

# The Gravity of Wildfire Smoke: Avoidance Costs Revealed by Mobility Data

Clara Berestycki\*  
*Job Market Paper*

[Click here for the most recent version](#)

October 26, 2026

## Abstract

This paper estimates the cost of wildfire smoke revealed by trip choices in Northern California in 2018. Using individual smartphone movement data and a discrete-choice model, I estimate the impact of wildfire smoke on activity participation and destination choice. I exploit the extensive choice set observed in the smartphone data to model trip flows as a gravity equation that includes smoke at the destination. People substitute away from smoky destinations and activities, and air quality alerts increase avoidance behavior. Outdoor trips are reduced the most, but trips to work and other destinations also decline with smoke. Based on estimated willingness to travel, these avoidance costs are about \$500 million, adding 24% to the mortality burden of the same wildfire smoke season.

---

\*Columbia University, School of International and Public Affairs, New York, NY, 10027. [csb2196@columbia.edu](mailto:csb2196@columbia.edu). I thank Keith Chen for access to anonymized smartphone data. Thanks to Doug Almond, Keith Chen, Jonathan Dingel, Hannah Farkas, Clément Herman, Jessie Lu, Nicolas Longuet-Marx, Robert Metcalfe, Suresh Naidu, Matthew Neidell, Maya Norman, Anna Papp, Jeffrey Shrader, and Stephan Thies for their helpful comments and suggestions. I also thank participants in the Camp Resources workshop, the AECCE workshop, Columbia Sustainable Development and Columbia Trade and Spatial colloquia, and Stanford Global Environment Policy student seminar for their input. I acknowledge research support from the Center for Political Economy at Columbia University. All errors are my own.

# 1 Introduction

In addition to causing billions of dollars in material damage, wildfires produce smoke plumes that can travel thousands of miles and generate extreme levels of air pollution (Borgschulte et al., 2024; Baylis et al., 2026). Exposure to fine particulate matter (PM<sub>2.5</sub>) from wildfire smoke causes thousands of deaths annually in the United States, a toll expected to rise as climate change increases the frequency and intensity of wildfires (Connolly et al., 2024; Miller et al., 2024; Qiu et al., 2025). By 2050, the health costs of wildfire smoke could even exceed the sum of all other estimated climate damages in the United States (Qiu et al., 2025). Simple behavioral adjustments, such as avoiding outdoor activities in smoky locations, can curtail exposure. This type of avoidance behavior is difficult to quantify and its costs remain largely unmeasured. The optimal level of investment in wildfire management policies, on which the U.S. federal government spent \$7.3 billion in 2025, depends on the damages these investments can avert, including avoidance costs (Congressional Research Service, 2025).

This paper develops a framework to measure the revealed preference cost of wildfire smoke. I estimate a discrete-choice model of daily activity and destination choice using smartphone movement data. The analysis focuses on Northern California in 2018, a densely populated area during one of the worst wildfire seasons on record. I find that people reallocate their activities toward staying in and substitute away from smoky destinations, and that the cost of smoke revealed by these choices amounts to about \$500 million for this single season.

I detect trips and activities in smartphone movement data, which consist of geocoded smartphone pings and trace individual mobility at a high spatio-temporal resolution. Based on visit frequency, I identify individual home and work locations for about 100,000 devices. Since outdoor activities induce the most direct exposure to wildfire smoke, I identify trips to outdoor locations by overlaying smartphone ping data with a GIS dataset of all outdoor protected areas in California, ranging from city to national parks. This approach identifies individual mobility patterns, distinguishing trips to home, work, and outdoor locations. Trips to destinations that do not fall in these categories, such as shopping, dining, or childcare, are recorded as “other”.

I use these mobility patterns to estimate a nested logit model of activity and destination choice where smoke can affect both what people do and where they go. In the upper tier, individuals pick among five activities several times a day: stay at their current location, go to work, visit an outdoor destination, go somewhere else, or return home. In the lower tier, individuals choose a destination for outdoor and “other” activities. I estimate the model on shares by aggregating individual trips to the ZIP code level. For each origin ZIP and period of the day, I compute the share of devices choosing each activity and, within each activity, aggregate trips into origin-destination (ZIP-to-ZIP)

flows.

The lower tier models origin-destination trip flows with a gravity equation, where the number of trips depends on smoke at the destination, travel cost, and baseline destination attractiveness. This approach draws on the gravity framework from the urban and trade literatures to extend the travel cost method, which uses the cost of reaching a site to value non-market amenities such as parks and beaches. Because I observe the distribution of trips across the full choice set, the model captures substitution across destinations. To illustrate, suppose someone decides to go hiking: if their usual spot is smoky, they may head to a cleaner area farther away. I include destination-by-period and origin-by-period fixed effects: the impact of smoke is identified from deviations from each destination’s usual number of visits, among travelers from the same origin in the same period. Since trips also respond to travel cost, these comparisons reveal people’s willingness to travel farther to avoid smoke, which I use to put a dollar value on avoidance.

I find that people both substitute away from smoky destinations and reallocate their activities toward staying in. Both margins start responding at moderate levels of wildfire-induced  $\text{PM}_{2.5}$ , around  $10\mu\text{g}/\text{m}^3$ , and strengthen with intensity. On the destination margin, trips to outdoor destinations fall by 14% above  $35\mu\text{g}/\text{m}^3$  and by 28% above  $75\mu\text{g}/\text{m}^3$ .<sup>1</sup> Trips to “other” destinations show little substitution, consistent with rerouting an indoor errand doing little to reduce exposure. On the activity margin, the effect of smoke is identified by comparing the activity mix of origins facing different smoke levels in the same period, holding fixed each origin’s usual mix. The share of devices choosing an outdoor activity falls by 23% on high smoke days (above  $75\mu\text{g}/\text{m}^3$ ). The response is not confined to discretionary activities: work also falls by 10%, and other activities by 8% on high smoke days.

For the 2018 fire season in Northern California, I estimate that the welfare cost of smoke revealed by trip choices is around \$500 million. This cost captures how much worse off smoke makes people based on their activity and destination choices. I use the estimated model parameters to recompute the utility of every alternative with smoke set to zero, and the difference between this counterfactual utility and the one on observed smoke days gives the compensating variation (Small and Rosen, 1981). The travel cost coefficient converts this loss into dollars: because the gravity equation measures how trips fall with the monetary cost of driving to a destination, the coefficient recovers the marginal utility of income.

More than half of the welfare loss accrues on days with smoke below  $35\mu\text{g}/\text{m}^3$ . These are mild to moderate smoke days, but they are frequent enough that their cumulative cost exceeds that of severe days. The pattern mirrors the health literature, where most of the mortality burden also

---

<sup>1</sup>Smoke above  $35\mu\text{g}/\text{m}^3$  corresponds to AQI above 100, where the air is deemed “unhealthy for sensitive groups” (U.S. Environmental Protection Agency, 2024).

falls on low and moderate exposure days rather than on the extremes (Miller et al., 2024; Heft-Neal et al., 2023). Through a similar mechanism, most of the welfare loss comes from cancelled work and “other” trips, which respond less strongly but represent a larger share of trips. Because each foregone trip is priced as a welfare loss, a trip merely postponed would inflate this figure. However, I find no evidence of intertemporal substitution: trips do not rebound in the weeks after smoke subsides.

To put the estimated avoidance costs into perspective, I estimate the mortality cost of the same wildfire smoke events. Using California death counts, I recover a dose-response of wildfire smoke on all-cause mortality. Applied to 2018 Northern California conditions and priced using the value of a statistical life year, this yields an acute mortality cost of \$2.1 billion. The mortality costs measure the health damages the smoke inflicts while avoidance costs reveal what individuals give up to reduce their exposure. Avoidance therefore constitutes an additional cost of wildfire smoke, representing about 24% of the mortality burden and presumably falling on a younger, more active population.

What drives avoidance behavior? A key constraint may be awareness: at low to moderate concentrations, smoke can go unnoticed, and these are the days that bear the brunt of the health cost (Miller et al., 2024; Heft-Neal et al., 2023). If awareness limits avoidance, providing information about wildfire smoke should increase it. In the final part of the paper, I study the effects of air quality alerts on avoidance behavior. Spare the Air alerts are issued in the Bay Area and Sacramento whenever forecast pollution crosses a fixed threshold. Because issuance is deterministic in the forecast, the setting supports a regression discontinuity design (Neidell, 2009; Cutter and Neidell, 2009). Days right above and below the threshold face nearly identical air quality but only the former receive an alert.

I find that alerts are salient, reducing outdoor trips by 6-10%. The alert should not change behavior for individuals who are already aware of the smoke and its dangers, so the discontinuity measures the behavior of those informed by the alert. I combine the RD estimate with the model’s smoke coefficients to recover how much an alert raises awareness relative to its baseline level. I then construct a counterfactual in which alerts are also issued on moderate smoke days below the threshold, assuming they raise awareness by the same proportion than at the threshold. Under this counterfactual, the incurred avoidance costs rise by around 20%. This represents a meaningful shift for an intervention that only changes information, and points to information provision as an important policy lever for mitigating the impacts of wildfire smoke.

This work contributes to several strands of the literature. First, it adds to work on adaptation to environmental risks. Much of the evidence on air pollution avoidance comes from specific

settings, such as zoo attendance (Neidell, 2009), movie theaters (He et al., 2022), and campground cancellations (Gellman et al., 2025). This paper builds on a growing literature that uses large-scale mobility data to show that air pollution changes time use and foot traffic (Berestycki and Chen, 2026; Sun, 2025). Whereas these studies estimate the effect of pollution separately on different activities, I embed these responses in a discrete-choice model where activities and destinations are chosen jointly. Using origin-destination trip flows rather than visit counts, I estimate substitution patterns and willingness to pay to avoid smoky destinations, which yields a revealed-preference estimate of wildfire smoke costs.

Second, this paper evaluates the effect of air quality alerts. Previous work on alerts has focused on specific sites or transportation modes, such as driving or cycling (Cutter and Neidell, 2009; Zivin and Neidell, 2009; Neidell, 2009; Saberian et al., 2017; Dangel and Goeschl, 2026). I use granular smartphone movement data to measure the effect of alerts on any outdoor activity, where exposure to smoke is most direct. My findings reinforce prior conclusions that alerts are salient and increase avoidance behavior, adding to mounting evidence that individuals respond to pollution information (Barwick et al., 2024; Ahmad et al., 2026). I also use the estimated effect of alerts to compute avoidance costs in a scenario with improved pollution awareness.

Third, this paper contributes to the travel cost literature on valuing non-market goods, connecting it to the gravity models used in trade and urban economics. Travel cost methods value recreation losses from environmental disasters (Hausman et al., 1995; English et al., 2018; Dundas and von Haefen, 2020; Gellman et al., 2025), but these approaches typically rely on surveys for a limited number of sites or activities. Trips taken somewhere outside the choice set or to another activity count as cancelled trips, which tends to overstate the damage.<sup>2</sup> Smartphone movement data lift this constraint by revealing trips across destinations and activities. Over this extensive choice set, the travel cost model I estimate is equivalent to a gravity equation, widely used in trade and urban economics.

Last, this paper relates to a growing literature in spatial economics that uses granular smartphone data to study consumption, commuting, and sorting decisions (Athey et al., 2021; Kreindler and Miyauchi, 2023; Miyauchi et al., 2025; Couture et al., 2025). In this literature, trip choices reflect a trade-off between travel costs and destination attributes. I adopt this framework by modeling wildfire smoke as a destination attribute, and this approach could be extended to other time-varying environmental hazards such as extreme heat.

The rest of the paper proceeds as follows. Section 2 presents the data, section 3 presents the

---

<sup>2</sup>For example, English et al. (2018) value shoreline recreation losses from the Deepwater Horizon oil spill with phone surveys covering 83 coastal sites, so trips to other areas fall outside their model. Gellman et al. (2025) study recreation losses due to wildfire smoke with campground cancellations, but the data does not include what campers do instead.

empirical model, section 4 presents results, section 5 the welfare computation, section 6 presents the health costs computation, and section 7 presents the air quality alerts analysis.

## 2 Data

I combine smartphone location data with daily smoke and weather data for California during the 2018 fire season. From anonymized individual smartphone pings I build an activity panel, which records what each device does in each period of the day, and a trip panel, which links each trip’s destination to its origin. Both are aggregated to the ZIP code level and matched to daily smoke  $\text{PM}_{2.5}$  from [Childs et al. \(2022\)](#).

### 2.1 Data Sources

**Smartphone Mobility Data.** I use individual Veraset smartphone ping data to measure mobility. This data comes from device geocoordinates recorded by phone operating systems. Depending on user settings, when open or in background mode, this information can be shared with apps. In 2018, the ability of users to control location sharing with apps was more limited than today. This data was sold to third parties and aggregated, resulting in an anonymized set of pings by device, with geocoordinates, timestamps, and GPS margins of error of usually a few meters.

**Wildfire-induced  $\text{PM}_{2.5}$ .** I use the [Childs et al. \(2022\)](#) dataset to measure daily  $\text{PM}_{2.5}$  produced by wildfires on a 10x10 kilometer grid. This dataset uses machine learning to predict  $\text{PM}_{2.5}$  concentrations from multiple sources, including ground-truth measurements from EPA monitors as well as remotely sensed and reanalysis products. The model also includes covariates such as pollution particle trajectories from the HYSPLIT model, meteorology, distance from fire, aerosol measurements, predicted aerosol optical depth, and cross-sectional variables like elevation.

Table [A1](#) reports summary statistics for wildfire-induced  $\text{PM}_{2.5}$  from the [Childs et al. \(2022\)](#) dataset from July to November 2018 in Northern California. The distribution is highly skewed. The median ZIP-day has no wildfire smoke at all, while the mean is around  $5 \mu\text{g}/\text{m}^3$  and the maximum  $294 \mu\text{g}/\text{m}^3$ . For scale, the current EPA annual standard for total  $\text{PM}_{2.5}$  from all sources is  $9 \mu\text{g}/\text{m}^3$ : smoke alone averaged more than half that over these five months ([U.S. Environmental Protection Agency, 2024](#)). Summer wildfire smoke pollution was mainly due to the Carr, Mendocino Complex, and Donnell fires which burned simultaneously at the end of July and early August (see Figure [A1](#)). The largest wildfire pollution event was caused by the November Camp Fire, which was the deadliest fire in California history and blanketed large areas of the state with smoke (see Figure [A2](#)). The two episodes give the sample complementary variation, prolonged moderate smoke from multiple fires over the summer and an extreme, sustained spike in the fall.

**Temperature and Precipitation.** Daily rainfall and daily minimum and maximum temperature come from the gridded PRISM weather product. This version uses a fixed number of monitoring stations and is obtained at the daily county level (Schlenker, 2024).

**Travel Cost.** Following Gellman et al. (2025), I construct a dollar-denominated, time-invariant travel cost combining travel time valued at a fraction of the hourly wage with a per-mile vehicle cost for each origin-destination ( $od$ ) pair in the model:

$$\text{travel cost}_{od} = 2 \times \left( \underbrace{\nu w_o \cdot \frac{\text{time}_{od}}{60}}_{\text{time cost}} + \underbrace{\rho \cdot \text{dist}_{od}}_{\text{vehicle cost}} \right) \quad (1)$$

where  $\text{time}_{od}$  is the ZIP-to-ZIP time-invariant driving time from Hu et al. (2020). The hourly value of travel time is a fraction  $\nu = \frac{1}{3}$  of the origin’s hourly wage  $w_o$  (U.S. Bureau of Labor Statistics, 2018).  $\nu = \frac{1}{3}$  reflects the convention that the opportunity cost of leisure travel time is typically well below the full wage (English et al., 2018; Gellman et al., 2025; Cesario, 1976). The per-mile vehicle cost  $\rho$  captures the marginal operating cost of driving: fuel, per-mile maintenance, tires, and marginal depreciation. I set  $\rho = \$0.30$  per mile (American Automobile Association, 2018; English et al., 2018; Gellman et al., 2025).<sup>3</sup> The factor of 2 converts the one-way origin–destination trip into a round trip.

**Outdoor Spaces.** I use the California Protected Areas Database (CPAD) and the California Conservation Easement Database (CCED) (GreenInfo Network, 2025) to define outdoor spaces suitable for outdoor activities. These GIS datasets collect boundaries from more than 1,000 public and non-profit organizations to map protected outdoor areas in California, from national parks to beaches, city parks, and conservation easements (see Figure A3).

## 2.2 Construction of Trips and Activities From the Mobility Data

**Sample Definition and Representativeness.** I restrict the sample to an area of Northern California containing the Bay Area and Sacramento that was the most affected by the 2018 smoke events (see Figure A4 for the delimitation). This records 268,975 distinct devices. So as not to confound the effect of smoke with the fire itself, I exclude 97 ZIPs (5.5% of the total) that had an active fire burning in them during the time period (Figure A5). I identify fire perimeters using the FRAP database (CAL FIRE, 2024). The baseline sample further excludes device-days on which the device travels more than 240 miles from home. In this exercise, I restrict the sample to device with an identified workplace in California. This restricts the sample to 98,054 devices in Northern California (see Table A2).

---

<sup>3</sup>Fuel cost uses the 2018 average price of regular-grade gasoline (\$2.52 per gallon). Maintenance and marginal depreciation use the per-mile figures in American Automobile Association (2018) as in English et al. (2018).

