate in its Donate Your Data program to share with researchers.[4] This dataset contains high frequency data (at five minute intervals) for every participating household in its database, providing measures of temperature, thermostat setting, and whether someone is at home as detected by one or more motion detection sensor(s).

For our analysis, we restrict the sample to 6,153 households in California. To obtain more precise location of the household, we match each household's city in the database to census place using a fuzzy match based on the recorded city name. Figure A3 shows the geographic distribution of users by county, with the color indicating the number of users on the left panel and the density of users on the right. There are users in most parts of California, except for Northern California and the central Sierra region. More users reside in populous areas such as the Bay Area and Southern California - their distribution is largely reflective of the underlying population. Nevertheless, it is important to note that smart thermostat users likely differ from general population significantly in socioeconomic conditions. Unfortunately, as the data do not contain demographic variables, we cannot quantitatively assess these differences.

Each Ecobee household in our final sample has at least one motion detection sensor that is attached to the main thermostat and some households also install additional remote sensors. The data contains five-minute readings from these sensors. For tractability, we collapse the data to the half-hourly level. The motion sensors can take either TRUE or FALSE values, conditional on the value being non-missing. If any one motion sensor recorded a

[4]See `https://www.ecobee.com/en-us/donate-your-data/` for more details on the program.

"TRUE" value in a half-hour block, a user was classified as being at home (indoors) in that block. Figure A1 shows the daily pattern of minutes at home by weekday/weekend in the left panel, and by season in the right. In general, the detected activity level is the lowest between midnight and early morning, ramps up significantly between 5 to 7am, then displays a U-shape during 7am-6pm, and finally falls again after 8pm. People spend substantially more time at home during daylight hours over weekends than weekdays, and slightly less time during the evening hours. The differences across seasons are less remarkable, with people spending slightly more time at home during winter and spring. Overall, these patterns are consistent with the typical work and recreation schedule of a household, suggesting that the data is reflective of users' time use behavior.

Based on the half hourly data, we then construct two main outcome variables. Our first outcome variable is time spent at home during work (8 a.m. - 5 p.m.) and non-work (5 p.m. - 8 p.m.) segments. We divide the total number of half-hours a user is home by the total number of half-hours observed to get the proportion of time a household is home i.e. indoors. Multiplying the calculated proportion by hundred gives our measure of "percentage time spent at home". Our second outcome variable measures whether users choose to vacate their residences temporarily or not on any given day. For a user to be away from home on any day, the sensors should record "FALSE" for all half-hour blocks in that day. If a user is away for at least 2 days in a row, we classify the first day of such a sequence as the first day of temporary residence vacancy. Table A1 provides summary statistics on these measures.

The Ecobee dataset consists of 6153 households in California. 105 households are in cities for which we do not have data on PM2.5, hence, they are dropped from our sample. Of the remaining, 6043 households, 1255 households have more than 5 percent of entirely missing days. These households are likely those with the older Ecobee models that do not come with sensors. After removing them, we are left with 4793 households. Of these, 2010 households have discontinuous daily data. On dropping days with less than 10 days in the initial and end periods for each household that causes discontinuity in data, we are able to retain 74 households, leaving us with a final sample of 2857 households.

2.2 Air Pollution

We extracted PM2.5 daily summary data from EPA's Air Quality System (AQS) database for monitors in California and its neighboring states: Arizona, Nevada and Oregon. To construct a balanced panel of census places, we matched each place polygon in California (ACS 2013-2017) with monitors[5] located within 20 miles using nearest neighbor matching.

[5]For comparability in PM records, we used monitors with parameter codes 88101 and 88502. 88101 monitors use Federal Reference Methods (FRM) and 88502 are "FRM-like".

We determined the PM record for every place-date combination as a weighted PM measure by calculating weights using inverse distance weighting[6].

$$P\bar{M}2.5_{jt} = \frac{\sum_{m\in\mathcal{M}_j} w_i PM2.5_{mt}}{\sum_{m\in\mathcal{M}_j} w_i}$$

Figure A2 shows the distribution of distance between each user's city centroid and its matched monitor(s), weighted using the same scheme. For the vast majority of the users, the weighted distance falls within 5 miles.

2.3 Smoke

We use smoke plume data from NOAA's Hazard Mapping System (HMS) Fire and Smoke Analysis product. Starting in 2005, the NOAA HMS program has identified the spatial extent of smoke plumes in the US on a daily basis. The data is constructed from twice-daily visible and infrared geostationary satellite observations (NASA Geostationary Operational Environmental Satellites, GOES) using an automated, manually quality-controlled process (Rolph et al., 2009). Smoke plumes are also categorized light, medium or heavy density based on thickness of the smoke. The result of the analysis is a series of polygon shapefiles identifying locations where smoke has been observed in the air column each day. Two important limitations of these data are that they cannot distinguish between smoke high in the air column and smoke impacting surface air quality, and that they may not recognize smoke plumes concealed by cloudy days (Childs et al., 2022). These limitations would be problematic for our IV strategy if the presence of clouds, or the location of smoke within the air column, were correlated with unobserved endogenous factors related to both time spent at home and air quality. We judge the existence of such a relationship to be unlikely.

We use smoke plume data to construct a daily indicator variable for whether smoke plumes were present at each census place on each day within our sample period. We use census place polygons to define place boundaries, and designate the place has having been affected by smoke on a given day if the place intersected any smoke plumes on that day, regardless of density.

2.4 Weather

Our weather measures come from PRISM Climate Data (Oregon State University, 2023). This is a nationwide gridded data with approximately 4km resolution containing daily

[6] weight $= \frac{1}{(1+miles)^2}$. Inverse distance weighting is used for all monitors.

weather observations. For every census place, we extracted daily mean temperature (in degrees Celsius) and precipitation (in millimeters) using its centroid.

2.5 Demographics and Predicted Home Values

To examine heterogeneous responses from different demographic groups, we obtain demographic data from the 2013-2017 American Community Survey at the Census place level, including family income, the share of White population, and the share of college graduates.

We also construct a measure of predicted home values for each user based on their home square footage, age, and city as reported in the Ecobee data. To generate this measure, we first create a model of log home prices as a function of log square footage, building age, year of sales. We estimate this model for each Census place or city using home transactions data from Zillow's Transaction and Assessment Database (ZTRAX).[7] The city-specific model is then combined with input from the Ecobee user data to predict their home values.[8] It should be noted that this measure is meant to proxy for the users' socioeconomic conditions rather than an accurate estimate of their property values.

3 Empirical Design

We estimate Ecobee users' response to pollution levels using the following specification for proportion of time spent at home:

$$Y_{ijt} = \beta \cdot \mathbb{1}(PM_{jt} \geq 35) + \mathbf{X}'\theta + \phi_{id(t)} + \eta_{w(t)} + \epsilon_{ijt} \quad (1)$$

Y_{ijk} is the proportion of time spent at home for user i in city j on day t during work (8 a.m. - 5 p.m.) and non-work (5 p.m. - 8 p.m.) hours. The key regressor $\mathbb{1}(PM_{jt} \geq 35)$ is an indicator variable which equals to one when PM2.5 is greater than or equal to 35 $\mu g/m^3$ in city j on day k. According to World Health Organization guidelines, PM2.5 concentrations above 35 μg/m^3 are unhealthy for sensitive groups; below 35 μg/m^3, PM2.5 concentrations are considered "moderate." As well, the EPA National Ambient Air Quality daily standard

[7]Property assessment and transaction data were provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at `http://www.zillow.com/ztrax`. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group. In accordance with the ZTRAX Access Agreement, we have deleted raw ZTRAX data and only retained the city-specific model coefficients.

[8]For a small number of users whose home built year is missing, we repeat these steps using a model without building age. We do not predict a home value if the square footage is missing.

for PM2.5 is set at 35 μg/m^3.[9]

X is a vector of controls which includes a national holiday dummy and polynomials to the second degree of temperature and precipitation. $\phi_{id(t)}$ is a user-by-day-of-week fixed effect to capture the unique time use pattern of each user. $\eta_{w(t)}$ is a week-of-sample fixed effect. To interpret β as a causal estimate in this specification, we assume that the variation in PM2.5 is plausibly exogenous conditional on the controls. In other words, there is no unobservable factors that drive both pollution and people's time use behavior. For example, this assumption may be violated when people are traveling more during a holiday or less during the Covid shutdown period. To address these concerns, we control for national holiday and exclude 2020 and beyond from our sample.

To further isolate variation in PM2.5 caused by wildfire smoke, we use an instrumental variable (IV) approach. Specifically, we instrument for $\mathbb{1}(PM_{jt} \geq 35)$ using an indicator of the presence of smoke plume. Under this framework, the identification of β as a causal parameter requires two assumptions. First, the instrument is relevant (i.e. wildfire smoke is a strong predictor of PM2.5). Second, the exclusion restriction is satisfied (i.e. smoke does not affect users' time use patterns through channels other than pollution). Instead of estimating behavioral responses to high-pollution days in general, this approach identifies responses to pollution generated by wildfire smoke. These responses might be different given different toxicity of the fine particulates from wildfires, the episodic nature of exposure, and greater salience (Aguilera et al., 2021; Gould et al., 2024). For these reasons, we consider the estimates under the IV framework as our preferred results.