Coverage of the mobility data does not vary systematically with smoke. Devices ping on average 100 times per day and across about seven hours, and these figures are stable across smoke levels: from clean days to days above  $75\mu\text{g}/\text{m}^3$  the mean number of pings varies by less than 15% and the number of hours with at least one ping by less than 7% (Table A3). Measured changes in activity therefore reflect behavior rather than changes in how devices are observed.

**From Pings to Locations.** I detect home and work location and check outdoor visits for each device. A device is considered outdoors if it pings within the bounds of a protected area in the CPAD and CCED databases (GreenInfo Network, 2025). The home location is where the device pings most frequently during night hours (Berestycki and Chen, 2026; Atkin et al., 2022; Athey et al., 2021). The location where the device accumulates the most work-hour time is designated as the workplace, unless that location is the device’s home.<sup>4</sup> With this procedure, I am able to assign a workplace to 36% of devices in the sample (see Table A2). This procedure identifies workplaces for office-schedule workers and leaves out workers with irregular schedules or sites.

I construct two separate objects to estimate the model: the activity and trip panels, which measure different objects (see Table A7 and Table A8 for descriptive statistics on both panels). The activity panel only records whether a device engaged in an activity during a given period of the day. The trip panel additionally requires observing where the trip started and ended, which leads to some attrition. The first tier of the model estimates participation and uses the device counts, while the second tier uses origin-destination trip flow to price destination choice.

Individual activities and trips are spatially aggregated to the ZIP level, which defines both the origins and destinations in the choice model. Finer units would make origin-destination flows sparse and would add little variation in smoke exposure, which is originally measured on a 10km grid (Childs et al., 2022). Aggregating to the ZIP level gives all individuals in the same ZIP a common choice set with identical distances, smoke exposure, and destination characteristics. Estimating on ZIP-level counts and flows makes the model much more computationally efficient.

**The Trips Panel.** I detect individual trips from the pings and aggregate them into a dataset of daily-period ZIP-to-ZIP flows. A trip occurs whenever two consecutive pings of a device fall in different ZIP codes. Because a person can cross several ZIP codes before reaching a destination, in particular in urban areas, I collapse consecutive ZIP crossings into a single trip chain if they are less than one hour apart. Figure A12 illustrates the procedure.

---

<sup>4</sup>A location is a grid cell of roughly 150 meters. Work hours are 8am to 6pm. To qualify, a location must be visited for at least 30 minutes of work-hour time on at least two distinct days in September. A device whose top location is its home is still assigned a workplace if it spends at least four hours a day, on at least three days, at another location.

I then classify each trip by its destination. A trip is classified a work trip if its destination is the work location of the device, and a trip is an outdoor trip if its destination pings fall within the boundaries of a protected area in the CPAD and CCED databases. Finally, I identify the origin ZIP of each trip with a backward search: I consider where the device spent time in the four hours prior to reaching the destination and select the most recent ZIP where the device spent more than 45 minutes. Figure A13 illustrates this procedure for an outdoor trip.

The resulting panel contains 1.69 million identified trips, of which 141,664 (8.4%) are to outdoor destinations (Table A7). The average trip covers 23 kilometers and 38 minutes of driving (Table A7). Outdoor trips are sparse at the daily level: the median origin ZIP records no outdoor trip on a given day, and 59% of ZIP-days have none, which motivates the Poisson estimator for the gravity equations. Note that this panel does not contain same-ZIP trips, because the gravity regression requires travel cost as a cost shifter, which is only available for trips to different ZIPs.

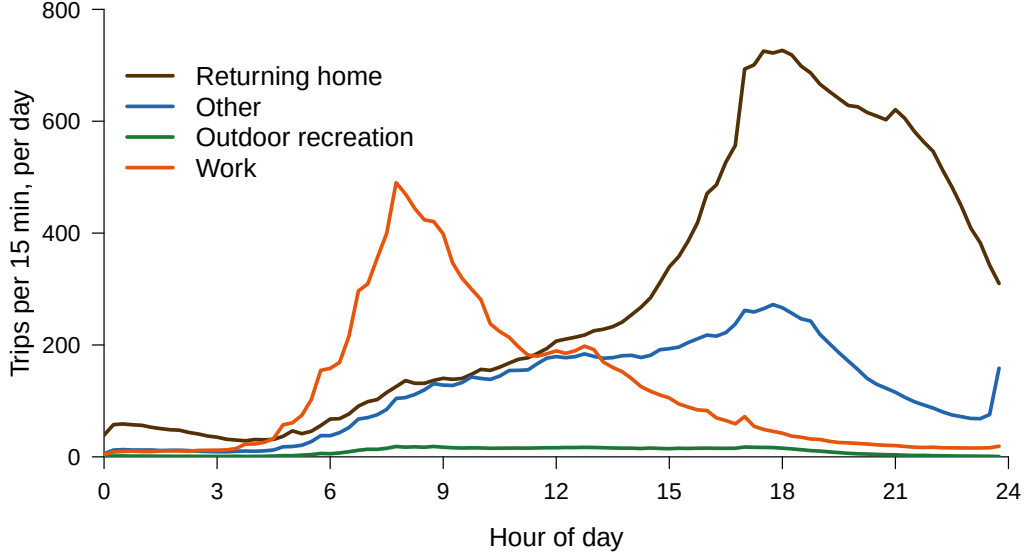
**The Activity Panel.** The first tier of the choice model requires a measure of daily participation: how many devices in each ZIP stay home, go to work, visit an outdoor destination, return home, or travel elsewhere. I classify each device-day-period into one of these five alternatives based on the locations the device visits, and aggregate to counts by home ZIP, day, and period of the day. This yields a panel of more than 1,900 ZIPs over 153 days from July 1 to November 30, 2018. On an average day, 7.9% of devices stay home, 3.3% visit an outdoor destination, 49.6% go to work, and 39.1% travel to other destinations (Table A8).

The type of activity is mutually exclusive within a time period. When a device qualifies for several activities during the same period, classification follows the priority order work, outdoor, other, home. Other trips can correspond to home production (childcare, eldercare), shopping, or leisure such as dining out. These ZIP-day-period counts are the estimation data for the first tier of the choice model.

## 2.3 Stylized Facts

**#1: The activity classification reproduces daily commuting patterns.** Devices follow a daily pendulum, going to work in the morning and coming home at night, and trip departures cluster around these commutes, as shown in Figure 1. Devices take about 2.3 trips per day and more than 96% of devices make at most 4 daily trips, so four occasions per day capture nearly all trips (see Figure A6). This rhythm motivates the structure of the model, which divides the day into four periods based on the density of trip departures: morning (12:00AM–11:00AM), midday

Figure 1: Distribution of trips during the day



(11:00AM–4:00PM), evening (4:00PM–7:00PM), and night (7:00PM–12:00AM).<sup>5</sup> The classification also captures the weekly rhythm: on weekends, the share of periods spent at work falls while outdoor activities more than double (Table A6).

I benchmark these patterns against the 2017 National Household Travel Survey (NHTS) (U.S. Department of Transportation, Federal Highway Administration, 2017). The classification closely matches the NHTS on work trips: on weekday mornings, 54.9% of devices go to work against 59.6% of NHTS respondents, and the share returning home matches almost exactly in the evening (10.9% versus 11.1%) (see Table A5). My activity classification assigns more periods to other activities and fewer to outdoor recreation and staying home, mostly because of how activities are defined. My definition of outdoor recreation only includes park visits, whereas the NHTS category also covers activities such as a walk around the neighborhood.<sup>6</sup> Since the model includes activity-specific constants, these level differences do not affect the estimated responses to smoke, which come from variation in smoke across days and ZIPs.

**# 2: Trips decay with distance.** Trips between ZIP codes fall sharply with distance (Figure A7). The trip flows detected in the smartphone data thus reproduce the familiar gravity pattern of origin-destination flows that is central to the urban and trade literature. This pattern motivates the second tier of the model, which takes the form of a gravity equation to price smoke

<sup>5</sup>The number of periods balances two errors. With too few periods, the model undercounts the true number of trips of a device, but with too many, most periods hold no trips, and the model treats that absence as a choice.

<sup>6</sup>The two samples also cover different periods: July–November 2018 for the mobility data, against a full year in 2016–17 for the NHTS.

avoidance. In my data, a 10% increase in distance reduces outdoor trips by 15% and other trips by about 25%, reflecting that people travel farther to reach parks than for routine activities such as dining.

**#3: Smoke sharply reduces activities.** I first show that the data capture smoke avoidance: when exposed to wildfire smoke, people sharply cut their outdoor trips, and also decrease trips to work and other destinations more modestly. Building on earlier mobility-data evidence that people reduce their activities on polluted days (Berestycki and Chen, 2026; Sun, 2025), I trace adaptation week by week around the onset of a wildfire smoke spell. A ZIP code enters a smoke spell in the first week with at least two days above  $35 \mu\text{g}/\text{m}^3$ , the EPA daily standard, after a week of clean air (below  $10 \mu\text{g}/\text{m}^3$ ). ZIP codes whose smoke spell starts in the same week paired with control (smoke-free) ZIP codes over the same window form a stack. Controls mostly come from Southern California ZIP codes, which stayed clean during these episodes (Figure A9). I estimate a stacked event-study regression (Cengiz et al., 2019):

$$y_{d\tau e} = \exp\left(\beta_\tau D_{de} + \alpha_{de} + \lambda_{\tau e} + X'_{d\tau e}\gamma\right) \varepsilon_{d\tau e} \quad (2)$$

where  $\tau \in \{-3, \dots, 4\}$  indexes the eight weeks around the onset of the smoke spell in stack  $e$ , with  $\beta_{-1} = 0$ .  $y_{d\tau e}$  is the number of outdoor trips to ZIP code  $d$  per device in week  $\tau$ <sup>7</sup> and  $D_{de}$  indicates smoky ZIP codes in stack  $e$ . ZIP-by-stack fixed effects  $\alpha_{de}$  and week-by-stack fixed effects  $\lambda_{\tau e}$  ensure that each smoky ZIP code is compared with clean ZIP codes in the same calendar week, and  $X_{d\tau e}$  contains weekly temperature and precipitation averages.

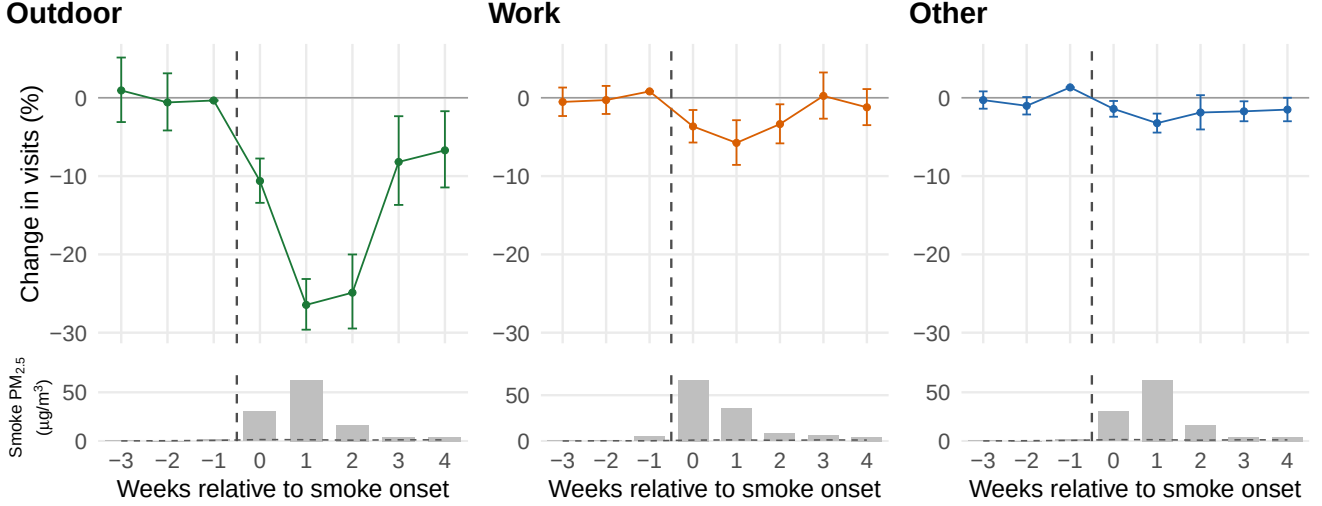
Figure 2 reports the jointly-estimated coefficients  $\beta_\tau$  relative to the pre-period mean. Before onset, smoky and clean ZIP codes follow the same path. Smoke then sharply reduces outdoor trips, by more than 25%, while work trips fall by 5–7% and other trips by 2–3%. The drop lasts as long as the smoke spell, about three weeks, and activities bounce back once it clears.

The event study shows that people avoid smoke, one activity at a time. The model brings these responses together in a single decision over activities and destinations, which separates smoke at the origin from smoke at the destination and makes it possible to value avoidance in dollars.

---

<sup>7</sup>I estimate equation (2) by Poisson pseudo-maximum likelihood with the log number of devices in the destination ZIP code as an offset. This amounts to modeling trips per device, weighting each ZIP code by its number of devices, while keeping weeks with zero trips.

Figure 2: Weekly coefficients from the event study of trips around smoke event onset



Notes: Stacked event-study estimates of the effect of smoke onset on activities, following [Cengiz et al. \(2019\)](#). The unit is the ZIP by week (Monday–Friday excluding holidays for work trips). A ZIP has an onset in week  $t$  if smoke  $\text{PM}_{2.5}$  exceeds  $35\mu\text{g}/\text{m}^3$  on at least two days of  $t$  after averaging below  $10\mu\text{g}/\text{m}^3$  in each of the three preceding weeks (700 events for outdoor and other trips, 413 for work trips, across 8 onset weeks). Each event dataset spans three weeks before to four weeks after onset and pairs treated ZIPs with control ZIPs whose weekly smoke stays below  $10\mu\text{g}/\text{m}^3$  throughout. Estimation is PPML on device-periods spent on the activity, with all device-periods as offset, ZIP-by-event

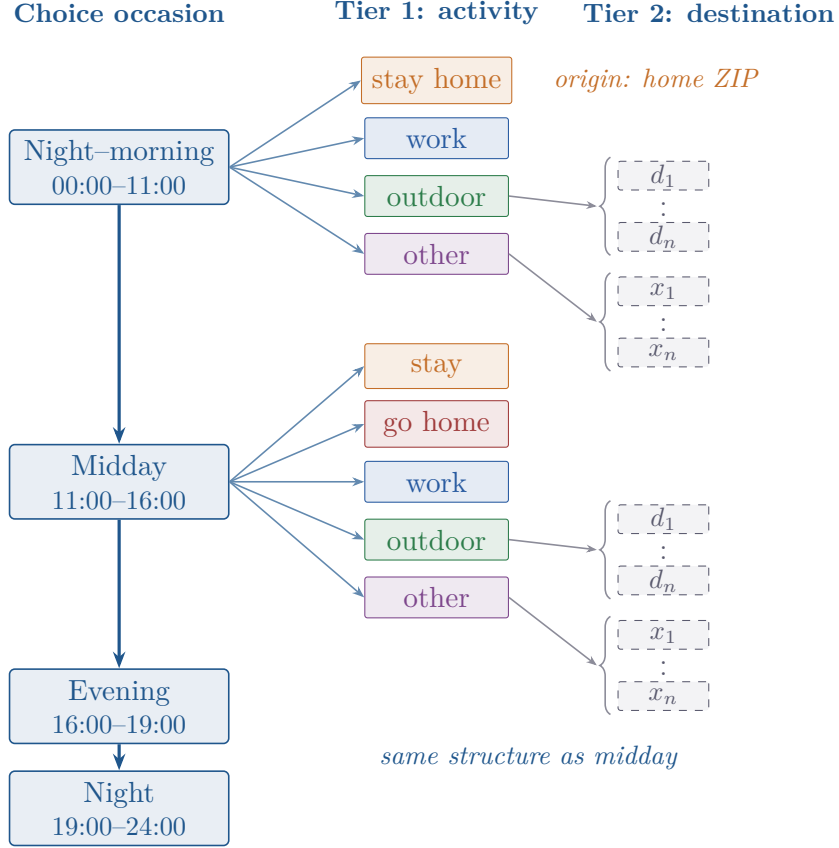
### 3 Empirical model

This section introduces a model of individuals’ time use decision and estimates how smoke influences this choice, both for smoke at their origin and for smoke at the destination. I write a nested-logit model of individuals’ decisions on how to allocate their time, which I estimate sequentially.<sup>8</sup> Individuals are indexed by  $i$ , origin ZIPs by  $o$ , destination ZIPs by  $d$ ,  $t$  denotes the day, and  $p$  denotes the period of the day. In the baseline model, individuals choose an activity–destination pair four times a day, with  $p \in \{\text{morning (0–11AM), midday (11AM–4PM), evening (4PM–7PM), night (7PM–12AM)}\}$ . Choices in each period are considered independent. Figure 3 represents the choice structure.

In each period  $p$  on day  $t$ , the individual chooses an activity  $a \in A_p$  and, for outdoor and other activities, a destination  $d$ . The choice set contains five alternatives,  $A_p = \{\text{stay, home, work, outdoor, other}\}$ : staying at the current location, going home, going to work, going to an outdoor destination, or going to another destination. Since every individual starts the day at home, going home is not available in the morning, and  $A_{\text{morning}} = \{\text{stay, work, outdoor, other}\}$ , where staying means staying

<sup>8</sup>The model estimates are consistent under sequential estimation ([Train, 2009](#)).

Figure 3: Choice structure across the four periods



home. The individual chooses  $(a, d)$  to maximize

$$U_{iadtp} = \sum_{b \geq 1} \gamma_a^b \mathbb{1}[\text{smoke}_{ot} \in b] + \mu V_{odtp}^a + \varepsilon_{iadtp} \quad (3)$$

where the first term captures smoke at the origin, and  $V_{odtp}^a$ , defined below, captures the attractiveness of destination  $d$  for activity  $a$ , including smoke at the destination.  $V_{odtp}^a$  is set to zero for staying, going home and work, which have no destination choice. Each activity–destination pair has a single error term  $\varepsilon_{iadtp}$ . These errors follow a generalized extreme value distribution in which errors are correlated across destinations within the same activity (McFadden, 1978; Train, 2009). Under this distribution, the probability of choosing a pair  $(a, d)$  factors into the probability of choosing destination  $d$  given activity  $a$  (lower tier) and the probability of choosing activity  $a$  (upper tier), in which destinations enter only through the inclusive value with coefficient  $\mu$ .

**Destination choice.** Conditional on choosing an outdoor or other activity, an individual selects a destination  $d$  from the activity-specific set  $D_a$ . The attractiveness of destination  $d$  for an individual living in origin  $o$  in period  $p$  of day  $t$  is summarized by the deterministic utility  $V_{odtp}^a$ . Individuals

living in the same origin ZIP have the same  $V_{odtp}^a$ , since all individuals share the same preferences and covariates are measured at the ZIP level.  $V_{odtp}^a$  depends on the travel cost of reaching  $d$  from  $o$ , on smoke and weather  $X_{dt}$  at the destination on day  $t$ , and on destination-by-period fixed effects  $\delta_{dp}^a$  capturing the time-invariant attractiveness of each destination in each period of the day:

$$V_{odtp}^a = \theta_a \log(\text{TravelCost}_{od}) + \sum_{b \geq 1} \beta_b^a \mathbb{1}[\text{smoke}_{dt} \in b] + X_{dt} \Gamma^a + \delta_{dp}^a. \quad (4)$$

The travel cost is the dollar cost of the round trip from  $o$  to  $d$ , combining the value of travel time with vehicle operating costs (its construction is detailed in Section 2.1). Coefficient  $\theta_a$  represents the disutility of spending an additional dollar on travel cost. Since travel cost is in dollars,  $\theta_a$  recovers the marginal utility of income, which I use to translate utility losses into dollar amounts in the welfare computation.  $b$  indexes five bins of smoke PM<sub>2.5</sub> concentrations ( $\mu\text{g}/\text{m}^3$ ):  $< 1$ ,  $1-10$ ,  $10-35$ ,  $35-75$ , and  $75+$ . The bin coefficients  $\beta_b^a$  capture the effect of destination smoke in bin  $b$  on the utility of choosing destination  $d$  for activity  $a$ , relative to the smoke-free reference bin ( $b = 0$ , smoke PM<sub>2.5</sub>  $< 1$ ). The controls  $X_{dt}$  are destination-by-day weather (temperature and precipitation), and  $\delta_{dp}^a$  is a destination-by-period fixed effect capturing time-invariant features of  $d$ , such as baseline attractiveness or amenity value, for each activity  $a$ .

The destination choice has the structure of a gravity model. It relates the probability of an  $o \rightarrow d$  trip to an origin term, a destination term and a bilateral travel cost, which links it to the trade and urban literature on gravity models of flows (Ahlfeldt et al., 2015; Couture and Handbury, 2020; Rossi-Hansberg, 2010). It can equivalently be read as the demand equation of a travel cost model. This method is widely used in environmental economics to value non-market amenities such as parks or beaches by relating the number of visits to the cost of reaching a site (English et al., 2018; McConnell, 1985; Gellman et al., 2025). My approach extends the travel cost logic to a setting with many destinations where the full choice set is observed. This makes it possible to measure substitution across destinations, a margin most travel cost studies cannot recover. Unlike the fixed site amenities valued in standard travel cost studies, smoke also varies from day to day at each destination.

Conditional on choosing activity  $a \in \{\text{outdoor}, \text{other}\}$ , the probability that an individual at origin  $o$  selects destination  $d$  in period  $p$  of day  $t$  is

$$P_{a|otp} = \frac{\exp\left(\sum_{b \geq 1} \gamma_a^b \mathbb{1}[\text{smoke}_{ot} \in b] + \mu \mathcal{V}_{otp}^a\right)}{\sum_{a'} \exp\left(\sum_{b \geq 1} \gamma_{a'}^b \mathbb{1}[\text{smoke}_{ot} \in b] + \mu \mathcal{V}_{otp}^{a'}\right)}. \quad (5)$$

Conditional on picking activity  $a$ , each individual chooses the destination with the highest utility. The expected utility of this best destination, averaged over taste shocks, takes a closed form, up to

an additive constant (McFadden, 1978; Train, 2009):

$$\mathcal{V}_{otp}^a = \log \sum_{d' \in D_a} \exp(V_{od'tp}^a).$$

The inclusive value  $\mathcal{V}_{otp}^a$  summarizes the expected utility from the destination choice, conditional on choosing activity  $a \in \{\text{outdoor, other}\}$ , and enters the upper-tier activity choice as a regressor. Since staying, going home and work have no destination choice,  $\mathcal{V}_{otp}^{\text{stay}}$ ,  $\mathcal{V}_{otp}^{\text{home}}$  and  $\mathcal{V}_{otp}^{\text{work}}$  are equal to zero: staying is the reference alternative, home is a single location, and each device’s workplace is fixed.

**Activity choice.** The upper tier models activity choice based on smoke at home, at origin  $o$ , and on the inclusive values of the outdoor and other activities. In equation (3), origin smoke enters through the parameters  $\gamma_a^b$ , which capture the effect of origin smoke in bin  $b$  on the utility of choosing activity  $a$ , relative to the smoke-free bin  $b = 0$ , using the same bins as for destination smoke. The inclusive value enters with coefficient  $\mu$ , the dissimilarity parameter, which measures how substitutable destinations are within an activity. A small value of  $\mu$  (close to zero) indicates that destinations within each nest are close substitutes, whereas a value close to 1 indicates that the model is close to a multinomial logit, in which destinations within a nest are no more substitutable than destinations across nests. Random-utility consistency requires  $\mu \in (0, 1]$ . Following standard practice, I estimate a single dissimilarity parameter  $\mu$  across nests. The  $\gamma_a^b$  parameters capture the effect of origin smoke in bin  $b$  on the utility of choosing activity  $a$ , relative to the smoke-free bin  $b = 0$ , using the same bins as for the destination choice. The probability of choosing activity  $a$  is

$$P_{a|otp} = \frac{\exp\left(\sum_{b \geq 1} \gamma_a^b \mathbb{1}[\text{smoke}_{otp} \in b] + \mu \mathcal{V}_{otp}^a\right)}{\sum_{a'} \exp\left(\sum_{b \geq 1} \gamma_{a'}^b \mathbb{1}[\text{smoke}_{otp} \in b] + \mu \mathcal{V}_{otp}^{a'}\right)}. \quad (6)$$

**Joint probability.** The probability that an individual at origin  $o$  chooses the activity–destination pair  $(a, d)$  in period  $p$  of day  $t$  is the product of the two tiers:

$$P_{da|otp} = P_{d|aotp} \cdot P_{a|otp}.$$

**Perceived smoke.** So far, the model has taken individuals’ information about smoke as given. Individuals, however, respond to smoke as they perceive it. This subsection makes the distinction between perceived and realized smoke explicit and derives the consequences for estimation and welfare. Let  $I_{ot}$  denote the information available at origin  $o$  on day  $t$ : visibility, news coverage,

forecasts, and alerts.<sup>9</sup> Individuals do not observe smoke directly, but form beliefs about it from the available information  $I_{ot}$ . An individual's belief that smoke falls in bin  $b$  is the conditional probability

$$\pi_{ot}^b = \Pr(\text{smoke}_{ot} \in b \mid I_{ot})$$

with  $\sum_{b \geq 0} \pi_{ot}^b = 1$ , and  $\pi_{dt}^b$  is defined analogously for smoke at destination  $d$ , conditional on the information at the origin.

Individuals maximize expected utility given the available information. Assuming smoke is imperfectly observed, individuals evaluate utility in expectation. Each indicator  $\mathbb{1}[\text{smoke}_{ot} \in b]$  from Equation (4) is replaced by the conditional probability, the belief  $\pi_{ot}^b$ . Because utility is linear in the bin indicators, this substitution is exact. In the destination tier, (4) becomes

$$V_{odtp}^a = \theta_a \log(\text{TravelCost}_{od}) + \sum_{b \geq 1} \beta_a^b \pi_{dt}^b + X_{dt} \Gamma_a + \delta_{dp}$$

and the origin smoke terms of the activity tier become  $\sum_{b \geq 1} \gamma_a^b \pi_{ot}^b$ . Note that if individuals have full information, this is equivalent to the baseline specification.

Given the information  $I_{ot}$ , beliefs are common to all individuals at origin  $o$  on day  $t$ . Perception therefore enters deterministic utility and does not interact with the error term.<sup>10</sup> Because the misperception does not affect the error term, the choice probabilities (5) and (6), the inclusive values, and the welfare formulas keep their closed forms, with perceived smoke now entering utility through  $\pi$ . In this setup, the estimated coefficients do not recover the response to realized smoke, but a blend of true responses and beliefs. The regression sorts observations by realized smoke, since beliefs are not observed. But behavior under each realized level follows the beliefs held at the time: on some bin  $b$  days the smoke is noticed, on others it is not. The coefficient on bin  $b$  therefore averages the belief-driven responses that prevail when smoke is truly at level  $b$ . Formally, the coefficient on the bin- $b$  indicator is the difference

$$\hat{\beta}_a^b = \mathbb{E} \left[ \sum_{b' \geq 1} \beta_a^{b'} \pi_{dt}^{b'} \mid \text{smoke}_{dt} \in b \right] - \mathbb{E} \left[ \sum_{b' \geq 1} \beta_a^{b'} \pi_{dt}^{b'} \mid \text{smoke}_{dt} \in 0 \right]$$

---

<sup>9</sup>This does not depend on the period, because the smoke variable is available at the daily level, and does not vary across periods.

<sup>10</sup>Equivalently, with realized smoke as the covariate, the error is the sum of a taste shock and a perception error,  $\sum_b \beta_a^b (\pi_{dt}^b - \mathbb{1}[\text{smoke}_{dt} \in b])$ . However, since beliefs are common within an origin-day, the perception error is a deterministic shifter of utility, not an idiosyncratic draw. The misperception can therefore be written as a covariate of  $V$  rather than as a component of the idiosyncratic draw

which by linearity equals

$$\hat{\beta}_a^b = \sum_{b' \geq 1} \beta_a^{b'} \left( \mathbb{E}[\pi_{dt}^{b'} \mid \text{smoke}_{dt} \in b] - \mathbb{E}[\pi_{dt}^{b'} \mid \text{smoke}_{dt} \in 0] \right).$$

The subtracted terms are mean beliefs in smoke on days without smoke. I consider that smoke is rarely perceived when absent and that these terms are negligible so that

$$\hat{\beta}_a^b \approx \sum_{b' \geq 1} \beta_a^{b'} \mathbb{E}[\pi_{dt}^{b'} \mid \text{smoke}_{dt} \in b].$$

If beliefs are correct, all weight lies on  $b' = b$  and  $\hat{\beta}_a^b = \beta_a^b$ . If the smoke goes unnoticed, weight shifts to the clean bin, which carries no coefficient, and the estimate is attenuated toward zero. The same holds for  $\hat{\gamma}_a$  with origin beliefs. The model therefore estimates decision utility: choices reveal responses to conditions as perceived, and the welfare measure of Section 3 prices the avoidance individuals actually undertook given what they perceived.

The relation between estimated and perfect-information welfare depends on the direction of the misperception. If smoke goes under-noticed, the estimated coefficients are attenuated and the welfare measure understates its perfect-information counterpart. The direction of the misperception cannot be detected from choice data alone, because observed behavior is the product of beliefs and preferences. To separate them, I use air quality alerts, which provide variation in information without changing air quality. If individuals are perfectly informed, an alert brings no additional information, and in the model outdoor trips are continuous through the alert threshold. However, I find that alerts reduce outdoor trips, suggesting that beliefs about smoke are below the true smoke level and that the full-information welfare consequence is bounded from below by the estimated welfare measure.

**Estimation of the destination choice.** The destination choice probability  $P_{d|aotp}$  maps directly onto the observed trip counts. The data record  $N_{odtp}^a$ , the number of  $o \rightarrow d$  trips for activity  $a$  in period  $p$  of day  $t$ . Each of the  $N_{otp}^a$  individuals at origin  $o$  who choose activity  $a$  picks destination  $d$  with probability  $P_{d|aotp}$  from equation (5). The expected number of  $o \rightarrow d$  trips is therefore

$$\mathbb{E}[N_{odtp}^a] = N_{otp}^a \frac{\exp(V_{odtp}^a)}{\sum_{d' \in D_a} \exp(V_{od'tp}^a)}$$

Neither  $N_{otp}^a$  nor the denominator depends on the destination  $d$ . I define  $\exp(\delta_{otp}) \equiv N_{otp}^a / \sum_{d' \in D_a} \exp(V_{od'tp}^a)$  and substitute  $V_{odtp}^a$  from equation (4). This gives the estimating equation

$$\mathbb{E}[N_{odtp}^a] = \exp\left(\theta_a \log \text{TravelCost}_{od} + \sum_{b \geq 1} \beta_b^a \mathbb{1}[\text{smoke}_{dt} \in b] + X_{dt}\Gamma^a + \delta_{dp}^a + \delta_{otp}\right) \quad (7)$$

I estimate equation (7) separately for each activity by Poisson pseudo-maximum likelihood (PPML). Each observation is an  $o \rightarrow d$  flow in one period of one day, and a single regression per activity stacks all four periods. Smoke effects  $\beta^a$  are thus common across periods, while destination appeal  $\delta_{dp}^a$  and the origin fixed effect  $\delta_{otp}$  vary by period. With origin-by-day-by-period fixed effects the PPML estimates coincide with those of the conditional logit (Guimarães et al., 2003).

The bin coefficients  $\beta_b^a$  recover the semi-elasticity of activity- $a$  trips to destination smoke in bin  $b$  relative to the smoke-free reference. The coefficient  $\theta_a$  on log travel cost recovers the elasticity of trips to travel cost. Conditional on picking an activity, the equation thus measures whether people substitute away from smoky destinations. The origin-by-day-by-period fixed effect  $\delta_{otp}$  restricts the identifying variation to within an origin, day and period. It absorbs anything that shifts the total number of trips out of origin  $o$  for activity  $a$  in that period, including the number of people choosing the activity and smoke at the origin. The destination-by-period fixed effect  $\delta_{dp}^a$  absorbs each destination's baseline appeal at each time of day (e.g. a park at noon versus at dusk), and I control for destination temperature and precipitation. Within these fixed effects,  $\theta_a$  is identified from the gradient in travel cost across destinations for a given origin, day and period. The smoke coefficients  $\beta_b^a$  are identified by comparing the same destination on smokier and cleaner days for travelers from the same origin in the same period. They therefore capture rerouting across destinations in response to smoke.

Equation (7) takes the form of a gravity equation. It relates bilateral trip flows to an origin term, a destination term and a bilateral travel cost, which links the destination choice to the trade and urban literature on gravity models of flows (Ahlfeldt et al., 2015; Couture and Handbury, 2020; Rossi-Hansberg, 2010). It can equivalently be read as the demand equation of a travel cost model. This method is widely used in environmental economics to value non-market amenities such as parks or beaches by relating the number of visits to the cost of reaching a site (English et al., 2018; McConnell, 1985; Gellman et al., 2025). My approach extends the travel cost logic to a setting with many destinations where the full choice set is observed. This makes it possible to measure substitution across destinations, a margin most travel cost studies cannot recover. Unlike the fixed site amenities valued in standard travel cost studies, smoke also varies from day to day at each destination.

**Estimation of the activity choice.** The activity choice is estimated with the same Poisson count approach. Each of the  $N_{otp}$  devices at origin  $o$  in period  $p$  of day  $t$  chooses activity  $a$  with probability  $P_{a|otp}$  from equation (6). The expected number choosing  $a$  is therefore

$$\mathbb{E}[N_{otp}^a] = N_{otp} \frac{\exp\left(\sum_{b \geq 1} \gamma_a^b \mathbb{1}[\text{smoke}_{ot} \in b] + \mu \mathcal{V}_{otp}^a\right)}{\sum_{a'} \exp\left(\sum_{b \geq 1} \gamma_{a'}^b \mathbb{1}[\text{smoke}_{ot} \in b] + \mu \mathcal{V}_{otp}^{a'}\right)}$$

Neither  $N_{otp}$  nor the denominator depends on the activity. As in the destination choice they are absorbed by an origin-by-day-by-period fixed effect  $\delta_{otp}$ . I also include an origin-by-activity-by-period fixed effect  $\delta_{oap}$  to absorb the baseline activity composition of each origin, and an activity-by-day-by-period fixed effect  $\delta_{atp}$  to capture common daily activity levels such as differences between weekdays and weekends. The estimating equation is

$$\mathbb{E}[N_{otp}^a] = \exp\left(\sum_{b \geq 1} \gamma_a^b \mathbb{1}[\text{smoke}_{ot} \in b] + \mu \mathcal{V}_{otp}^a + \delta_{otp} + \delta_{oap} + \delta_{atp}\right) \quad (8)$$

I estimate equation (8) by PPML in a single regression over five alternatives (outdoor, work, other, going home and staying home) that stacks the four periods. As for the destination choice, the PPML estimates coincide with those of the multinomial logit (Guimarães et al., 2003).