To examine whether households travel away from their homes as a pollution avoidance strategy, we use the following specification:

$$Y_{ijt} = \beta \cdot \sum_{\tau=t-7}^{t-1} \mathbb{1}(PM_{j\tau} \geq 35) + \mathbf{X}'\theta + \phi_{id(t)} + \eta_{w(t)} + \epsilon_{ijt} \tag{2}$$

Here, the outcome is whether user i in city j vacate their home on day t, as defined in Section 2.1. The sample excludes the following days when they are away. The key regressor $\sum_{\tau=t-7}^{t-1} \mathbb{1}(PM_{j\tau} \geq 35)$ is the number of days PM2.5 is greater than or equal to 35 $\mu g/m^3$ in the previous seven days. This measure is intended for representing variation in PM2.5 over a longer period. As traveling from home is a costlier avoidance strategy, we expect this decision to be motivated by a prolonged period of air pollution. The other controls are the

[9]Under the Clean Air Act, an area falls out of compliance with the Clean Air Act if 24-hour air quality falls below 35 μg/m^3 more than 2 percent of days per year over a three-year period. However, days on which smoke causes air quality to fall below this threshold can generally be excluded from assessments of regulatory compliance under the Clean Air Act's Exceptional Events Rule.

same as before, as does the identification assumption.

Similar to above, we also apply the IV framework to this outcome. This framework is potentially more important here because people are more likely to travel away from home to avoid a period of smoke pollution rather than idiosyncratic pollution days. We use two separate sets of instruments. For our just-identified IV specification, the instrument is the number of smoke days in the previous seven days. For our over-identified IV specification, we create seven indicators for having 1, 2, ..., 7 smoke days in the previous seven days.

Throughout the analysis, we cluster the standard error at the census place level to account for correlations between users in the same city and over time.

4 Results

4.1 Pollution and Time Spent at Home

Table 1 reports the results on percent time spent at home during weekdays. Column (1) presents the OLS estimates for day time (8 a.m.–5 p.m.), which suggest that a day with PM2.5 exceeding 35 $\mu g/m^3$ is associated with a 1.6 percentage point (p.p.) increase in time spent at home, equivalent to a 2.8 percent increase when compared to the outcome mean. Column (2) reports the first-stage regression under the instrumental variable framework, which uses an indicator for a smoke day to instrument for a high PM day. The coefficient is positive and highly statistically significant, with a F-statistic that is greater than 25, supporting the relevance and power of the instrument. Column (3) contains the IV estimate which implies that high PM leads users to spend 8.1 p.p. more time at home, or 14.3 percent of the mean time spent at home. A number of reasons might account for the IV estimate being much larger than the OLS estimate. For one, the instrument might have corrected the attenuation bias due to measurement errors in the high-PM day indicator by isolating relevant variation. It is also possible that the instrument picks up polluting days that are due to a nearby wildfire rather than other reasons. People might have a stronger behavioral response on these days because the pollution might be more salient. In columns (4)-(6), we use the same specifications to examine evening time spent at home (5 p.m.–8 p.m.). Here, the OLS estimate is also positive and statistically significant but only half the magnitude of the previous estimate on day time. The IV estimate is opposite in sign and statistically insignificant. In general, the response in evening hours is much weaker, possibly because people already spend most of weekday evening hours at home and there is less room for adjustment.

For the same reason as above, we would expect users who routinely spend more time at

Table 1: Pollution and Proportion of Time Spent at Home During Weekdays

	Dependent variable:					
	Day Time (8 a.m. - 5 p.m.)			*Evening Time (5 p.m. - 8 p.m.)*		
	OLS	IV 1st Stage	IV 2nd Stage	OLS	IV 1st Stage	IV 2nd Stage
	(1)	(2)	(3)	(4)	(5)	(6)
PM $\geq$ 35	1.595***		8.107***	0.753**		−0.865
	(0.282)		(2.578)	(0.298)		(2.276)
Smoke Day		0.060***			0.060***	
		(0.005)			(0.005)	
Precipitation	0.091	−0.001***	0.099***	0.040***	−0.001***	0.038***
	(0.009)	(0.0001)	(0.010)	(0.011)	(0.0001)	(0.011)
$(\text{Precipitation})^2$	−0.001***	0.00002***	−0.001***	−0.0001	0.00002***	−0.00005
	(0.0002)	(0.00000)	(0.0002)	(0.0002)	(0.00000)	(0.0002)
Temperature	0.071	−0.007***	0.116	0.091	−0.007***	0.079
	(0.076)	(0.001)	(0.079)	(0.058)	(0.001)	(0.063)
$(\text{Temperature})^2$	−0.002	0.0001***	−0.002	−0.002	0.0001***	−0.002
	(0.002)	(0.00002)	(0.002)	(0.002)	(0.00002)	(0.002)
Holiday	6.701***	−0.002	6.717***	−4.526***	−0.001	−4.528***
	(0.234)	(0.004)	(0.230)	(0.191)	(0.004)	(0.193)
Mean	56.5	0.01	56.5	71.43	0.01	71.43
User by day-of-week FE	Yes	Yes	Yes	Yes	Yes	Yes
Week-of-sample FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,346,392	1,346,392	1,346,392	1,335,638	1,335,638	1,335,638
R^2	0.432	0.220	0.432	0.325	0.219	0.325
F-stat	-	26.1	-	-	25.75	-

Notes: this table contains results from estimation of Equation (1) for time spent at home on weekdays. Vacation days (as defined by the authors previously), including one day prior to vacating residence, have been excluded. The dependent variable is the percentage of time a user was home each day during the defined time windows. Number of half hour blocks the user was home divided by the observed number of half hour blocks times 100 gives the percentage of time a user was home. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01

Table 2: Time Spent at Home by Quintiles of Baseline Time Use

	Q1	*Q2*	*Q3*	*Q4*	*Q5*
A. Day time (8 a.m. - 5 p.m.)					
$\widehat{PM \geq 35}$	14.055**	9.919*	6.935	8.059**	3.454
	(5.412)	(5.558)	(5.539)	(3.809)	(3.696)
F-stat	19.85	23.07	26.78	29.35	26.73
Mean	25.7	42.87	56.52	68.79	81.82
Observations	232,961	266,590	273,671	284,899	288,271
R^2	0.184	0.116	0.087	0.078	0.123
Controls	Yes	Yes	Yes	Yes	Yes
B. Evening time (5 p.m. - 8 p.m.)					
$\widehat{PM \geq 35}$	7.894	6.173	−5.298	−4.204	−5.802**
	(7.056)	(5.341)	(5.183)	(3.828)	(2.818)
F-stat	19	23.52	25.02	26.84	29.58
Mean	41.72	62.73	73.71	81.72	90.42
Observations	227,269	261,856	271,691	287,067	287,755
R^2	0.151	0.066	0.060	0.060	0.072
Controls	Yes	Yes	Yes	Yes	Yes

Notes: this table contains results from estimation of Equation (1) for time spent at home on weekdays by users divided into quintiles based on their mean time spent at home. Vacation days (as defined by the authors previously), including one day prior to vacating residence, have been excluded. The dependent variable is the percentage of time a user was home each day during the defined time windows. Number of half hour blocks the user was home divided by the observed number of half hour blocks times 100 gives the percentage of time a user was home. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01.

home to adjust less in response to pollution. In Table 2, we explore whether this is true by estimating the changes in time spent at home separately for users in different quintiles of baseline time spent at home. The overall pattern is as expected: users in the lowest two quintiles increase their time spent at home the most during both day and evening time. While the estimates for evening hours are not statistically significant, their magnitudes are substantial, showing a 7.9 p.p. increase for the 20% users who usually spent the least time at home and a 6.2 p.p. increase for the next 20%.

In Figure A6, we find that the increase of the share of time spent at home during day time is even across time of day. In Figure A7, we examine the dynamic effect of pollution for up to seven days using a distributed lag model. A high-pollution day significantly increases time spent at home for up to 3 days after, with the effect decreasing over time. Days 5 and 6 after see a significant decrease in time spent at home, potentially due to a compensating effect.

Overall, we find evidence that people increase their time spent at home during the day on weekdays when air pollution reach high levels, which is consistent with a motivation to avoid exposure.

4.2 Pollution and Traveling Away from Home

Next, we examine another avoidance behavioral margin: whether people leave their homes and travel elsewhere. Table 3 reports the main estimates on the effect of pollution in the previous 7 days on whether the household vacate their residence. Similar to before, we estimate Equation (2) using both OLS and IV specifications. In particular, we use two sets of instruments under the IV framework. The first is a single instrument of the number of smoke days in the last 7 days and the second consists of a set of seven indicators of having 1, 2, ..., 7 smoke days within the last 7 days, respectively. Thus, we will have two IV estimates, one just-identified and the other over-identified.

We find positive and statistically significant estimates across the OLS and IV specifications. The OLS estimate in Column (1) suggests that an additional high-pollution (*i.e.* PM2.5 exceeding $35\mu g/m^3$) day among the previous 7 days increases the probability of vacating residence by 3 basis points (b.p.). Columns (2) and (3) show that the first-stage regression for both sets of instruments is strong. In Columns (4) and (5), we find that the corresponding IV estimates are at 17 and 8 b.p., respectively. To put these magnitudes in context, these effects are equivalent to 29 and 14% of the mean probability of vacating residence, respectively.

Our results provide substantial evidence in support of short-term avoidance behavior in response to wildfire induced pollution by individuals. They align with the existing work on wildfire avoidance behavior that also suggest that people spend more time at home on days with high PM2.5. We are able to provide more detailed estimates based on the time of day when individuals engage in said avoidance. We find that most of the effect is driven by altering time at home during day time hours as opposed to evening hours. Furthermore, these avoidance behaviors are more evident during weekdays as opposed to weekends.

4.3 Heterogeneity

Households with different characteristics might have different behavioral responses to wildfire smoke due to differences in their ability to adjust time use and financial resources for travel (Holloway and Rubin, 2022; Burke et al., 2022). In this section, we explore these potential heterogeneous responses by splitting the sample of users along several key socio-economic dimensions and separately estimating the effect of wildfire-induced pollution on

Table 3: Pollution and Traveling Away from Home

	Dependent variable: Vacated Residence				
	OLS	IV 1st stage		IV 2nd stage	
	(1)	(2)	(3)	(4)	(5)
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(PM_{j\tau} \geq 35)$	0.027*			0.169***	0.081**
	(0.015)			(0.048)	(0.033)
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(\text{Smoke Day})$		0.158***			
		(0.009)			
$\mathbb{1}$(Smoke = 1 day)			−0.015***		
			(0.005)		
$\mathbb{1}$(Smoke = 2 days)			−0.072***		
			(0.009)		
$\mathbb{1}$(Smoke = 3 days)			0.083***		
			(0.028)		
$\mathbb{1}$(Smoke = 4 days)			0.243***		
			(0.059)		
$\mathbb{1}$(Smoke = 5 days)			0.604***		
			(0.038)		
$\mathbb{1}$(Smoke = 6 days)			1.354***		
			(0.069)		
$\mathbb{1}$(Smoke = 7 days)			1.635***		
			(0.072)		
Temperature	0.032***	−0.045***	−0.043***	0.037***	0.034***
	(0.010)	(0.004)	(0.004)	(0.010)	(0.010)
(Temperature)^2	−0.001***	0.001***	0.001***	−0.001***	−0.001***
	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.0003)
Precipitation	0.004*	0.002***	0.002***	0.004*	0.004*
	(0.002)	(0.0004)	(0.0004)	(0.002)	(0.002)
(Precipitation)^2	−0.0001**	−0.00003***	−0.00002*	−0.0001**	−0.0001**
	(0.00004)	(0.00001)	(0.00001)	(0.00004)	(0.00004)
Holiday	0.389***	−0.098***	−0.096***	0.405***	0.396***
	(0.046)	(0.008)	(0.008)	(0.047)	(0.046)
Mean	0.56	0.1	0.1	0.57	0.57
Week-of-sample FE	Yes	Yes	Yes	Yes	Yes
User by Day of week FE	Yes	Yes	Yes	Yes	Yes
Observations	1,905,765	1,885,720	1,885,720	1,885,720	1,885,720
R^2	0.041	0.410	0.457	0.041	0.041
F-stat	-	64.46	77.86	-	-

Notes: this table represents results from estimation of Equation (2). The dependent variable is an indicator for first day of leaving home. It takes a value 1 on the first day of the user vacating residence and 0 on all non-vacation days. The instrument for the number of days in the last 7 days when PM $\geq$ 35 µg/m^3 in column (2) is the total number of smoke days in the last 7 days. In column (3), the instrument variables are a sum of indicators, each turning on if the number of smoke days in the last 7 days is equal to the indicator number. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01

Table 4: Heterogeneous Responses Based on Census Place Demographics

	Family Income (Median = 87967)		White Population (Median = 0.62)		Bachelors or More (Median = 0.36)	
	≤ Median	*> Median*	*≤ Median*	*> Median*	*≤ Median*	*> Median*
A. Percentage time at home between 8 a.m. - 5 p.m.						
$\widehat{PM \geq} 35$	14.578**	3.792*	8.219***	9.032**	11.995**	4.506**
	(6.051)	(2.211)	(3.016)	(3.622)	(5.920)	(2.155)
Stage 1 F-stat	18.94	44.99	32.46	20.97	19.56	42.31
Mean	55.18	57.72	55.53	57.48	55.95	57.02
Observations	647,536	698,856	678,432	667,960	660,373	686,019
R^2	0.441	0.420	0.413	0.450	0.435	0.428
Controls	Yes	Yes	Yes	Yes	Yes	Yes
B. Vacated residence						
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(\widehat{PM}_{j\tau} \geq 35)$	0.137	0.185***	0.121*	0.247***	0.143	0.213***
	(0.091)	(0.056)	(0.063)	(0.085)	(0.090)	(0.052)
Stage 1 F-stat	41.42	121.05	89.57	46.59	43.25	113.37
Mean	0.54	0.59	0.56	0.57	0.49	0.64
Observations	906,460	979,260	947,664	938,056	931,182	954,538
R^2	0.047	0.036	0.034	0.048	0.044	0.039
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: This table contains IV estimates from subsamples of users residing in places with below and above median family income, proportion of white population and proportion of individuals with bachelors or a higher degree. The dependent variable is time spent at home on weekdays in Panel A and leaving home in Panel B. Vacation days (as defined by the authors previously), including one day prior to vacating residence, have been excluded. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01.

the two main outcomes under the IV framework.

In Table 4, we examine three demographic variables of the users' Census place. Column (1) presents estimates from users living in Census place where family income is below median, while column (2) shows those above median.[10] We find a much larger increase in percent of time at home during weekday day time for users from lower-income communities when compared to their counterparts from higher-income ones, but a reverse pattern for vacating residence. We observe very similar pattern in columns (5)-(6) where users in more educated communities respond less in hours at home but more in traveling away from home. When we split the sample by the share of White population in columns (3)-(4), we find that users in both sets of communities are similarly responsive in their time spent at home, but a higher share of users in Whiter communities travel away from home. These patterns are indicative of different avoidance strategies and behavioral adjustments taken by different demographic groups. In particular, we consistently find that users from more marginalized communities

[10]The median refers to the sample median, not the median income of California.

Table 5: Heterogeneous Responses Based on Predicted Home Value

	Home Value	
	$\leq$ *County Median*	$>$ *County Median*
A. Percentage time at home between 8 a.m. - 5 p.m.		
$\widehat{PM \geq} 35$	5.814*	10.835***
	(3.088)	(3.371)
Stage 1 F-stat	26.28	27.94
Mean	54.05	58.93
Observations	629,565	646,418
R^2	0.432	0.418
Controls	Yes	Yes
B. Vacated residence		
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(\widehat{PM_{j\tau}} \geq 35)$	0.068	0.251***
	(0.060)	(0.074)
Stage 1 F-stat	63.47	70.5
Mean	0.53	0.61
Observations	879,487	908,068
R^2	0.036	0.044
Controls	Yes	Yes

Notes: This table contains IV estimates from subsamples of users with predicted home values below and above their county median. The dependent variable is time spent at home on weekdays in Panel A and leaving home in Panel B. Home values are predicted using square footage. Vacation days (as defined by the authors previously), including one day prior to vacating residence, have been excluded. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01.

(i.e. lower-income, less White and educated) are significantly less likely to travel away from home, which might require significant financial means and flexibility.

An important limitation of these results, however, is that the aggregate demographics of the Census places might not be representative for Ecobee users. To address this problem, we next examine a user-specific characteristic – predicted home value. While we do not directly observe users' demographics, the Ecobee data record the square footage and age of the users' homes, which we use to predict their property values as described in Section 2.5.

In Table 5, we split the sample of users based on whether their predicted home value is below or above the median in their county and re-estimated the effects under the IV framework. For both time use outcomes, we find a significantly stronger response among users with higher property values. On high pollution days, they increase the percent of time spent at home by 10.8 p.p. as opposed to 5.8 p.p. for those living in lower-valued homes. Their probability of traveling away from home increases by 25 b.p. for an additional high-

PM day in the past seven days, compared to a much smaller and statistically insignificant increase for their peers.

In conclusion, the average estimated effects in the previous sections mask important heterogeneity in the responses. Our findings consistently show that households living in higher-income, more White and educated communities as well as those with higher-value homes have a stronger avoidance response to high particulate matter pollution due to wildfire smoke by traveling away from home. In terms of time spent at home, we find a mixed pattern, with users in marginalized communities adjust more while those lower-valued homes adjust less.