Origin smoke takes the same value for all activities within an origin and day, so its direct effect is absorbed by  $\delta_{otp}$ . The coefficients  $\gamma_a^b$  are therefore identified only through how smoke shifts the choice between activities relative to staying home. With  $\delta_{oap}$  and  $\delta_{atp}$  they are identified from how an origin's activity shares change on smokier days compared with cleaner days in the same period. The dissimilarity parameter  $\mu$  is identified from variation in the inclusive value  $\mathcal{V}_{otp}^a$  across activities within an origin, day and period.

## 4 Results

Smoke changes behavior on two margins. People reroute away from smoky destinations, and when smoke reaches their own neighborhood they stay home. Outdoor recreation bears most of the adjustment on both margins. I first present the destination choice estimates from equation (7) and then the activity choice estimates from equation (8).

I find that individuals substitute away from outdoor trips, and that responses increase with the intensity of the smoke (see Figure 4 and Table 1 for coefficients). Trips to outdoor destinations show a clear dose-response. Smoke below  $10\mu\text{g}/\text{m}^3$  has no detectable effect, while smoke between 10 and  $35\mu\text{g}/\text{m}^3$  reduces trips to a destination by about 8.3% relative to other destinations, smoke

between 35 and 75 $\mu\text{g}/\text{m}^3$  by 13.3% and smoke above 75 $\mu\text{g}/\text{m}^3$  by 24.3%.<sup>11</sup> I find no detectable substitution patterns for trips to “other” destinations.

The coefficient on log travel cost is  $-2.59$  for outdoor trips and  $-3.10$  for other trips (Table 1), so trips are highly sensitive to the cost of reaching a destination. Most comparable estimates in the literature are elasticities with respect to travel time. To compare with them I also estimate the specification in log travel time and obtain  $-2.93$  for outdoor and  $-3.55$  for other trips. These magnitudes are in line with the upper end of existing estimates (Table A9).

The effect of smoke on destination choice is robust to different specifications. First, to check that the smoke coefficients are robust to how travel cost is specified, I re-estimate the model with origin-by-destination fixed effects. These absorb every time-invariant feature of a route, including its travel cost. The smoke coefficients are essentially unchanged (Table A10), so the effect of destination smoke on trips is robust to the specification of travel costs. I also run the model separately for each period. While I cannot reject that the response is the same across all periods, the destination substitution is most precisely estimated during the midday and afternoon periods of the model (see Figure A15). I also verify the sensitivity of results to different choices for the partition of smoke-induced  $\text{PM}_{2.5}$ , considering a finer partition with 1, 1-5, 5-10, 10-15, 15-25, 25-35, and 35+  $\mu\text{g}/\text{m}^3$ , a partition with more weight on heavy smoke with 1, 1-25, 25-50, 50-75, 75-100, 100-150, and 150+  $\mu\text{g}/\text{m}^3$  (see Figure A14). Finally, a spline in smoke concentration yields a dose-response very close to the binned estimates (Figure A14).

---

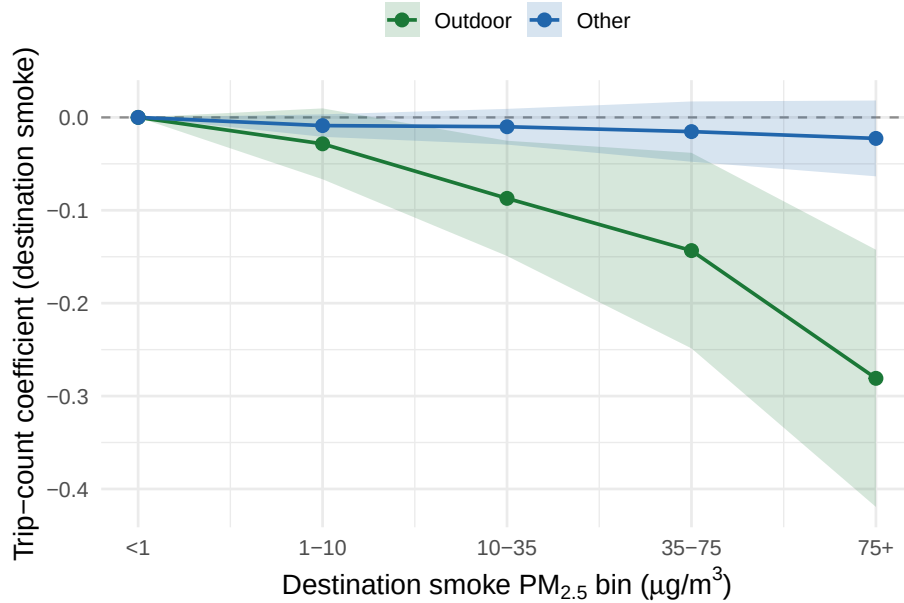
<sup>11</sup>Percentages are computed as  $\exp(\beta_b^a) - 1$ . Since each destination accounts for a small share of trips, they closely approximate the percentage change in trips to that destination.

Table 1: Effect of smoke on destination choice

| Smoke PM <sub>2.5</sub>                    | Outdoor                | Other                  |
|--------------------------------------------|------------------------|------------------------|
| 1–10 $\mu\text{g}/\text{m}^3$              | -0.0293<br>(0.0193)    | -0.0087<br>(0.0060)    |
| 10–35 $\mu\text{g}/\text{m}^3$             | -0.0865**<br>(0.0320)  | -0.0100<br>(0.0098)    |
| 35–75 $\mu\text{g}/\text{m}^3$             | -0.1431**<br>(0.0533)  | -0.0160<br>(0.0160)    |
| 75+ $\mu\text{g}/\text{m}^3$               | -0.2783***<br>(0.0702) | -0.0232<br>(0.0201)    |
| log(travel cost)                           | -2.5880***<br>(0.0265) | -3.0990***<br>(0.0155) |
| Observations                               | 6,418,673              | 56,383,751             |
| Pseudo $R^2$                               | 0.188                  | 0.336                  |
| Period $\times$ destination ZIP FE         | Yes                    | Yes                    |
| Origin ZIP $\times$ day $\times$ period FE | Yes                    | Yes                    |

*Notes:* PPML estimates of equation (7), estimated separately for outdoor and other trips. The dependent variable is the number of  $o \rightarrow d$  trips in a period of a day. Observations stack the four periods of the day (0–11AM, 11AM–4PM, 4–7PM, 7PM–12AM). Smoke coefficients are common across periods and measured relative to destination smoke PM<sub>2.5</sub> below  $1\mu\text{g}/\text{m}^3$ . All specifications control for destination temperature and precipitation. Standard errors clustered by ZIP in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Figure 4: Effect of Smoke on Destination Choice



*Notes:* Smoke coefficients  $\beta_b^a$  from Table 1, estimated separately for outdoor and other trips and pooled across the four periods of the day. The reference bin is destination smoke PM<sub>2.5</sub> below  $1\mu\text{g}/\text{m}^3$ . Shaded areas are 95% confidence intervals.

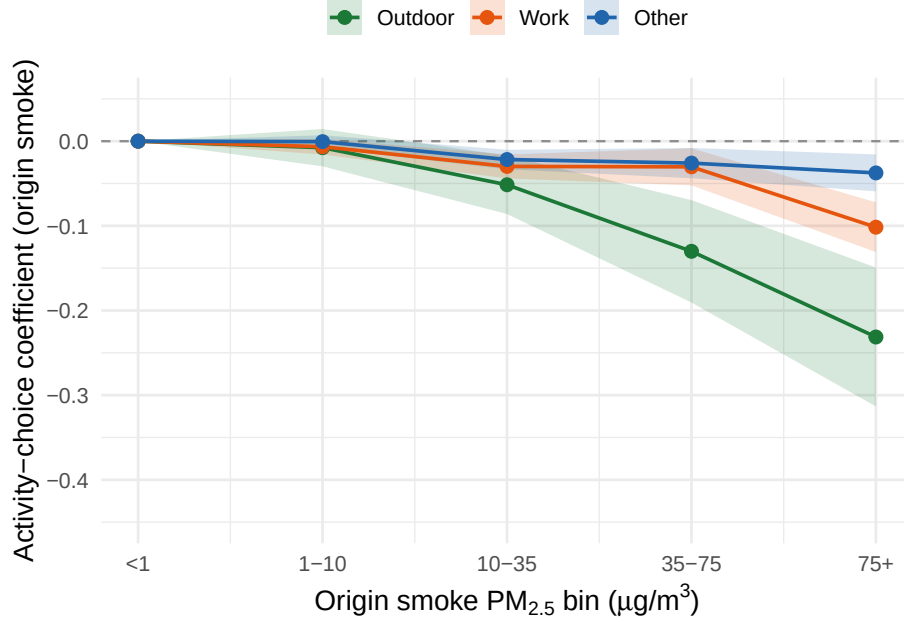
**Tier 1 - Effect of Smoke on Activities** Smoke at the origin location keeps people from heading out. Table 2 and Figure 5 present the estimates of equation (8), where coefficients are changes in the log odds of each activity relative to staying put. Light smoke ( $1\text{--}10\mu\text{g}/\text{m}^3$ ) leaves activity choices unchanged. Beyond  $10\mu\text{g}/\text{m}^3$  every trip type declines, and the decline steepens with smoke intensity. Outdoor recreation is the most sensitive. Relative to staying put, the odds of an outdoor trip fall by 5.0% when origin smoke is between 10 and  $35\mu\text{g}/\text{m}^3$ , by 12.2% between 35 and  $75\mu\text{g}/\text{m}^3$  and by a fifth (20.6%) above  $75\mu\text{g}/\text{m}^3$ . Work and other trips are more resilient. Even above  $75\mu\text{g}/\text{m}^3$  the odds of a work trip fall by 9.7% and those of an other trip by only 3.7%.<sup>12</sup> People thus cut back most on outdoor activities, which are both the most discretionary and the most exposed to smoke.

Beyond the smoke effects, the activity choice estimates the dissimilarity parameter  $\mu$ , the coefficient on the inclusive value that links the two tiers of the model. It measures how correlated the unobserved utilities of destinations within an activity are. Its estimate of 0.26 lies inside the unit interval required for consistency with random utility maximization. This low value indicates strong correlation within activities, which supports the nesting structure: destinations of the same type are much closer substitutes for each other than for other activities. A smoky destination is therefore mostly replaced by another destination for the same activity rather than by a different activity. This also implies that destination smoke has only a muted effect on the activity choice, which the welfare decomposition reflects. Finally, the activity choice estimates are very similar on the full sample that includes devices without an identified workplace (Figure A18), so the results are robust to sample selection.

---

<sup>12</sup>Percentages are computed as  $\exp(\gamma_a^b) - 1$ . Trips home are estimated as a separate alternative so that returns home appear as their own category. They are not valued in the welfare computation.

Figure 5: Activity choice by origin smoke exposure



= 95% CI. Tier-1 (activity choice); coefficients pooled across periods (0-24). Trips to home alt. estimated but not shown.

Notes: Origin smoke coefficients  $\gamma_a^b$  from Table 2, pooled across the four periods of the day. Each coefficient is relative to staying at origin ZIP location and to origin smoke PM<sub>2.5</sub> below 1µg/m³. Trips home are estimated but not shown. Shaded areas are 95% confidence intervals.

Table 2: Activity choice: origin smoke reduces trips out of the home

| Smoke PM <sub>2.5</sub>              | Outdoor                | Work                   | Other                  | Trips home             |
|--------------------------------------|------------------------|------------------------|------------------------|------------------------|
| 1-10 µg/m³                           | -0.0077<br>(0.0113)    | -0.0065<br>(0.0047)    | -0.0006<br>(0.0039)    | -0.0061<br>(0.0059)    |
| 10-35 µg/m³                          | -0.0515**<br>(0.0176)  | -0.0298***<br>(0.0073) | -0.0217***<br>(0.0059) | -0.0190*<br>(0.0090)   |
| 35-75 µg/m³                          | -0.1301***<br>(0.0309) | -0.0303**<br>(0.0112)  | -0.0259**<br>(0.0091)  | -0.0646***<br>(0.0143) |
| 75+ µg/m³                            | -0.2313***<br>(0.0418) | -0.1016***<br>(0.0150) | -0.0375***<br>(0.0111) | -0.0780***<br>(0.0182) |
| $\mu$ (V_ot)                         | 0.2574*** (0.0651)     |                        |                        |                        |
| Observations                         | 1,743,419              |                        |                        |                        |
| Pseudo R²                            | 0.810                  |                        |                        |                        |
| Alternative × day × period FE        | Yes                    | Yes                    | Yes                    | Yes                    |
| Origin ZIP × day × period FE         | Yes                    | Yes                    | Yes                    | Yes                    |
| Alternative × origin ZIP × period FE | Yes                    | Yes                    | Yes                    | Yes                    |

Notes: Estimates of the activity choice in equation (6). Coefficients are relative to staying at origin ZIP location and to origin smoke PM<sub>2.5</sub> below 1µg/m³, and are pooled across the four periods of the day (0-11AM, 11AM-4PM, 4-7PM, 7PM-12AM).  $\mu$  is the coefficient on the inclusive value  $\mathcal{V}_{otp}^a$  from the destination choice. Standard errors clustered by ZIP×date in parentheses. \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

## 5 Welfare Computation

**Compensating Variation.** To monetize the welfare effect of smoke, I construct a counterfactual that is entirely smoke-free, i.e. where there is no smoke neither at the origin nor at the destination. I use this counterfactual to compute the compensating variation associated with moving from observed smoke conditions to a smoke-free counterfactual. Smoke lowers the utility of each away alternative. Let  $\mathcal{V}_{ot}^{a,cf}$  denote the counterfactual inclusive value—the log-sum recomputed with destination smoke set to its smoke-free level—so that  $\mathcal{V}_{ot}^{a,cf} - \mathcal{V}_{ot}^a$  is the gain in destination-tier value from removing smoke at the destinations. The utility change from removing smoke is

$$\begin{aligned}\Delta U_{ot}^a &= U_{ot}^{a,cf} - U_{ot}^a \\ \Delta U_{ot}^a &= \underbrace{\mu(\mathcal{V}_{ot}^{a,cf} - \mathcal{V}_{ot}^a)}_{\text{destination channel}} - \underbrace{\beta_a(\text{smoke}_{ot})}_{\text{origin channel}}.\end{aligned}$$

Smoke reduces utility both through the destination channel (change in the inclusive value when destination smoke is removed) and the origin channel (direct shift in the activity choice based on origin smoke). On clean-air days,  $\Delta U_{ot}^a = 0$ . Following the standard log-sum welfare formula for discrete choice (Small and Rosen, 1981; Train, 2009), the average per-person utility loss at origin  $o$  on day  $t$  is

$$W_{ot} = \log\left(\sum_a P_{ot}^a \exp(\Delta U_{ot}^a)\right) \quad (9)$$

where the  $P_{ot}^a$  are observed choice shares rather than fitted probabilities. Each activity's loss is weighted by the fraction of people actually choosing it, so the aggregate is driven as much by the distribution of shares as by the size of the per-activity losses due to smoke PM<sub>2.5</sub>.

**Conversion to dollars.**  $W_{ot}$  is measured in utils. I convert it to dollars using the travel cost coefficient as the numeraire. Because travel cost is expressed in dollars, differentiating utility with respect to travel cost gives the marginal utility of a dollar  $\frac{\partial U_{ot}^a}{\partial \text{cost}}$ . To obtain the dollar value of a util  $\lambda_a$ :

$$\lambda_a = 1 / \frac{\partial U_{ot}^a}{\partial \text{cost}}.$$

We compute  $\frac{\partial U_{ot}^a}{\partial \text{cost}}$  following the expression of  $U_{ot}^a$  in (3). Travel cost affects the activity choice through the inclusive value  $\mathcal{V}_{ot}^a$ , which enters the upper tier scaled by the nest parameter  $\mu$ . By the chain rule,

$$\frac{\partial U_{ot}^a}{\partial \text{cost}} = \underbrace{\frac{\partial U_{ot}^a}{\partial \mathcal{V}_{ot}^a}}_{=\mu} \cdot \underbrace{\frac{\partial \mathcal{V}_{ot}^a}{\partial \text{cost}}}_{=\theta/\text{cost}}.$$

Because travel cost enters  $V_{odt}^a$  in a log form, the derivative  $\frac{\partial V_{odt}^a}{\partial \text{cost}_{od}}$  needs to be evaluated at a certain point of the cost distribution. The travel-cost coefficient itself is held common across activities: I use the single coefficient  $\theta$  estimated for the “other” nest as the numeraire for all channels, since the welfare conversion requires one marginal utility of income rather than an activity-specific one. The dollar value of a util for activity  $a$   $\lambda_a$  is then

$$\begin{aligned}\frac{\partial U_{otp}^a}{\partial \text{cost}} &= \frac{\mu \theta}{\bar{c}_a} \\ \lambda_a &= \frac{\bar{c}_a}{\mu \theta}.\end{aligned}$$

**Aggregation.** The per-person dollar loss at origin  $o$  on day  $t$  in period  $p$  converts the log-sum loss in (9) to dollars:

$$\text{CV}_{otp} = \lambda_p W_{otp} \quad (10)$$

where  $\text{CV}_{otp}$  is the loss for a representative individual at that origin–day–period. Welfare is computed using period-specific coefficients, so  $\lambda_p = \bar{c}_{\text{other},p}/(\mu \theta_{\text{other},p})$ . Because the utility of staying in and going home is normalized to zero, the measure prices forgone trips. Scaling by population and summing over all origins, days, and periods gives the total welfare cost of smoke:

$$\mathcal{W} = \sum_o \text{pop}_o \sum_t \sum_p \text{CV}_{otp}.$$

**Results.** Summing across periods gives a total welfare cost of approximately \$500M, or \$5 per person per smoke day above  $30 \mu\text{g}/\text{m}^3$ .

The loss is concentrated in the periods with the most travel: morning, midday, and evening together account for 99% of the total (see Table 3 and Figure 1 for trip density across the day). Per-person losses rise steadily with concentration, from \$0.45 in the  $1\text{--}10 \mu\text{g}/\text{m}^3$  range to \$6.6 above  $75 \mu\text{g}/\text{m}^3$  (see Table A13). However, because moderate-smoke days are far more common than severe ones,  $10\text{--}35 \mu\text{g}/\text{m}^3$  days weigh almost as much in the total welfare loss as days above  $75 \mu\text{g}/\text{m}^3$  (29% and 26% respectively). Table 4 decomposes the baseline welfare loss by activity type. Outdoor recreation exhibits the largest proportional reduction in trips, yet accounts for only 8.2% (\$41.6M) of total welfare losses. Although they respond less to smoke than outdoor trips, work and “other” trips make up a large share of all trips and thus account for a larger share of the welfare loss (53.9% and 37.9% respectively). Valuing avoidance using a comprehensive choice-set across activities therefore reveals that the largest welfare costs of smoke come from small adjustments to everyday trips rather than recreation choices. Because the conversion from utils to dollars is proportional to the mean travel cost  $\bar{c}$  (10), the welfare estimates are sensitive to the assumptions used to compute

Table 3: Welfare decomposition by period

| Period            | Welfare<br>(\$M) | Share<br>(%) | Per person,<br>smoky day (\$) |
|-------------------|------------------|--------------|-------------------------------|
| Morning (0–11)    | 177.7            | 35.2         | 1.544                         |
| Midday (11–16)    | 255.6            | 50.6         | 2.355                         |
| Afternoon (16–19) | 65.7             | 13.0         | 1.233                         |
| Evening (19–24)   | 6.4              | 1.3          | -0.103                        |
| Total             | 505.5            | 100.0        |                               |

*Notes:* Annual welfare losses by time of day, computed with period-specific coefficients and summed across periods. Each period is treated as an independent choice occasion. Sample restricted to workers; home valuation excluded.

travel costs. The baseline travel cost computation (1) follows conventions in the recreation demand literature and uses one-third of the wage rate (U.S. EPA, 2020), which is low compared to other revealed preferences value of time estimates (Goldszmidt et al., 2020). Valuing time at half the wage rate ( $w/2$  instead of  $w/3$ ) raises the total welfare loss from \$505.5M to \$637.6M.

I consider whether the welfare loss stems mainly from destination switching within the nest or from changes in the composition of activities at the origin. I find that reallocation of activities at the origin dominates. The two channels can be separated because each utility change is itself the sum of a destination term,  $\mu(\mathcal{V}_{otp}^{a,cf} - \mathcal{V}_{otp}^a)$ , and an origin term,  $-\beta_a(\text{smoke}_{ot})$ . The contribution  $c_{ot}^a$  is linear in  $\Delta U_{ot}^a$ , so it splits along the same lines into a destination part and an origin part, which sum across alternatives, origins, and days to  $\mathcal{W} = \mathcal{W}^{\text{dest}} + \mathcal{W}^{\text{orig}}$ . The origin channel represents 89% of the total welfare loss, indicating that the bulk of the welfare loss comes from cancelling activities rather than from switching destinations.

Finally, I also test the sensitivity of the welfare cost to the number of choice occasions. In the baseline model, individuals make four choices per day, reflecting that 98% of devices make four trips or less per day (see Figure A6). In Table A14, I consider six alternative partitions of the day, from one to six periods. The estimates range from \$266M to \$675M. The number of choice occasions balances two issues. If occasions are too few, the welfare is understated because the model misses some trips that were cancelled: the daily model with one choice occasion gives a welfare number of \$266M, roughly half of the baseline estimate. With too many occasions, the welfare is also understated because the model adds periods in which individuals have no realistic opportunity to travel, which artificially inflates the share of staying in. Consistent with this, the welfare loss peaks at 3 periods and then declines as additional periods are added.

Table 4: Welfare decomposition by channel

| Channel            | Welfare (\$M) | Share (%) |
|--------------------|---------------|-----------|
| Outdoor recreation | 41.6          | 8.2       |
| Work               | 272.4         | 53.9      |
| Other              | 191.4         | 37.9      |
| Total              | 505.5         | 100.0     |

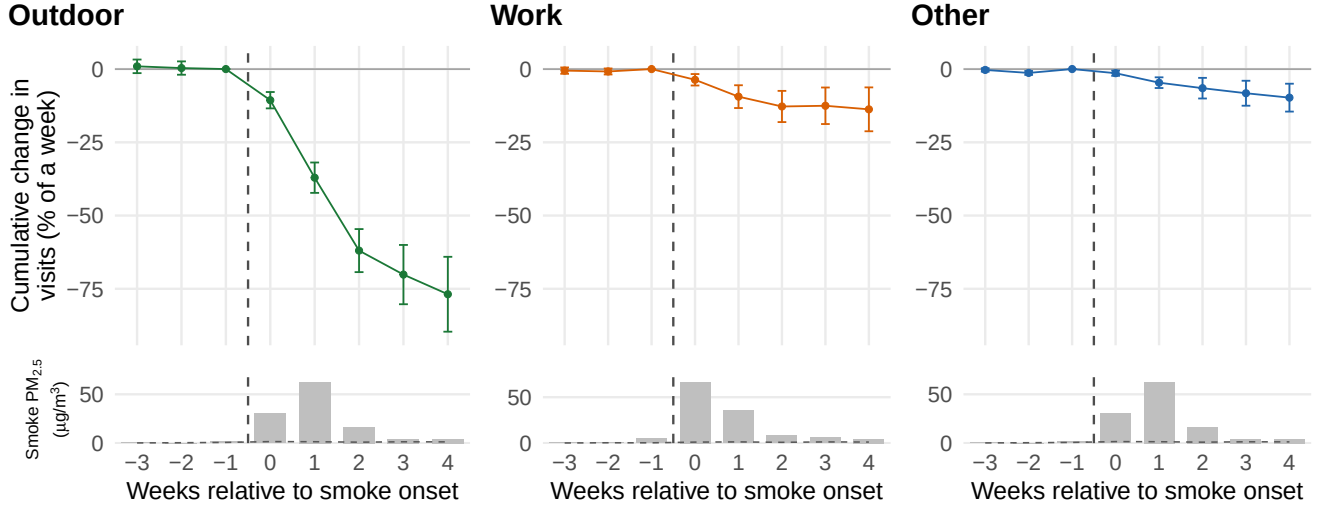
*Notes:* Annual welfare losses by trip purpose. Shares reflect both the size of the behavioral response and the baseline volume of trips of each type, so they need not track the ranking of estimated responses. Sample restricted to workers; home valuation excluded. Components may not sum to the total due to rounding.

**No evidence of intertemporal substitution** The welfare calculation treats every forgone trip as a loss. This would overstate the cost if people shift trips over time instead of cancelling them, for instance by going to the park next week instead of today. Delaying a trip is arguably less costly than cancelling it. The event study of equation (2) (Section 2.3) can tell the two apart, because it follows activity for several weeks after the smoke clears: delayed trips would appear as extra activity in those later weeks. Figure 6 plots the cumulative change in each activity. At each week, it sums the weekly effects since onset, expressed as a percentage of a typical pre-onset week. If people made up for lost trips once the smoke clears, the cumulative change would climb back toward zero. It does not, for any activity: losses build up while smoke is present and stay flat once the air clears. Four weeks after onset, the cumulative loss equals 75% of a typical week of outdoor activity, 13% of a typical week of work trips and 10% of a typical week of other trips. Trips are therefore cancelled rather than delayed, which supports treating them as lost in the welfare calculation.

## 6 Health Damages from the Smoke

How does this avoidance cost compare to other impacts of wildfire smoke? To put these revealed preferences welfare estimate into perspective, I compare them to the estimated contemporaneous mortality cost of exposure to the same smoke. I use California death-count data at the county-month level and estimate the effect of monthly smoke  $\text{PM}_{2.5}$  on the all-cause mortality rate. I include county $\times$ year, month $\times$ year, and county $\times$ month fixed effects, so that identification comes from within-county deviations in smoke net of seasonal and county-specific trends. Following Qiu et al. (2025), I estimate a Poisson model on the death rate, using data from 2002 to 2020 for the entire state. I find that a 1  $\mu\text{g}/\text{m}^3$  increase in monthly smoke raises the mortality rate by roughly 0.19% (see Table 5). The magnitude of this estimate appears in line with effects found by Miller

Figure 6: Cumulative response of trips around smoke event onset



Notes: Stacked event-study estimates of the effect of smoke onset on residents' activities, following [Cengiz et al. \(2019\)](#). The unit is the ZIP by week (Monday–Friday excluding holidays for work trips). A ZIP has an onset in week  $t$  if smoke  $\text{PM}_{2.5}$  exceeds  $35\mu\text{g}/\text{m}^3$  on at least two days of  $t$  after averaging below  $10\mu\text{g}/\text{m}^3$  in each of the three preceding weeks (700 events for outdoor and other trips and 413 for work trips, across 8 onset weeks). Each event dataset spans three weeks before to four weeks after onset and pairs treated ZIPs with control ZIPs whose weekly smoke stays below  $10\mu\text{g}/\text{m}^3$  throughout. The outcome is the number of device-periods residents spend on the activity. Estimation is PPML with the log of all device-periods as offset, so coefficients are proportional changes in the share of the week spent on the activity. Regressions include ZIP-by-event and week-by-event fixed effects and weekly temperature and precipitation. Coefficients are relative to the three pre-onset weeks and cumulated from three weeks before onset. 95% confidence intervals use the delta method with standard errors clustered by ZIP. Lower panels show mean smoke  $\text{PM}_{2.5}$  in treated (grey bars) and control ZIPs (dashed line). Sample: California, June–December 2018, excluding ZIPs with at least 1% of their area inside a 2018 fire perimeter.

[et al. \(2024\)](#).<sup>13</sup>

Applying this dose-response to my 2018 Northern California sample,<sup>14</sup> I find that wildfire caused approximately 576 excess deaths from July to November 2018. This estimate is a lower bound since it does not take into effect lagged mortality effects of wildfire smoke ([Qiu et al., 2025](#); [Connolly et al., 2024](#)).

To monetize the death counts, I use a value-of-statistical-life (VSL) approach, which assigns a monetary value to every life lost. I follow the EPA and use a value of a statistical life of \$8.9 millions for every life lost. Large evidence suggests that acute smoke mortality is concentrated among the elderly with limited remaining life expectancy ([Qiu et al., 2025](#)). Following [Deryugina et al. \(2019\)](#) and [Miller et al. \(2024\)](#), I estimate that the remaining life expectancy of the elderly is

<sup>13</sup>[Miller et al. \(2024\)](#) find that a  $1\mu\text{g}/\text{m}^3$  increase in daily  $\text{PM}_{2.5}$  raises three-day elderly mortality by 0.20 deaths per million for a mean daily mortality of 122.8 deaths.  $0.20/122.8$  implies a roughly 0.16% increase in mortality per  $\mu\text{g}/\text{m}^3$ , close to my 0.19% estimate.

<sup>14</sup>I exclude Butte county from this exercise, since 84 people directly died from the Camp Fire in that county in 2018.

about 3.5 years and value this using a value of a statistical life year (VSL-Y) of \$ 385,000, while valuing nonelderly deaths at the VSL of \$8.9M. Using the age-specific concentration-response estimates of [Qiu et al. \(2025\)](#) to allocate deaths across age groups, I attribute roughly 70% of deaths to those aged 65 and above. This yields a total immediate mortality cost of approximately \$2 billion.

The mortality costs are concentrated on the elderly population, while my revealed preferences estimates concentrate on a sample of full-time workers, which skews younger. The revealed preference cost of avoidance behavior for this population amounts to \$500M, which represents about 24% of the mortality cost of the same smoke plumes, or a significant additional cost of wildfire smoke that was previously unaccounted for. In addition to activity reallocation, there are many other costs of adapting to wildfire smoke that this model does not take into account, such as mask wearing, air filter purchase, or improving HVAC systems. Last, this estimated revealed preference cost of wildfire smoke is likely to be an underestimate, because individuals only react to the smoke they perceive. The health literature documents that low to moderate levels of smoke are the most harmful to human health, consistent with these levels of smoke being the most prevalent and harder to perceive. I test this hypothesis in the next section by considering the impact of air quality alerts to ask whether information shifts behavior.

Table 5: Effect of monthly smoke  $PM_{2.5}$  on the all-cause mortality rate

|                            | Death Rate <sub>m</sub> |
|----------------------------|-------------------------|
| average smoke $PM_{2.5_m}$ | 0.00192**<br>(0.000901) |
| county x year FE           | yes                     |
| month x year FE            | yes                     |
| county x month FE          | yes                     |
| N                          | 8893                    |

## 7 Air quality alerts

Most of the welfare loss happens on days with low or moderate smoke. These are also the days when smoke is least noticeable, and the modest responses are consistent with individuals not minding smoke at these levels or not being aware of it. In the mobility data, choices are observed under a single information environment, so that awareness and preferences enter as a product and cannot be separated. Separating the two mechanisms requires variation in information provision itself. I use air quality alerts for that variation. Alerts are issued on a deterministic rule on the pollution forecast, so crossing the alert threshold changes information provision but not air quality. If the information carried no news, outdoor trips would be continuous through the threshold. I estimate this discontinuity, and then feed the response back into the choice model to compute the welfare cost under improved smoke awareness.

I study the Spare the Air program, which operates in the air districts of the Bay Area and Sacramento. The alerts bundle messaging about awareness, health, and emissions. When pollution is forecast to reach unhealthy levels, the program issues an alert warning residents about the health impacts of high pollution and asks residents to reduce trips and stay indoors if smoke is present. In the Sacramento basin, the program was designed for ozone. Since ozone forms with heat and sunlight, it is a summer issue and the alert program only runs through the summer months. The Bay Area program covers all pollutants, including wildfire smoke. During the summer of 2018, smoke and ozone episodes frequently coincided in both basins. I also restrict the period of study to the summer of 2018 because of the running window of the Sacramento program. Because air pollution in Northern California in 2018 was so elevated, 33 spare the air alert days were called in Sacramento (20) and in the Bay Area (13) during the period of interest.

Spare the Air alerts are issued based on discretionary thresholds, making the setting suitable for a regression discontinuity design (Neidell, 2009; Cutter and Neidell, 2009). Each air district produces a daily basin-level forecast of next-day pollution and issues an alert when the forecast exceeds its threshold: 100 AQI in the Bay Area and 126 AQI in Sacramento. I obtained the forecast data directly from the air districts, so the running variable is the same series the districts act on. I find that the forecast is deterministic for alert issuance, so that assignment is sharp at the cutoff (see Figure A19). The difference in thresholds also means the discontinuity falls at different points of the pollution distribution in the two basins, additional evidence that the estimated effect is not an artifact of behavior at one particular AQI level.