4.4 Robustness

We conduct several robustness checks of our main results.

First, we test the robustness of our result against the confounding effect of evacuation from an active fire nearby. If an active wildfire nearby leads to evacuation notices which in turn changes the time spent at home and decisions to vacate residence, our estimate would reflect not only avoidance of smoke but also direct wildfire impacts, as the two are correlated. To alleviate this concern, we conduct a simple robustness check by dropping user-days with active wildfires within 5 miles from our sample. Tables A2 and A3 report these results, which remain very similar to our main ones regarding the general direction and magnitude.

Our main results are also robust to controlling for year-month fixed effects instead of week-of-sample ones, as shown in Tables A4 and A5 in the appendix.

5 Comparison between avoidance measures

Our results provide evidence regarding two channels through which households reduce exposure to very high levels of PM2.5: by increasing time spent at home, and through temporary travel away from home, presumably to higher air quality locations. Which of these strategies is of greater consequence for reductions of PM2.5 exposure depends on the extent of uptake and the extent to which the strategy results in reductions in exposure. To compare reduced exposure due to intensive and extensive margin avoidance, we perform a back-of-the-envelope calculation in which we calculate changes in the number of hours individuals were exposed to PM2.5 greater than 35 μg/m^3 as a result of each strategy. For time spent at home, we calculate the reduction in exposure hours (ΔH^1) as:

$$\Delta H^1 = \sum_t \sum_j \mathbb{1}(\text{weekday}) \cdot \phi \cdot \overline{PM_{jt} \geq 35} \cdot N_j \cdot \frac{9\beta^{\text{D,Weekday}}}{100} \tag{3}$$

where N_j is the count of households in our sample in place j and ϕ is a parameter representing the share of time spent away from home individuals are assumed to spend outdoors. $\beta^{D,Weekday}$ represents estimated average change in the percent of time spent at home during daytime hours (D) on weekdays. Though indoor PM levels can vary widely, we assume for the purpose of this exercise that on when days PM2.5 is above 35 μg/m^3 outdoors, it is below this threshold in indoor locations. We find that users in our sample avoid 7484 hours of exposure to high PM2.5 per year by increasing time spent at home.

For temporarily vacating smoke-impacted areas, we calculate the reduction in exposure hours (ΔH^2) as:

$$\Delta H^2 = \sum_t \sum_j \mathbb{1}\big(PM_{jt} \geq 35\big) \cdot \beta \cdot \mu \cdot N_{jt} \cdot \phi \cdot \big(9(1 - \bar{Y}^{D,d(t)}) + 3(1 - \bar{Y}^{E,d(t)}) + 12\big) \tag{4}$$

Here, β represents the change in the probability a household will choose to leave home due to high PM2.5 levels, μ represents the median number of days away from home for households who leave, N_{jt} represents the number of households in the sample at place j, and ϕ again represents the share of time individuals are assumed to spend outdoors. The final term in equation 4 provides the average number of hours that a household spends outside the home on a given day, which we assume is also representative of days on which the household has left the area. We find that users in our sample avoid 5828.325 hours[11] by vacating residences.

[11]The hours vary by year. 1591.46 hours in 2017, 4151.70 hours in 2018 and 85.16 hours in 2019. The calculation assumes the number of users is constant across our sample years.

6 Conclusion

We find that individuals engage in avoidance behavior by increasing time spent at home and by traveling away from home to avoid wildfire smoke pollution. The increase in time spent at home is particularly evident during weekday daytime hours, when users may have greater margin to adjust their time spent at home. Moreover, our IV estimates indicate a particularly strong avoidance response toward smoke, as compared to high PM2.5 more generally. Those estimates indicate that high PM2.5 ($\geq 35\mu g/m^3$) caused by smoke increases the share of time spent at home by Ecobee users during weekday daytime hours by 8 percentage points on average—a 14 percent increase above users' baseline average.

Users are also more likely to spend a day or more away from home during periods of heavy air pollution due to smoke. We find that whereas on any given low pollution day the likelihood typical household leaves home for more than a day is 0.56 percentage points, this probability increases by 3 basis points on average on high pollution days, and by 8-17 basis points on high pollution days caused by smoke.

While these behaviors are undertaken voluntarily in order to minimize health and other consequences from smoke, they are nevertheless costly, and some populations may be better able to bear these costs than others. We find that smoke causes users with predicted home value (based on square footage) above the county median to more strongly increase the amount of time at home during day time hours, and to more substantially increase the likelihood of leaving home for more than one day, than for users living in homes with below median home value. Heterogeneity results for family income, white population, and college education are more equivocal, but these are based on Census place-level demographics, rather than household-level data.

A key limitation of our study is that the data are drawn from users with smart thermostats, who are likely to be wealthier than California's broader population and thus not fully representative.[12] Previous studies have used cellphone tracking data, which, due to the greater penetration of smartphone use in the overall population, is more broadly representative. However, while available cellphone data provides some information at the level of individual user, it is generally not possible to track users over time, nor do these data provide information about how time spent at home changes over the course of the day. In contrast, smart thermostat data allows us to track individual households' behavior at a fine temporal scale, enabling use of household-level fixed effects and allowing us to analyze the times of day and times of week when smoke changes behavior the most. We therefore view

[12] These data also limit the interpretation of our heterogeneity analyses, which must be viewed as studies of heterogeneous responses among the population of Ecobee users, rather than of heterogeneous responses within the population overall.

our study as a complement to these previous studies.

While this and related studies have consistently demonstrated that, in the face of significant degradations in air quality due to smoke, populations engage in costly avoidance behavior to a significant degree, there is currently little work available quantifying these costs. This is an important area for future research, which could enable inclusion of avoidance costs in estimates of overall damage costs associated with smoke. A second important area for future research is more detailed examination of behavior of smoke exposure and the capacity for avoidance behavior within disadvantaged groups. While our study and previous studies have typically pursued heterogeneity analyses using aggregate data (our estimated home value analysis is an exception), these data may mask heterogeneity at the individual or household levels. Therefore, there is a need for researchers to identify new data sources that would enable understanding responses to smoke within disadvantaged groups at the individual level.

References

Addoum, Jawad M, Dimitrios Gounopoulos, Matthew T Gustafson, Ryan Lewis, and Tam Nguyen, "The Impact of Wildfire Smoke on Real Estate Market."

Aguilera, Rosana, Thomas Corringham, Alexander Gershunov, and Tarik Benmarhnia (2021) "Wildfire smoke impacts respiratory health more than fine particles from other sources: observational evidence from Southern California," *Nature communications*, 12 (1), 1493.

An, Xudong, Stuart A Gabriel, and Nitzan Tzur-Ilan (2023) "The effects of extreme wildfire and smoke events on household financial outcomes," *Available at SSRN*, 4353113.

Borgschulte, Mark, David Molitor, and Eric Yongchen Zou (2022) "Air pollution and the labor market: Evidence from wildfire smoke," *Review of Economics and Statistics*, 1–46.

Burke, Marshall, Sam Heft-Neal, Jessica Li et al. (2022) "Exposures and behavioural responses to wildfire smoke," *Nature human behaviour*, 6 (10), 1351–1361.

Cabral, Marika and Marcus Dillender (2024) "Air Pollution, Wildfire Smoke, and Worker Health," Technical report, National Bureau of Economic Research.

Childs, Marissa L., Jessica Li, Jeffrey Wen et al. (2022) "Daily Local-Level Estimates of Ambient Wildfire Smoke PM2.5 for the Contiguous US," *Environmental Science & Technology*, 56 (19), 13607–13617, 10.1021/acs.est.2c02916.

Feng, Shaolong, Dan Gao, Fen Liao, Furong Zhou, and Xinming Wang (2016) "The health effects of ambient PM2. 5 and potential mechanisms," *Ecotoxicology and environmental safety*, 128, 67–74.

Gellman, Jacob, Margaret Walls, and Matthew Wibbenmeyer (2023) "Welfare Losses from Wildfire Smoke: Evidence from Daily Outdoor Recreation Data," Technical report, Resources for the Future.

Gould, Carlos F, Sam Heft-Neal, Mary Johnson, Juan Aguilera, Marshall Burke, and Kari Nadeau (2024) "Health effects of wildfire smoke exposure," *Annual Review of Medicine*, 75 (1), 277–292.

Heft-Neal, Sam, Carlos F Gould, Marissa L Childs, Mathew V Kiang, Kari C Nadeau, Mark Duggan, Eran Bendavid, and Marshall Burke (2023) "Emergency department visits respond nonlinearly to wildfire smoke," *Proceedings of the National Academy of Sciences*, 120 (39), e2302409120.

Holloway, M Steven and Edward Rubin (2022) "Unequal avoidance: Disparities in smoke-induced out-migration," Technical report, Department of Economics, University of Oregon.

Iglesias, Virginia, Jennifer K. Balch, and William R. Travis (2022) "U.S. fires became larger, more frequent, and more widespread in the 2000s," *Science Advances*, 8 (11), eabc0020, 10.1126/sciadv.abc0020, Published online 2022 Mar 16.