I follow the approach of Neidell (2009) and estimate the effect of an alert on outdoor trips using a regression discontinuity in the forecast. For ZIP code  $i$  in basin  $r$  on day  $t$ , I estimate the Poisson

model

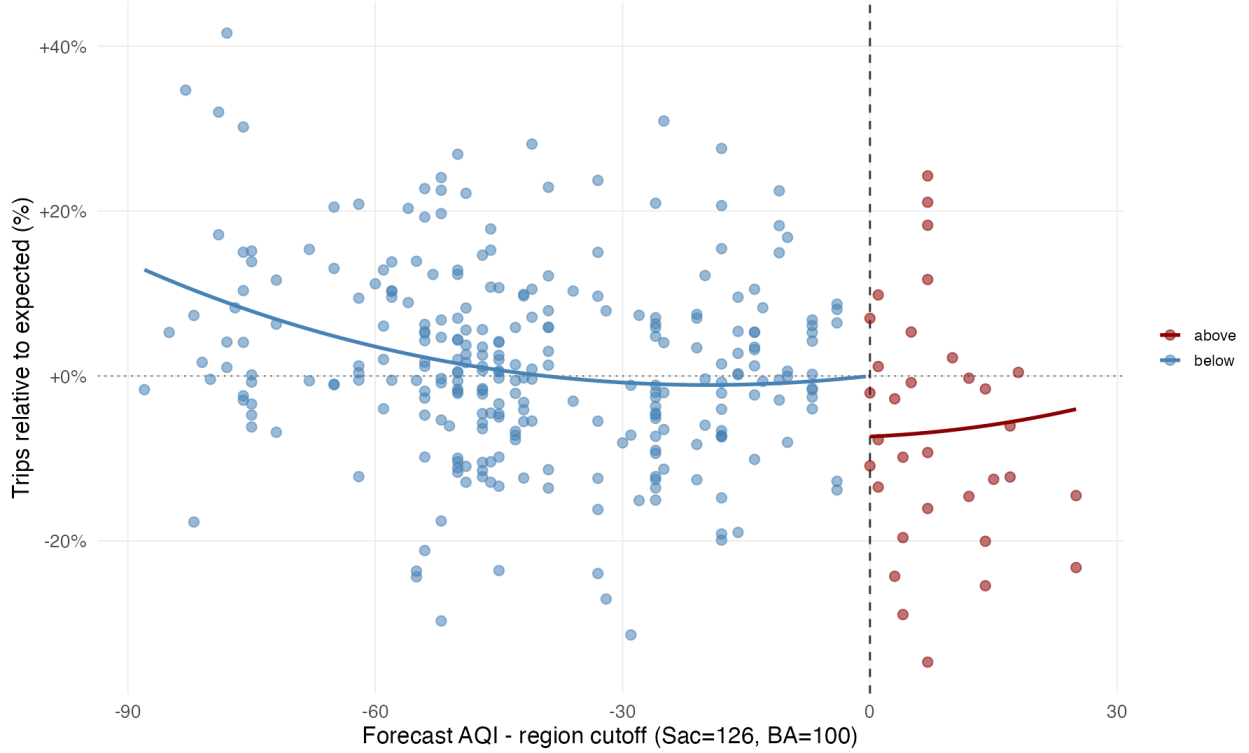
$$\log E[y_{irt}] = \beta \text{alert}_{rt} + \gamma_1 (AQI_{rt}^f - c_r) + \gamma_2 \text{alert}_{rt} \cdot (AQI_{rt}^f - c_r) + X'_{it}\delta + \eta_{it}$$

where  $y_{irt}$  is the number of outdoor trips originating from ZIP  $i$  on day  $t$  — the same outcome that enters the welfare construction —  $AQI_{rt}^f$  is the basin-level forecast, and  $c_r$  is the basin-specific alert threshold (100 in the Bay Area, 126 in Sacramento). The running variable  $AQI_{rt}^f - c_r$  is centered at each basin’s cutoff. The coefficient  $\gamma_1$  captures the slope of trips in the forecast below the threshold, and  $\gamma_2$  allows the slope to differ above the threshold. The slope above the cutoff is  $\gamma_1 + \gamma_2$  so that trips may respond more (or less) steeply to the forecast once an alert is in effect. Once  $\gamma_1$  and  $\gamma_2$  absorb the average relationship between the AQI forecast and trips,  $\beta$  isolates the discontinuity in trips at the threshold itself. Days with forecasts just above and just below the cutoff face nearly identical air quality, but only days above the cutoff receive an alert. The discontinuity therefore isolates the effect of the alert itself rather than pollution, which varies smoothly through the cutoff. Because the threshold applies to the forecast AQI regardless of the pollutant driving it,  $\beta$  identifies the effect of an alert per se, not of any particular pollutant.  $X_{it}$  contains temperature and precipitation, and  $\eta_{it}$  denotes ZIP  $\times$  day of week, and week fixed effects. ZIP  $\times$  day of week fixed effects absorb persistent differences in outdoor trips across ZIPs and days of the week (for instance, whether certain ZIPs have a higher share of outdoor trips on weekends). Week fixed effects control for seasonality in outdoor trips. Standard errors are clustered at the basin  $\times$  date level.

An alert reduces expected outdoor trips by 6–10% across bandwidths of 20, 25, and 30 AQI points (Figure 7, Table A17). Results are robust to alternative fixed effects structures. To confirm that the drop is specific to the alert threshold, I re-estimate the discontinuity at placebo cutoffs below the true cutoff point. Estimated coefficients are small, changing sign, and insignificant, so the discontinuity at the true alert threshold passes this placebo test (see Table A19). Existing evidence on air quality alerts come from single activities: zoo and observatory attendance (Neidell, 2009), cycling (Saberian et al., 2017), traffic counters (Dangel and Goeschl, 2026). My setting measures outdoor activity at a high resolution, the primary margin of adjustment for air quality alerts, and aggregates different types of outdoor trips across destinations. The magnitude of the response is in line with Neidell (2009) and Saberian et al. (2017).

I also find that the response is heterogeneous across ZIP codes, with the effect of the alert being largest in higher-income areas. This finding is interesting because research on scientific information suggests that higher income individuals are also the most likely to know about air pollution and its impacts on health (Lange, 2011; Cutler et al., 2011). If they were already informed, they should have the least to learn from the alert. That the opposite holds suggests that socioeconomic status

Figure 7: Effect of Air Quality Alerts on Outdoor Trips



also determines the ability to take up information. Consistent with this hypothesis, the effect of the alert is also weaker in areas with larger shares of Hispanic population. Taken together, these patterns suggest that the information channel operates differentially across communities.

## 7.1 Alerts and welfare

The alert response has implications for the welfare estimates. Section 5 showed that the welfare cost prices smoke as perceived, so that if awareness is imperfect estimated responses are attenuated. The regression discontinuity analysis establishes that it is: individuals respond more when information provision increases, implying that the welfare cost is understated compared to a counterfactual where individuals are more aware. How much does the welfare cost increase with improved awareness? To answer this question, I incorporate the alert response as an awareness parameter into the model. This counterfactual requires a scale factor on the smoke coefficients, since awareness mediates the behavioral response to smoke. The discontinuity measures a change in trips. Converting one into the other requires knowing how much of a trip response a given change in awareness produces.

I impose a two-point structure on beliefs, governed by a single parameter  $\lambda$ , the awareness share.  $\lambda$  is the probability an individual assigns to smoke being at its realized level. On a day with smoke in bin  $b$ , an individual believes with probability  $\lambda$  that smoke is in bin  $b$ , and with probability

$1 - \lambda$  that the air is clean. Individuals therefore either perceive smoke at its realized level or perceive clean air.<sup>15</sup> In this model, alerts change the awareness level by  $\Delta\lambda$ . Every estimated smoke coefficient is its full-awareness counterpart scaled by a baseline awareness level  $\lambda_0$ , the level prevailing on days without an alert, such that  $\hat{\gamma}_a = \lambda_0\gamma_a$  and  $\hat{\beta}_a = \lambda_0\beta_a$ .

Under this parametrization, individuals act on a share  $\lambda_0$  of the smoke present below the cutoff and on  $\lambda_0 + \Delta\lambda$  above it. Crossing the threshold only changes the perceived smoke, raising it by  $\Delta\lambda$ . In the model, this increase lowers the utility of each activity and trips fall. I denote  $\bar{B}$  the trip response per unit of awareness, computed from smoke coefficients and fitted choice probabilities. The discontinuity recovers the awareness gain  $\Delta\lambda$  relative to its baseline  $\lambda_0$ . Appendix A.1 shows that the jump in trips at the threshold divided by  $\bar{B}$  is

$$\frac{\beta_{RD}}{\bar{B}} = \frac{\Delta\lambda}{\lambda_0} \equiv r. \quad (11)$$

The strategy is to divide an observed behavioral response ( $\beta_{RD}$ ) by the response the model predicts from a given change in awareness ( $\bar{B}$ ). This follows the approach of Chetty et al. (2009), who use a change in the salience of a sales tax to identify how much attention consumers pay to taxes. The discontinuity gives  $\beta_{RD} = -0.069$  and the model gives  $\bar{B} = -0.024$  so  $r = 2.45$ . An alert raises awareness by roughly 2.5 times its baseline level, implying that individuals perceive on the order of 40% of the smoke present on days below the alert threshold. I construct a counterfactual scenario where I use  $r$  to scale responses on days with moderate smoke (10-35  $\mu\text{g}/\text{m}^3$ ). For instance, the outdoor destination response to moderate smoke moves from a 6.7% to a 15.5% reduction in trips. Figure A22 shows the shift for each channel. Aggregating across channels, the welfare cost rises by 20% with improved awareness.<sup>16</sup>

The large increase in smoke responses in this counterfactual scenario suggest that awareness is a binding constraint for adaptation, and that individuals avoid little at least in part because they perceive little. In addition, this counterfactual prices a policy-attainable level of awareness, not full information. It prices the awareness an existing alert program already produces, extended to days that currently receive no alert. An alert reaches only part of the unaware population, so it closes only part of the gap between baseline awareness and full information. That a partial intervention produces a large effect suggests substantial benefits to extending public health efforts to improve information provision on days that currently receive no alert. A limitation of this analysis is that the alert bundles information about different pollutants with appeals to reduce driving. I therefore

<sup>15</sup>This implies that they cannot perceive an intermediate level and that individuals never overestimate smoke. See Appendix A.1 for further discussion.

<sup>16</sup>For this counterfactual, I restrict the analysis to areas within the air districts during the common alert season (June-September). For this sample, I scale days when smoke is above  $10\mu\text{g}/\text{m}^3$ . The correction applies to 3.7% days in the sample.

cannot separate belief revision about smoke from these other channels, so  $r$  should be read as an upper bound on the awareness gain.

## 8 Conclusion

This paper prices the behavioral cost of wildfire smoke. Using smartphone mobility data for Northern California over the 2018 season, I estimate a nested logit model of activity and destination choice in which smoke enters both where people are and where they might go. I use the travel cost margin to convert the resulting utility losses into dollars.

Three results stand out. First, avoidance is not confined to recreation. Outdoor activity responds most strongly, but work and other trips respond as well, and because there are far more trips to work and “other” destinations, they account for most of the welfare loss. Second, the burden is concentrated as much on mild and moderate smoke days as on severe ones: more than half of the loss accrues below  $35\mu\text{g}/\text{m}^3$ , mirroring the distribution of health damages. Third, information matters for how much avoidance occurs. Air quality alerts reduce outdoor trips at the issuance threshold, and extending that awareness to smoke days that fall below the threshold would raise avoidance costs by around 20%.

Together these findings put this avoidance cost revealed through location choice at roughly \$500M for a single season in one region, adding about 24% to the mortality burden of the same smoke and falling on a younger and more active population. Avoidance costs are revealed by behavior, so they capture only the smoke people notice and respond to, and they omit protective margins that mobility data cannot observe. This framework could be used to estimate the welfare costs of different types of environmental damages such as extreme heat.

## References

- Gabriel M. Ahlfeldt, Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf. The economics of density: Evidence from the berlin wall. *Econometrica*, 83(6):2127–2189, 2015. doi: <https://doi.org/10.3982/ECTA10876>. URL <https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA10876>.
- Husnain F Ahmad, Matthew Gibson, Fatiq Nadeem, Sanval Nasim, and Arman Rezaee. Expectations and adaptation to environmental threats. *The Economic Journal*, 136(674):750–779, 02 2026. ISSN 0013-0133. doi: 10.1093/ej/ueaf071. URL <https://doi.org/10.1093/ej/ueaf071>.
- American Automobile Association. Your driving costs: How much are you really paying to drive? Technical report, AAA Association Communication, Heathrow, FL, 2018.
- Susan Athey, Billy Ferguson, Matthew Gentzkow, and Tobias Schmidt. Estimating experienced racial segregation in us cities using large-scale gps data. *Proceedings of the National Academy of Sciences*, 118(46):e2026160118, 2021. doi: 10.1073/pnas.2026160118. URL <https://www.pnas.org/doi/abs/10.1073/pnas.2026160118>.
- David Atkin, M. Keith Chen, and Anton Popov. The returns to face-to-face interactions: Knowledge spillovers in silicon valley, 6 2022.
- Panle Jia Barwick, Shanjun Li, Liguang Lin, and Eric Yongchen Zou. From fog to smog: The value of pollution information. *American Economic Review*, 114(5):1338–81, 5 2024. doi: 10.1257/aer.20200956. URL <https://www.aeaweb.org/articles?id=10.1257/aer.20200956>.
- Patrick Baylis, Judson Boomhower, Jonathan M Colmer, and John L Voorheis. The distribution and consequences of natural disasters: Evidence from u.s. wildfire victims. Working Paper 35675, National Bureau of Economic Research, August 2026. URL <http://www.nber.org/papers/w35675>.
- Clara Berestycki and M. Keith Chen. Behavioral responses to wildfire smoke: Insights from smartphone location data. *Proceedings of the National Academy of Sciences*, 123(18):e2527320123, 2026. doi: 10.1073/pnas.2527320123. URL <https://www.pnas.org/doi/abs/10.1073/pnas.2527320123>.
- Mark Borgschulte, David Molitor, and Eric Zou. Air pollution and the labor market: Evidence from wildfire smoke. *Review of Economics and Statistics*, pages 1–46, 9 2024. ISSN 0034-6535. doi: 10.1162/rest\_a\_01243.

- CAL FIRE. California historical fire perimeters. Fire and Resource Assessment Program (FRAP), 2024. URL <https://www.fire.ca.gov/what-we-do/fire-resource-assessment-program/fire-perimeters>. California Department of Forestry and Fire Protection. GIS dataset, version fire23\_1.
- Doruk Cengiz, Arindrajit Dube, Attila Lindner, and Ben Zipperer. The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454, 2019. doi: 10.1093/qje/qjz014. URL <https://doi.org/10.1093/qje/qjz014>.
- Frank J. Cesario. Value of time in recreation benefit studies. *Land Economics*, 52(1):32–41, 1976. doi: 10.2307/3145598.
- Raj Chetty, Adam Looney, and Kory Kroft. Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–1177, 2009. doi: 10.1257/aer.99.4.1145.
- Marissa L. Childs, Jessica Li, Jeffrey Wen, Sam Heft-Neal, Anne Driscoll, Sherrie Wang, Carlos F. Gould, Minghao Qiu, Jennifer Burney, and Marshall Burke. Daily local-level estimates of ambient wildfire smoke pm<sub>2.5</sub> for the contiguous us. *Environmental Science & Technology*, 56: 13607–13621, 10 2022. ISSN 0013-936X. doi: 10.1021/acs.est.2c02934.
- Congressional Research Service. Funding for wildfire management: FY2025 appropriations for Forest Service and Department of the Interior. In Focus IF13102, Congressional Research Service, 2025. URL <https://www.congress.gov/crs-product/IF13102>.
- Rachel Connolly, Miriam E. Marlier, Diane A. Garcia-Gonzales, Joseph Wilkins, Jason Su, Claire Bekker, Jihoon Jung, Eimy Bonilla, Richard T. Burnett, Yifang Zhu, and Michael Jerrett. Mortality attributable to pm<sub>2.5</sub> from wildland fires in california from 2008 to 2018. *Science Advances*, 10(23):eadl1252, 2024. doi: 10.1126/sciadv.adl1252. URL <https://www.science.org/doi/abs/10.1126/sciadv.adl1252>.
- Victor Couture and Jessie Handbury. Urban revival in america. *Journal of Urban Economics*, 119:103267, 2020. ISSN 0094-1190. doi: <https://doi.org/10.1016/j.jue.2020.103267>. URL <https://www.sciencedirect.com/science/article/pii/S0094119020300383>.
- Victor Couture, Jonathan Dingel, Allison Green, and Jessie Handbury. Demographic preferences and income segregation, 1 2025. URL <http://www.jdingel.com/research/CoutureDingelGreenHandbury.pdf>.
- David M. Cutler, Fabian Lange, Ellen Meara, Seth Richards-Shubik, and Christopher J. Ruhm. Rising educational gradients in mortality: The role of behavioral risk factors. *Journal of Health*

- Economics*, 30(6):1174–1187, 2011. ISSN 0167-6296. doi: <https://doi.org/10.1016/j.jhealeco.2011.06.009>. URL <https://www.sciencedirect.com/science/article/pii/S0167629611001147>.
- W. Bowman Cutter and Matthew Neidell. Voluntary information programs and environmental regulation: Evidence from ‘spare the air’. *Journal of Environmental Economics and Management*, 58(3):253–265, 2009. ISSN 0095-0696. doi: <https://doi.org/10.1016/j.jeem.2009.03.003>. URL <https://www.sciencedirect.com/science/article/pii/S0095069609000527>.
- Alexander Dangel and Timo Goeschl. Air quality alerts and don’t drive appeals: Evidence on voluntary pollution mitigation dynamics from germany. *Journal of Environmental Economics and Management*, 135:103248, 2026. ISSN 0095-0696. doi: <https://doi.org/10.1016/j.jeem.2025.103248>. URL <https://www.sciencedirect.com/science/article/pii/S0095069625001329>.
- Donald R. Davis, Jonathan I. Dingel, Joan Monras, and Eduardo Morales. How segregated is urban consumption? *Journal of Political Economy*, 127(4):1684–1738, 2019. doi: 10.1086/701680.
- Tatyana Deryugina, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif. The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, 109(12):4178–4219, December 2019. doi: 10.1257/aer.20180279. URL <https://www.aeaweb.org/articles?id=10.1257/aer.20180279>.
- Steven J. Dundas and Roger H. von Haefen. The effects of weather on recreational fishing demand and adaptation: Implications for a changing climate. *Journal of the Association of Environmental and Resource Economists*, 7(2), 2020. doi: 10.1086/706343.
- Eric English, Roger H. von Haefen, Joseph Herriges, Christopher Leggett, Frank Lupi, Kenneth McConnell, Michael Welsh, Adam Domanski, and Norman Meade. Estimating the value of lost recreation days from the Deepwater Horizon oil spill. *Journal of Environmental Economics and Management*, 91:26–45, 2018.
- Jacob Gellman, Margaret Walls, and Matthew Wibbenmeyer. Welfare losses from wildfire smoke: Evidence from daily outdoor recreation data. *Journal of Environmental Economics and Management*, 132:103166, 2025. doi: 10.1016/j.jeem.2025.103166.
- Ariel Goldszmidt, John A. List, Robert D. Metcalfe, Ian Muir, V. Kerry Smith, and Jenny Wang. The value of time in the united states: Estimates from nationwide natural field experiments. Working Paper 28208, National Bureau of Economic Research, 2020. URL <https://www.sciencedirect.com/science/article/pii/S0959378024000554>.

- GreenInfo Network. California Protected Areas Database (CPAD), Version 2025b. [https://calands.org/wp-content/uploads/2025/12/CPAD\\_2025b\\_Database\\_Manual.pdf](https://calands.org/wp-content/uploads/2025/12/CPAD_2025b_Database_Manual.pdf), December 2025. URL [https://calands.org/wp-content/uploads/2025/12/CPAD\\_2025b\\_Database\\_Manual.pdf](https://calands.org/wp-content/uploads/2025/12/CPAD_2025b_Database_Manual.pdf). Database Manual available at [https://calands.org/wp-content/uploads/2025/12/CPAD\\_2025b\\_Database\\_Manual.pdf](https://calands.org/wp-content/uploads/2025/12/CPAD_2025b_Database_Manual.pdf).
- Paulo Guimarães, Octavio Figueiredo, and Douglas Woodward. A tractable approach to the firm location decision problem. *Review of Economics and Statistics*, 85(1):201–204, 2003.
- Jerry A. Hausman, Gregory K. Leonard, and Daniel McFadden. A utility-consistent, combined discrete choice and count data model: Assessing recreational use losses due to natural resource damage. *Journal of Public Economics*, 56(1):1–30, 1995.
- Xiaobo He, Zijun Luo, and Junjie Zhang. The impact of air pollution on movie theater admissions. *Journal of Environmental Economics and Management*, 112:102626, 2022. ISSN 0095-0696. doi: <https://doi.org/10.1016/j.jeem.2022.102626>. URL <https://www.sciencedirect.com/science/article/pii/S0095069622000134>.
- Sam Heft-Neal, Carlos F. Gould, Marissa L. Childs, Mathew V. Kiang, Kari C. Nadeau, Mark Duggan, Eran Bendavid, and Marshall Burke. Emergency department visits respond nonlinearly to wildfire smoke. *Proceedings of the National Academy of Sciences*, 120(39):e2302409120, 2023. doi: 10.1073/pnas.2302409120. URL <https://www.pnas.org/doi/abs/10.1073/pnas.2302409120>.
- Yujie Hu, Changzhen Wang, Ruiyang Li, and Fahui Wang. Estimating a large drive time matrix between zip codes in the united states: A differential sampling approach. *Journal of Transport Geography*, 86:102770, 2020. ISSN 0966-6923. doi: <https://doi.org/10.1016/j.jtrangeo.2020.102770>. URL <https://www.sciencedirect.com/science/article/pii/S0966692319310051>.
- Gabriel E. Kreindler and Yuhei Miyauchi. Measuring commuting and economic activity inside cities with cell phone records. *Review of Economics and Statistics*, 105(4):899–909, 2023.
- Fabian Lange. The role of education in complex health decisions: Evidence from cancer screening. *Journal of Health Economics*, 30(1):43–54, 2011. ISSN 0167-6296. doi: <https://doi.org/10.1016/j.jhealeco.2010.08.008>. URL <https://www.sciencedirect.com/science/article/pii/S0167629610001177>.
- Kenneth E. McConnell. Chapter 15 the economics of outdoor recreation. In *Handbook of Natural Resource and Energy Economics*, volume 2 of *Handbook of Natural Resource and Energy*

- Economics*, pages 677–722. Elsevier, 1985. doi: [https://doi.org/10.1016/S1573-4439\(85\)80022-1](https://doi.org/10.1016/S1573-4439(85)80022-1). URL <https://www.sciencedirect.com/science/article/pii/S1573443985800221>.
- Daniel McFadden. Modeling the choice of residential location. In Anders Karlqvist, Lars Lundqvist, Folke Snickars, and Jörgen W. Weibull, editors, *Spatial Interaction Theory and Planning Models*, pages 75–96. North-Holland, Amsterdam, 1978.
- Nolan Miller, David Molitor, and Eric Zou. The nonlinear effects of air pollution on health: Evidence from wildfire smoke, 9 2024.
- Yuhei Miyauchi, Kentaro Nakajima, and Stephen J. Redding. The economics of spatial mobility: Theory and evidence using smartphone data. *The Quarterly Journal of Economics*, 140(4): 2507–2570, 2025.
- Matthew Neidell. Information, avoidance behavior, and health. *Journal of Human Resources*, 44: 450–478, 2009. ISSN 0022-166X. doi: 10.3368/jhr.44.2.450.
- Minghao Qiu, Jessica Li, Carlos F Gould, Renzhi Jing, Makoto Kelp, Marissa Childs, Mathew Kiang, Sam Heft-Neal, Noah Diffenbaugh, and Marshall Burke. Wildfire smoke exposure and mortality burden in the us under climate change. *Nature*, 9 2025. ISSN 0028-0836. doi: 10.1038/s41586-025-09611-w.
- Esteban Rossi-Hansberg. A spatial theory of trade. *American Economic Review*, 100(4):1464–1491, 2010.
- Soodeh Saberian, Anthony Heyes, and Nicholas Rivers. Alerts work! air quality warnings and cycling. *Resource and Energy Economics*, 49:165–185, 2017. ISSN 0928-7655. doi: <https://doi.org/10.1016/j.reseneeco.2017.05.004>. URL <https://www.sciencedirect.com/science/article/pii/S0928765516300707>.
- Wolfram Schlenker. Daily weather data - continental usa, February 2024. URL <https://doi.org/10.5281/zenodo.10625288>.
- Kenneth A. Small and Harvey S. Rosen. Applied welfare economics with discrete choice models. *Econometrica*, 49(1):105–130, 1981.
- Xinhui Sun. Air pollution and economic activity: Evidence from foot traffic patterns. Job market paper, 2025.
- Kenneth E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, Cambridge, 2nd edition, 2009.

- U.S. Bureau of Labor Statistics. Quarterly census of employment and wages. <https://www.bls.gov/cew/>, 2018. U.S. Department of Labor. County-level annual average wages, 2018. Accessed [DATE].
- U.S. Department of Transportation, Federal Highway Administration. 2017 national household travel survey. <https://nhts.ornl.gov>, 2017.
- U.S. Environmental Protection Agency. Reconsideration of the national ambient air quality standards for particulate matter; final rule. Technical Report 89 FR 16202, Federal Register, 2024. URL <https://www.epa.gov/system/files/documents/2024-02/pm-naaqs-air-quality-index-fact-sheet.pdf>. March 6, 2024.
- U.S. EPA. Handbook on valuing changes in time use induced by regulatory requirements and other EPA actions. Technical Report EPA 236-B-15-001, U.S. EPA, National Center for Environmental Economics, Washington, DC, 2020.
- Joshua Graff Zivin and Matthew Neidell. Days of haze: Environmental information disclosure and intertemporal avoidance behavior. *Journal of Environmental Economics and Management*, 58 (2):119–128, 2009. doi: 10.1016/j.jeem.2009.03.001.