Krebs, Benjamin and Matthew Neidell (2024) "Wildfires exacerbate inequalities in indoor pollution exposure," *Environmental Research Letters*, 19 (2), 024043.

Lopez, Luis A and Nitzan Tzur-Ilan (2023) "Air pollution and rent prices: Evidence from wildfire smoke," *Available at SSRN 4537395*.

Moretti, Enrico and Matthew Neidell (2011) "Pollution, health, and avoidance behavior: evidence from the ports of Los Angeles," *Journal of human Resources*, 46 (1), 154–175.

Naeher, Luke P, Michael Brauer, Michael Lipsett, Judith T Zelikoff, Christopher D Simpson, Jane Q Koenig, and Kirk R Smith (2007) "Woodsmoke health effects: a review," *Inhalation toxicology*, 19 (1), 67–106.

Oregon State University (2023) "PRISM Climate Group," `https://prism.oregonstate.edu`.

O'Dell, Katelyn, Bonne Ford, Jesse Burkhardt, Sheryl Magzamen, Susan C Anenberg, Jude Bayham, Emily V Fischer, and Jeffrey R Pierce (2022) "Outside in: The relationship between indoor and outdoor particulate air quality during wildfire smoke events in western US cities," *Environmental Research: Health*, 1 (1), 015003.

Richardson, Leslie A, Patricia A Champ, and John B Loomis (2012) "The hidden cost of wildfires: Economic valuation of health effects of wildfire smoke exposure in Southern California," *Journal of Forest Economics*, 18 (1), 14–35.

Rolph, Glenn D, Roland R Draxler, Ariel F Stein et al. (2009) "Description and verification of the NOAA smoke forecasting system: the 2007 fire season," *Weather and Forecasting*, 24 (2), 361–378.

Shi, Liuhua, Qiao Zhu, Yifan Wang et al. (2023) "Incident dementia and long-term exposure to constituents of fine particle air pollution: A national cohort study in the United States," *Proceedings of the National Academy of Sciences*, 120 (1), e2211282119.

Thangavel, Prakash, Duckshin Park, and Young-Chul Lee (2022) "Recent insights into particulate matter (PM2. 5)-mediated toxicity in humans: an overview," *International journal of environmental research and public health*, 19 (12), 7511.

Wegesser, Teresa C, Kent E Pinkerton, and Jerold A Last (2009) "California wildfires of

2008: coarse and fine particulate matter toxicity," *Environmental health perspectives*, 117 (6), 893–897.

Wen, Jeff, Sam Heft-Neal, Patrick Baylis, Judson Boomhower, and Marshall Burke (2023) "Quantifying fire-specific smoke exposure and health impacts," *Proceedings of the National Academy of Sciences*, 120 (51), e2309325120.

Zivin, Joshua Graff and Matthew Neidell (2009) "Days of haze: Environmental information disclosure and intertemporal avoidance behavior," *Journal of Environmental Economics and Management*, 58 (2), 119–128.

Zivin, Joshua Graff, Matthew Neidell, and Wolfram Schlenker (2011) "Water quality violations and avoidance behavior: Evidence from bottled water consumption," *American Economic Review*, 101 (3), 448–453.

A Appendix

A.1 Descriptive Statistics

Table A1: Percentage Time at Home and Probability of Vacating Residence

	Days	Time at home (Morning)	Time at home (Evening)	Prob(vacate)
1	Monday-Thursday	54.76	70.62	0.006
2	Friday	56.78	66.59	0.008
3	Saturday	65.26	63.35	0.010
4	Sunday	65.82	69.61	0.007

Notes: time at home was calculated for each user-day group pair and then averaged across users. Trip days including the day prior to leaving was not included in the calculation. Probability of vacating residence was based on first day of leaving home divided by the total number of non-trip days.

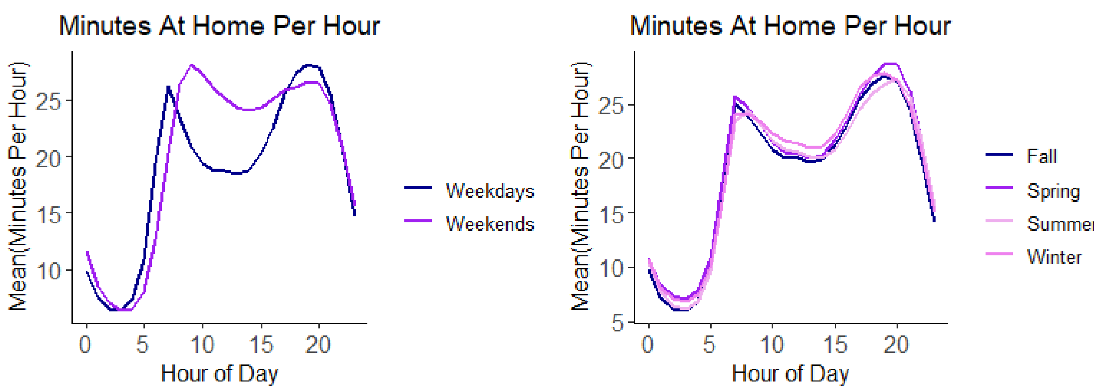


Figure A1: Daily Patterns of Minutes at Home per Hour

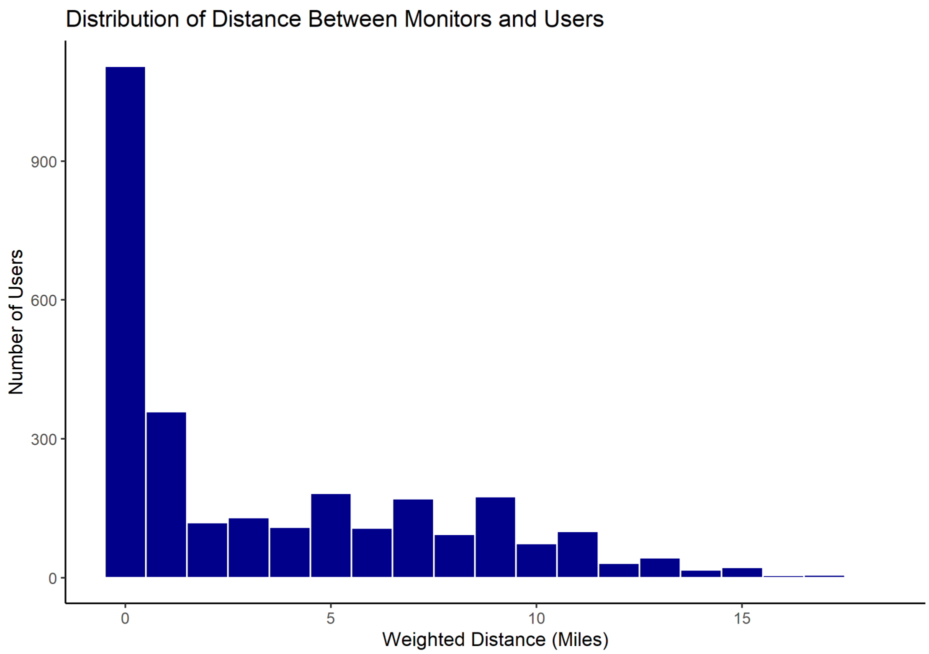
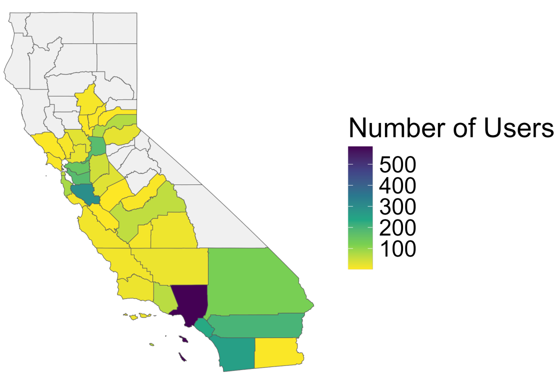
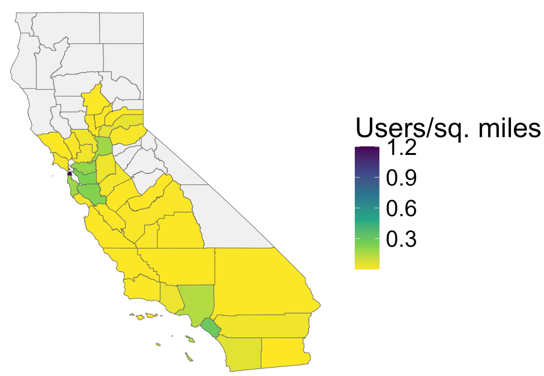


Figure A2: Distribution of Weighted Distance Between Monitors and User Cities





Number of Users in Each County **Number of Users Per Square Mile**

Figure A3: Comparison of User Metrics Across Counties

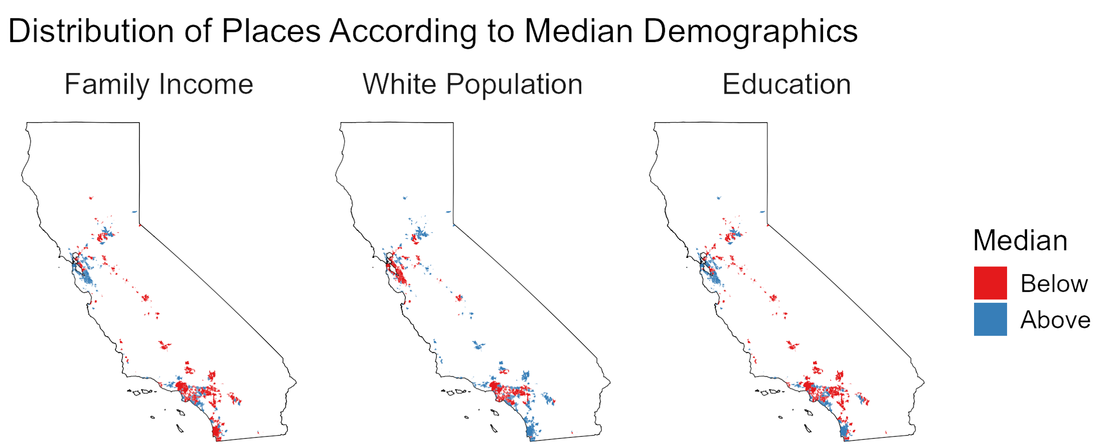


Figure A4: Distribution of Places by Demographics

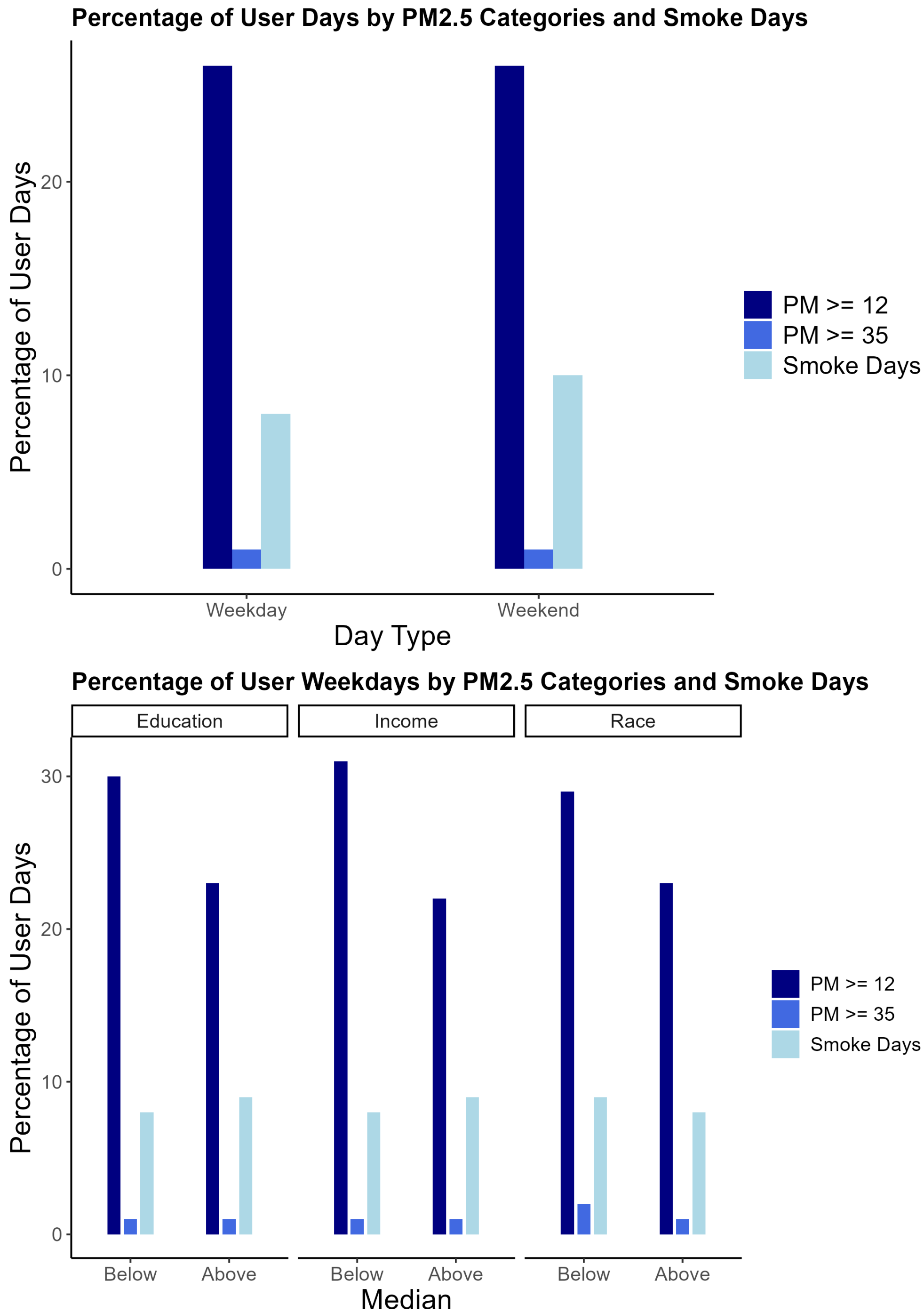


Figure A5:

Note: The statistics in this figure has been calculated using days the users did not vacate their residences. PM $\geq$ 12 corresponds to days PM2.5 was at least 12 $\mu g/m^3$ but less than 35 $\mu g/m^3$. PM $\geq$ 35 corresponds to days PM2.5 was at least as high as 35 $\mu g/m^3$. The figure in the bottom panel is filtered to include only weekdays.

A.2 Additional Tables

Table A2: Robustness Check Excluding Nearby Fire - Time Spent at Home

	Dependent variable:					
	Day Time (8 a.m. - 5 p.m.)			*Evening Time (5 p.m. - 8 p.m.)*		
	OLS	IV 1st Stage	IV 2nd Stage	OLS	IV 1st Stage	IV 2nd Stage
	(1)	(2)	(3)	(4)	(5)	(6)
$PM \geq 35$	1.475***		7.731***	0.720**		−1.435
	(0.276)		(2.157)	(0.321)		(2.111)
Smoke Day		0.061***			0.061***	
		(0.005)			(0.005)	
Precipitation	0.090***	−0.001***	0.097***	0.039***	−0.001***	0.036***
	(0.009)	(0.0001)	(0.010)	(0.011)	(0.0001)	(0.011)
$(\text{Precipitation})^2$	−0.001***	0.00002***	−0.001***	−0.0001	0.00002***	−0.00002
	(0.0002)	(0.00000)	(0.0002)	(0.0002)	(0.00000)	(0.0002)
Temperature	0.079	−0.007***	0.118	0.097	−0.007	0.083
	(0.078)	(0.001)	(0.078)	(0.061)	(0.001)	(0.064)
$(\text{Temperature})^2$	−0.002	0.0001***	−0.003	−0.002	0.0001***	−0.002
	(0.002)	(0.00003)	(0.002)	(0.002)	(0.00003)	(0.002)
Holiday	6.615	−0.004	6.645***	−4.554***	−0.003	−4.562***
	(0.231)	(0.003)	(0.230)	(0.183)	(0.003)	(0.184)
Mean	56.54	0.01	56.54	71.46	0.01	71.46
User by day-of-week FE	Yes	Yes	Yes	Yes	Yes	Yes
Week-of-sample FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,308,055	1,308,055	1,308,055	1,297,699	1,297,699	1,297,699
R^2	0.432	0.237	0.432	0.325	0.236	0.325

Notes: the sample has been modified to exclude user-days with a fire that is within 5 miles. This table contains results from estimation of Equation (1) for time spent at home on weekdays. Vacation days (as defined by the authors previously), including one day prior to vacating residence, have been excluded. The dependent variable is the percentage of time a user was home each day during the defined time windows. Number of half hour blocks the user was home divided by the observed number of half hour blocks times 100 gives the percentage of time a user was home. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01.

Table A3: Robustness Check Excluding Nearby Fire - Traveling Away from Home

	Dependent variable:				
	Vacated Residence				
	OLS	IV 1st stage		IV 2nd stage	
	(1)	(2)	(3)	(4)	(5)
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(PM_{j\tau} \geq 35)$	0.027*			0.179***	0.096***
	(0.015)			(0.048)	(0.034)
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(SmokeDay)$		0.167***			
		(0.010)			
I(Smoke = 1 day)			−0.011*		
			(0.006)		
I(Smoke = 2 days)			−0.080***		
			(0.011)		
I(Smoke = 3 days)			0.117***		
			(0.030)		
I(Smoke = 4 days)			0.308***		
			(0.039)		
I(Smoke = 5 days)			0.637***		
			(0.048)		
I(Smoke = 6 days)			1.411***		
			(0.098)		
I(Smoke = 7 days)			1.684***		
			(0.079)		
Temperature	0.032***	−0.041***	−0.039***	0.036***	0.034***
	(0.011)	(0.006)	(0.005)	(0.011)	(0.011)
$(\text{Temperature})^2$	−0.001***	0.001***	0.001***	−0.001***	−0.001***
	(0.0003)	(0.0001)	(0.0001)	(0.0003)	(0.0003)
Precipitation	0.004*	0.002***	0.002***	0.004*	0.004*
	(0.002)	(0.0004)	(0.0004)	(0.002)	(0.002)
$(\text{Precipitation})^2$	−0.0001**	−0.00003**	−0.00002*	−0.0001**	−0.0001**
	(0.00004)	(0.00001)	(0.00001)	(0.00004)	(0.00004)
Holiday	0.392***	−0.104***	−0.101***	0.413***	0.404***
	(0.047)	(0.007)	(0.007)	(0.048)	(0.048)
Mean	0.56	0.1	0.1	0.56	0.56
Week-of-sample FE	Yes	Yes	Yes	Yes	Yes
User by day-of-week FE	Yes	Yes	Yes	Yes	Yes
Observations	1,851,388	1,790,216	1,790,216	1,790,216	1,790,216
R^2	0.041	0.436	0.480	0.042	0.042

Notes: this table represents results from estimation of Equation (2). The dependent variable is an indicator for first day of leaving home. It takes a value 1 on the first day of the user vacating residence and 0 on all non-vacation days. The instrument for the number of days in the last 7 days when PM $\geq$ 35 μg/m³ in column (2) is the total number of smoke days in the last 7 days. In column (3), the instrument variables are a sum of indicators, each turning on if the number of smoke days in the last 7 days is equal to the indicator number. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01

Table A4: Robustness Check Using Alternative Fixed Effects - Time Spent at Home

	Dependent variable:					
	Day Time (8 a.m. - 5 p.m.)			*Evening Time (5 p.m. - 8 p.m.)*		
	OLS	IV 1st Stage	IV 2nd Stage	OLS	IV 1st Stage	IV 2nd Stage
	(1)	(2)	(3)	(4)	(5)	(6)
PM $\geq$ 35	2.436***		6.671***	0.284		−1.373
	(0.282)		(1.706)	(0.309)		(1.516)
Smoke Day		0.080***			0.081***	
		(0.008)			(0.009)	
Precipitation	0.074***	−0.002***	0.083***	0.014	−0.002***	0.010
	(0.009)	(0.0002)	(0.009)	(0.011)	(0.0002)	(0.010)
(Precipitation)^2	−0.001***	0.00003***	−0.001***	0.0001	0.00003***	0.0002
	(0.0002)	(0.00000)	(0.0002)	(0.0002)	(0.00000)	(0.0002)
Temperature	−0.343***	−0.004***	−0.324***	0.243***	−0.005***	0.235***
	(0.058)	(0.001)	(0.060)	(0.048)	(0.001)	(0.049)
(Temperature)^2	0.007***	0.0001**	0.007***	−0.005***	0.0001***	−0.005***
	(0.002)	(0.00003)	(0.002)	(0.001)	(0.00003)	(0.001)
Holiday	8.750***	0.022***	8.653***	−5.198***	0.022***	−5.159***
	(0.267)	(0.003)	(0.273)	(0.192)	(0.003)	(0.188)
Mean	56.5	0.01	56.5	71.43	0.01	71.43
User by day-of-week FE	Yes	Yes	Yes	Yes	Yes	Yes
Year-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,346,392	1,346,392	1,346,392	1,335,638	1,335,638	1,335,638
R^2	0.429	0.119	0.429	0.323	0.121	0.323

Notes: this table contains results from estimation of Equation (1) for time spent at home on weekdays. Vacation days (as defined by the authors previously), including one day prior to vacating residence, have been excluded. The dependent variable is the percentage of time a user was home each day during the defined time windows. Number of half hour blocks the user was home divided by the observed number of half hour blocks times 100 gives the percentage of time a user was home. Standard errors are clustered at the census place level.

Table A5: Robustness Check Using Alternative Fixed Effects - Traveling Away from Home

	Dependent variable: Vacated Residence				
	OLS	IV 1st stage		IV 2nd stage	
	(1)	(2)	(3)	(4)	(5)
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(PM_{j\tau} \geq 35)$	0.068***			0.192***	0.123***
	(0.012)			(0.035)	(0.029)
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(SmokeDay)$		0.163***			
		(0.014)			
I(Smoke = 1 day)			0.019***		
			(0.007)		
I(Smoke = 2 days)			0.014		
			(0.013)		
I(Smoke = 3 days)			0.120***		
			(0.032)		
I(Smoke = 4 days)			0.291***		
			(0.065)		
I(Smoke = 5 days)			0.636***		
			(0.055)		
I(Smoke = 6 days)			1.420***		
			(0.098)		
I(Smoke = 7 days)			1.620***		
			(0.084)		
Temperature	−0.017*	−0.027***	−0.025***	−0.012	−0.014
	(0.009)	(0.004)	(0.004)	(0.010)	(0.010)
$(\text{Temperature})^2$	0.0003	0.0004***	0.0004***	0.0002	0.0003
	(0.0002)	(0.0001)	(0.0001)	(0.0002)	(0.0002)
Precipitation	0.006***	0.001*	0.0005	0.006***	0.006***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
$(\text{Precipitation})^2$	−0.0001**	−0.00002*	−0.00001	−0.0001**	−0.0001**
	(0.00004)	(0.00001)	(0.00001)	(0.00004)	(0.00004)
Holiday	0.523***	0.061***	0.063***	0.512***	0.519***
	(0.047)	(0.007)	(0.008)	(0.048)	(0.048)
Mean	0.56	0.1	0.1	0.57	0.57
Year-month FE	Yes	Yes	Yes	Yes	Yes
User by day-of-week FE	Yes	Yes	Yes	Yes	Yes
Observations	1,905,765	1,885,720	1,885,720	1,885,720	1,885,720
R^2	0.040	0.259	0.303	0.040	0.040
F-stat	-	32.64	40.6	-	-

Notes: this table represents results from estimation of Equation (2). The dependent variable is an indicator for first day of leaving home. It takes a value 1 on the first day of the user vacating residence and 0 on all non-vacation days. The instrument for the number of days in the last 7 days when PM ¿= 35 µg/m³ in column (2) is the total number of smoke days in the last 7 days. In column (3), the instrument variables are a sum of indicators, each turning on if the number of smoke days in the last 7 days is equal to the indicator number. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01

Table A6: Robustness Check Using Alternative PM threshold - Time Spent at Home

	Dependent variable: Percentage Time at Home					
	Day Time (8 a.m. - 5 p.m.)			*Evening Time (5 p.m. - 8 p.m.)*		
	OLS	IV 1st Stage	IV 2nd Stage	OLS	IV 1st Stage	IV 2nd Stage
PM $\geq$ 12	0.079		4.176***	0.099		-0.438
	(0.083)		(1.392)	(0.086)		(1.144)
Smoke Day		0.116***			0.119***	
		(0.011)			(0.012)	
Precipitation	0.091***	-0.014***	0.147***	0.040***	-0.013***	0.033*
	(0.009)	(0.001)	(0.023)	(0.011)	(0.001)	(0.018)
$(\text{Precipitation})^2$	-0.001***	0.0002***	-0.002***	-0.0001	0.0002***	0.00002
	(0.0002)	(0.00002)	(0.0004)	(0.0002)	(0.00002)	(0.0003)
Temperature	0.063	-0.036***	0.206**	0.089	-0.034***	0.071
	(0.076)	(0.004)	(0.092)	(0.058)	(0.004)	(0.076)
$(\text{Temperature})^2$	-0.002	0.001***	-0.006**	-0.002	0.001***	-0.001
	(0.002)	(0.0001)	(0.003)	(0.002)	(0.0001)	(0.002)
Holiday	6.697***	-0.001	6.705***	-4.527***	-0.0005	-4.528***
	(0.236)	(0.008)	(0.232)	(0.191)	(0.008)	(0.192)
Mean	56.5	0.28	56.5	71.43	0.28	71.43
User by Day of Week FE	Yes	Yes	Yes	Yes	Yes	Yes
Week-of-sample FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,346,392	1,346,392	1,346,392	1,335,638	1,335,638	1,335,638
R^2	0.432	0.347	0.430	0.325	0.346	0.325
F-stat	-	49.09	-	-	48.61	-

Notes: This table contains results from estimation of Equation (1) for time spent at home on weekdays. Vacation days (as defined by the authors previously), including one day prior to vacating residence, have been excluded. The dependent variable is the percentage of time a user was home each day during the defined time windows. Number of half hour blocks the user was home divided by the observed number of half hour blocks times 100 gives the percentage of time a user was home. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01

Table A7: Robustness Check Using PM $\geq$ 12 Indicator - Travelling Away from Home

	Dependent variable: Vacated Residence				
	OLS	IV 1st stage		IV 2nd stage	
	(1)	(2)	(3)	(4)	(5)
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(PM_{j\tau} \geq 12)$	-0.003			0.106^{***}	0.058^{**}
	(0.004)			(0.030)	(0.023)
$\sum_{\tau=t-7}^{t-1} \mathbb{1}(SmokeDay)$		0.253^{***}			
		(0.014)			
I(Smoke = 1 day)			0.031		
			(0.029)		
I(Smoke = 2 days)			0.048		
			(0.060)		
I(Smoke = 3 days)			0.106		
			(0.080)		
I(Smoke = 4 days)			0.724^{***}		
			(0.116)		
I(Smoke = 5 days)			1.233^{***}		
			(0.082)		
I(Smoke = 6 days)			1.696^{***}		
			(0.087)		
I(Smoke = 7 days)			2.574^{***}		
			(0.142)		
Temperature	0.030^{***}	-0.262^{***}	-0.260^{***}	0.057^{***}	0.045^{***}
	(0.010)	(0.020)	(0.020)	(0.011)	(0.010)
$(\text{Temperature})^2$	-0.001^{***}	0.006^{***}	0.006^{***}	-0.001^{***}	-0.001^{***}
	(0.0002)	(0.001)	(0.001)	(0.0003)	(0.0003)
Precipitation	0.004^{*}	0.020^{***}	0.020^{***}	0.002	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$(\text{Precipitation})^2$	-0.0001^{**}	-0.0003^{***}	-0.0003^{***}	-0.0001	-0.0001^{*}
	(0.00004)	(0.00004)	(0.00004)	(0.00004)	(0.00004)
Holiday	0.386^{***}	-0.150^{***}	-0.146^{***}	0.404^{***}	0.396^{***}
	(0.046)	(0.022)	(0.023)	(0.046)	(0.046)
Mean	0.56	1.92	1.92	0.57	0.57
Week-of-sample FE	Yes	Yes	Yes	Yes	Yes
User by Day of week FE	Yes	Yes	Yes	Yes	Yes
Observations	1,905,765	1,885,720	1,885,720	1,885,720	1,885,720
R^2	0.041	0.584	0.590	0.041	0.041
F-stat		130.31	133.04		

Notes: This table represents results from estimation of Equation (2). The dependent variable is an indicator for first day of leaving home. It takes a value 1 on the first day of the user vacating residence and 0 on all non-vacation days. The instrument for the number of days in the last 7 days when PM $\geq$ 12 µg/m^3 in column (2) is the total number of smoke days in the last 7 days. In column (3), the instrument variables are a sum of indicators, each turning on if the number of smoke days in the last 7 days is equal to the indicator number. Standard errors are clustered at the census place level. *p<0.1; **p<0.05; ***p<0.01

A.3 Additional Figures

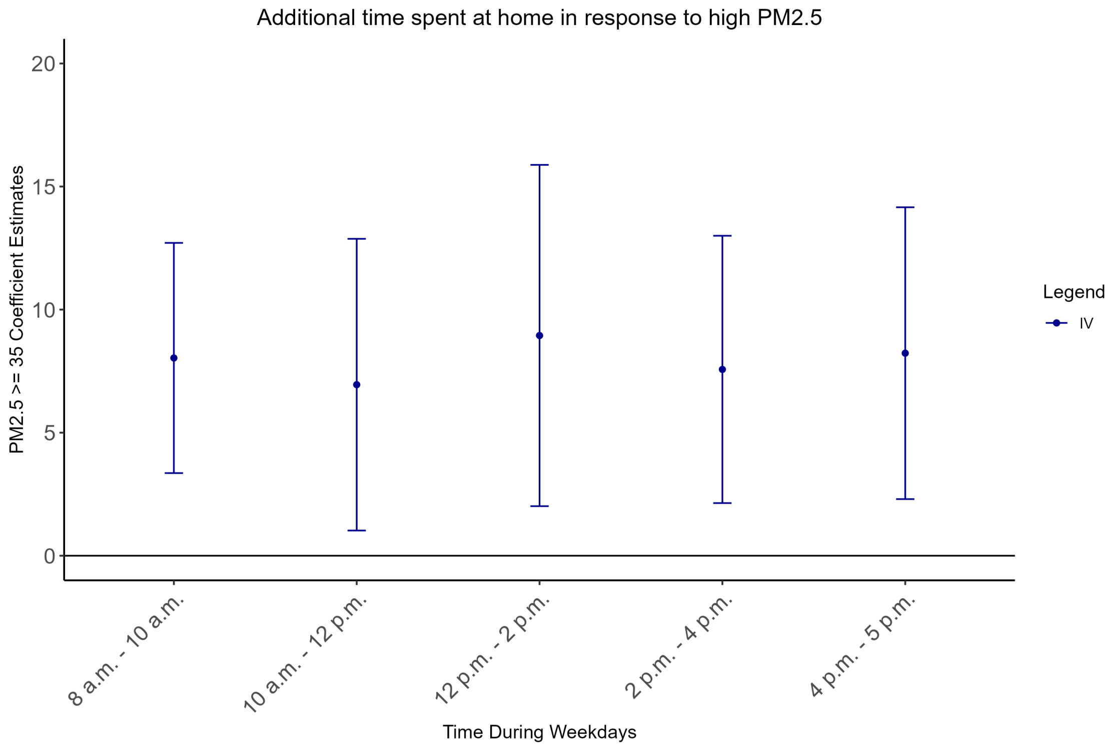


Figure A6: Changes in Time Spent at Home by Time of Day

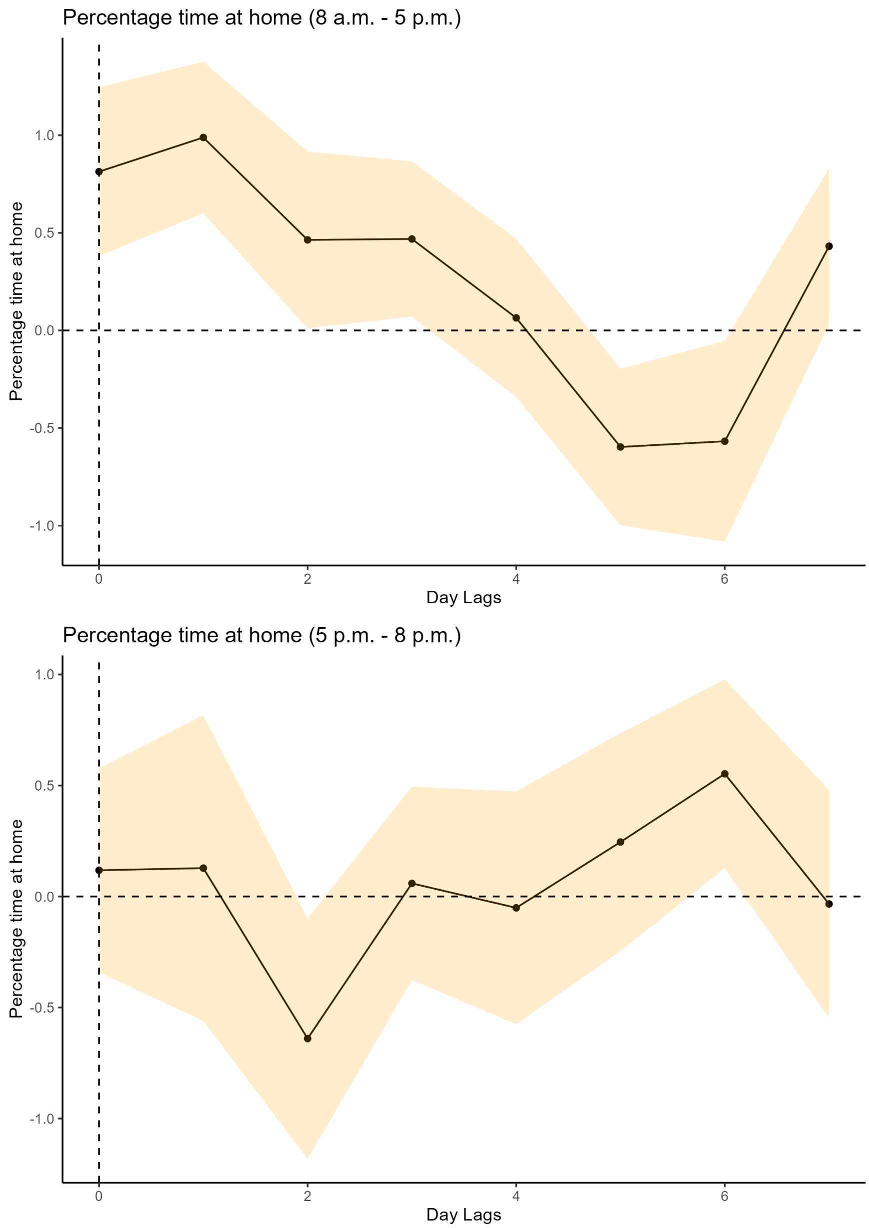


Figure A7: Time Spent at Home - Distributed Lag Model