# A Appendix

## Figures

Figure A1: California Wildfire Smoke and Fires, August 8, 2018

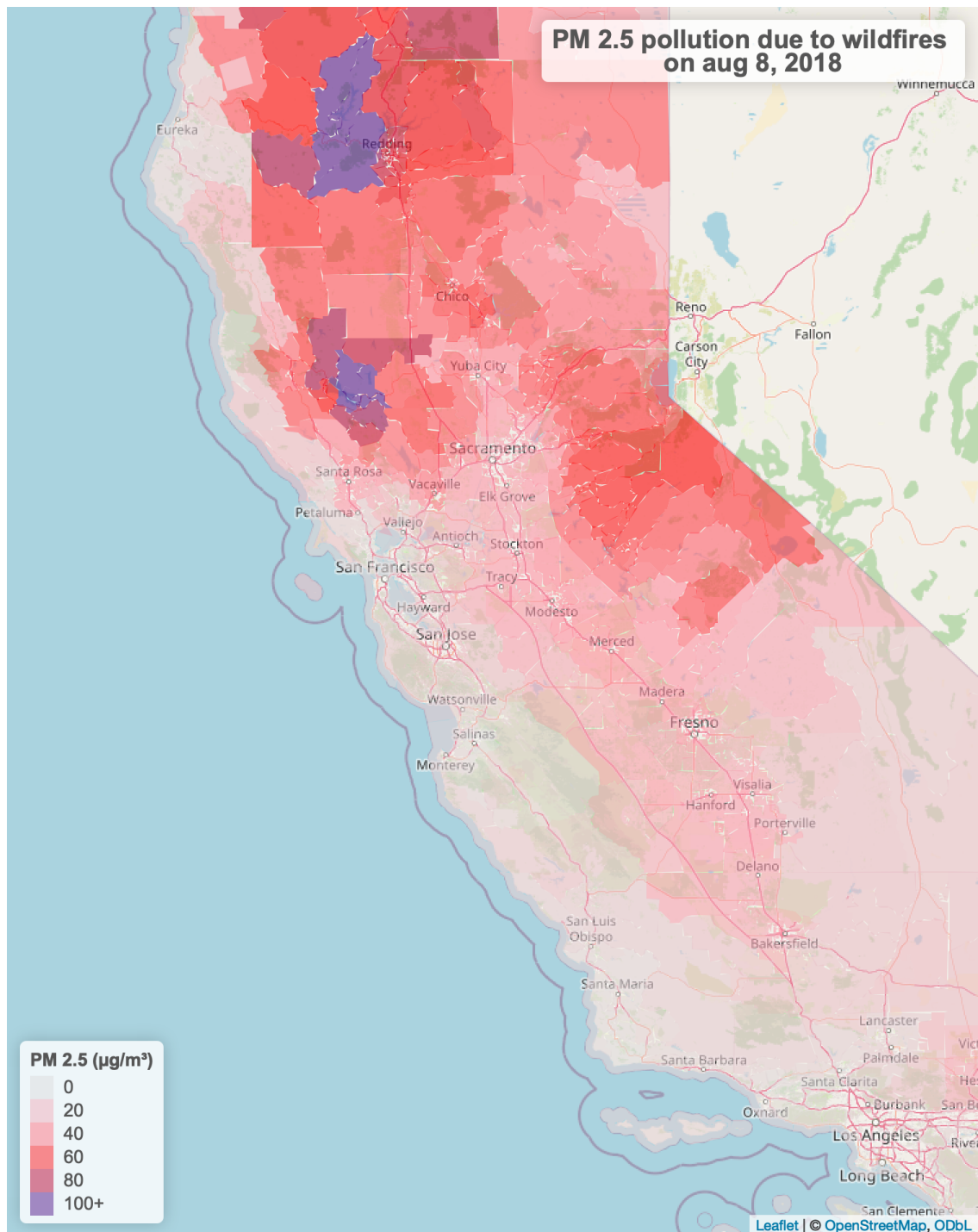


Figure A2: California Wildfire Smoke and Fires, November 15, 2018

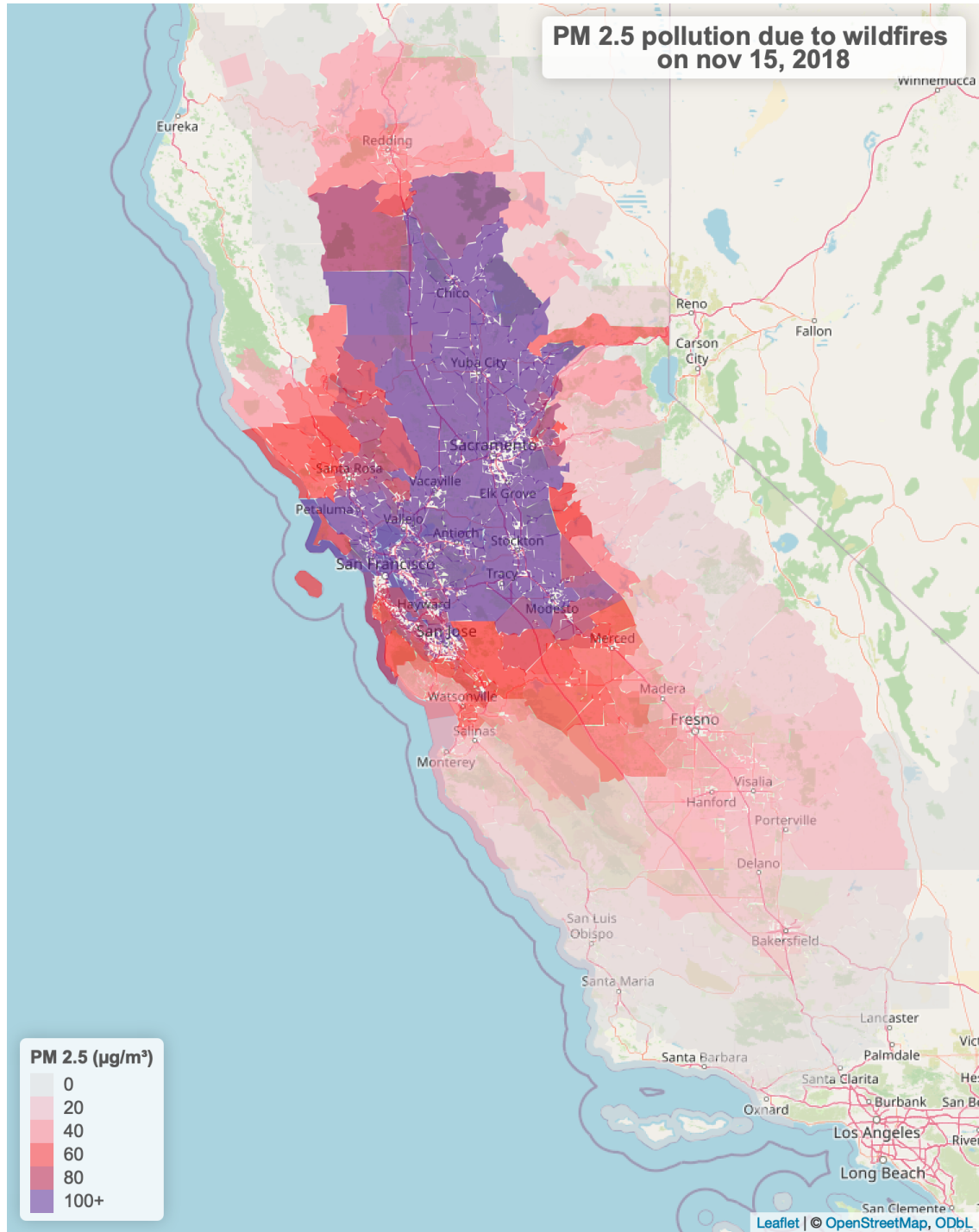


Figure A3: CPAD + CCED Area in the Bay Area

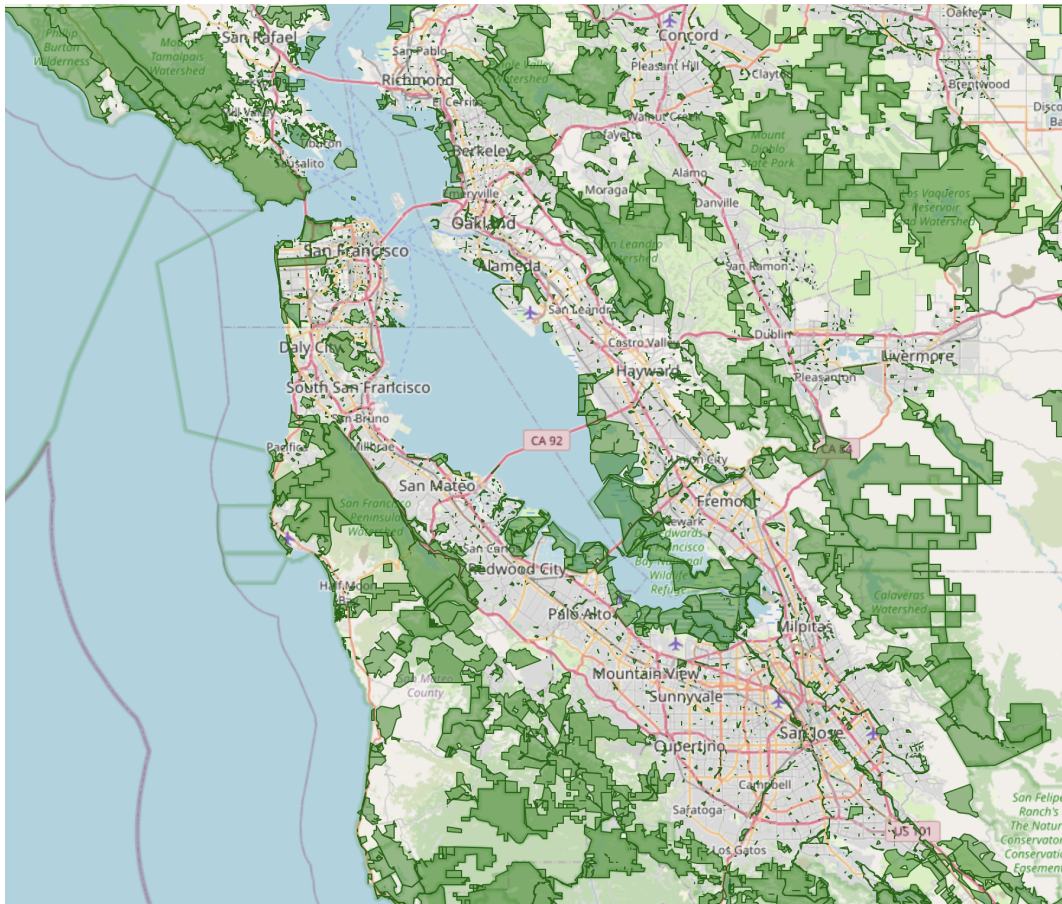


Figure A4: Bounding Box for Device Home Locations

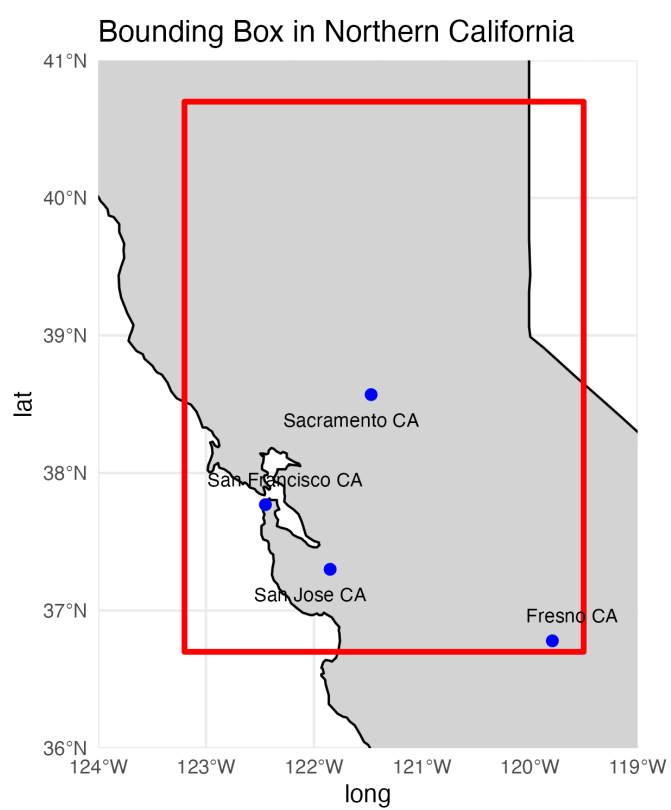


Figure A5: ZIP codes excluded because of active fires, 2018

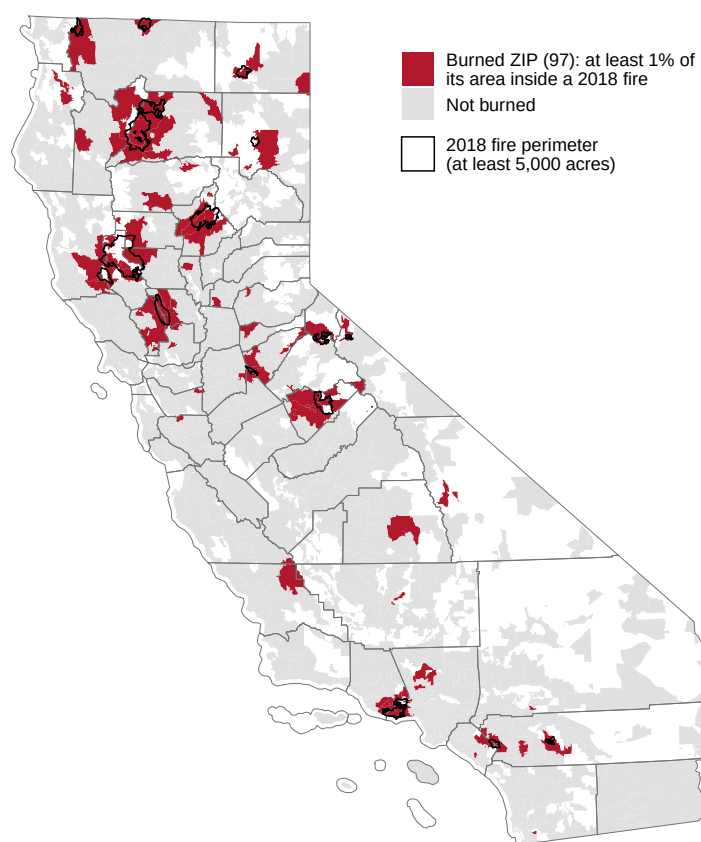


Figure A6: Distribution of # of trips by device-day

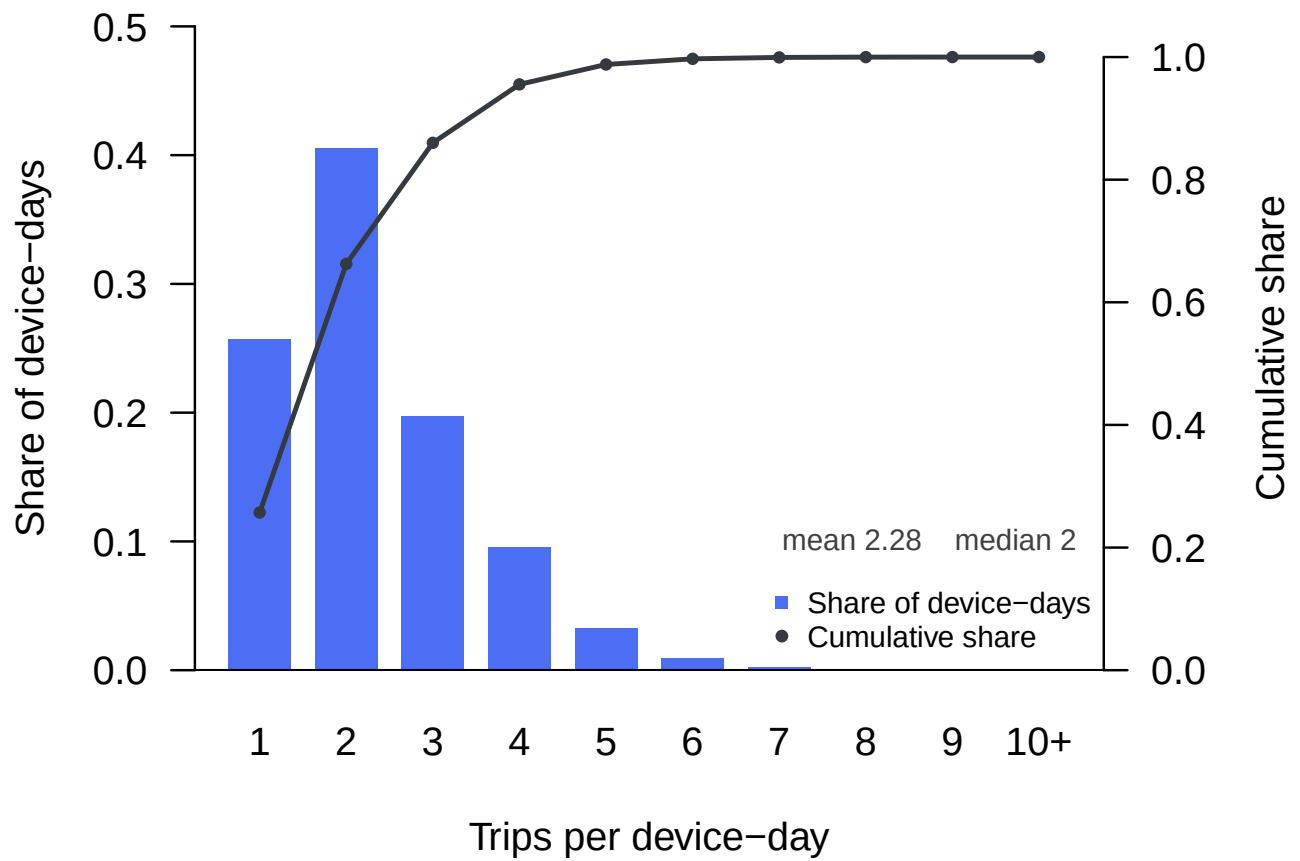
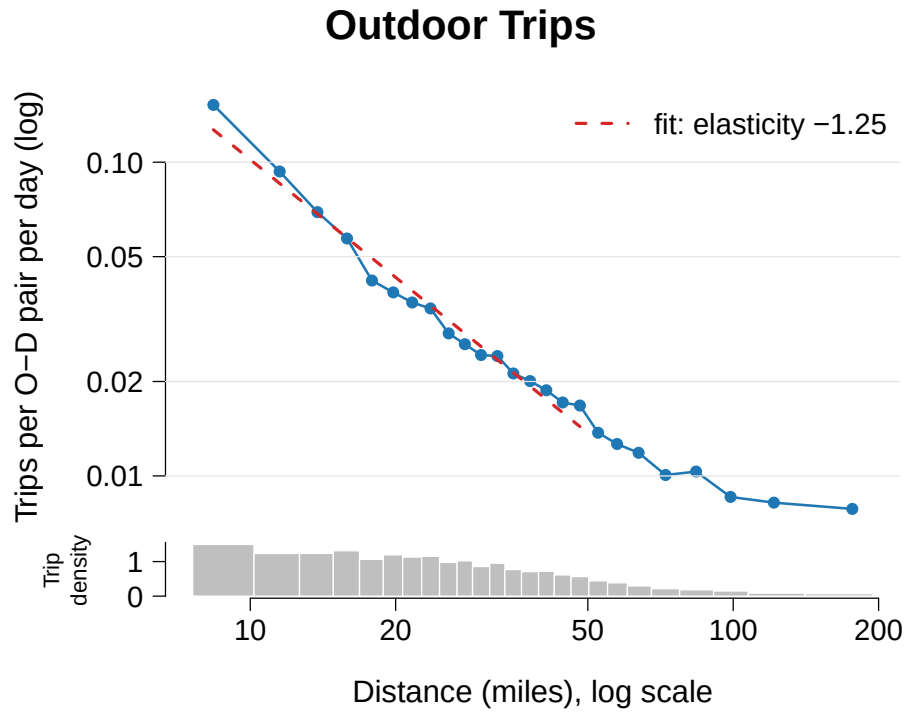
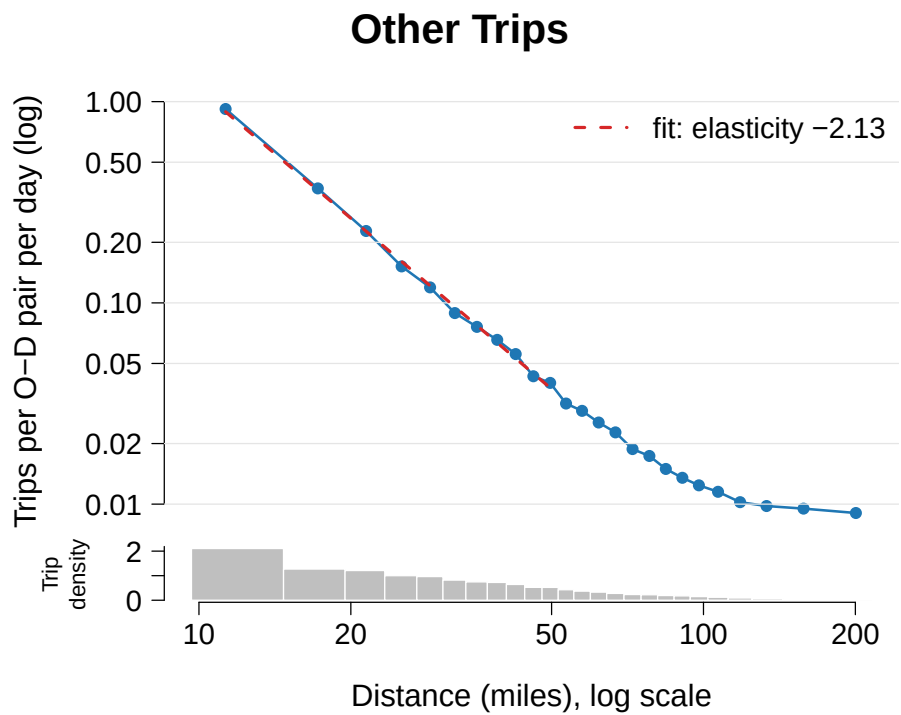


Figure A7: Trips decay with distance



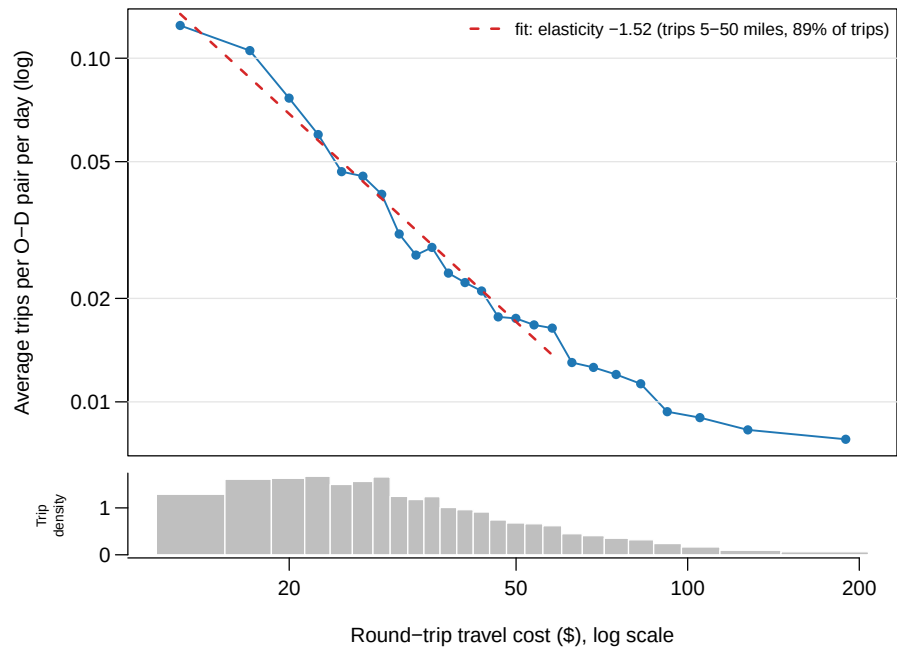
(a) Outdoor trips



(b) Other trips

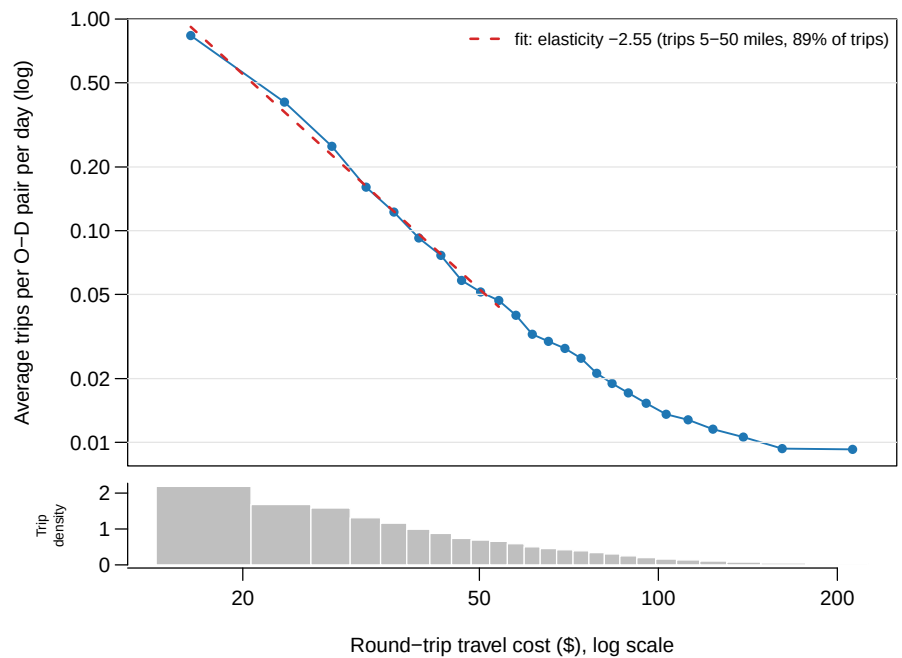
Figure A8: Trips decay with travel cost

### Outdoor Trips



(a) Outdoor trips

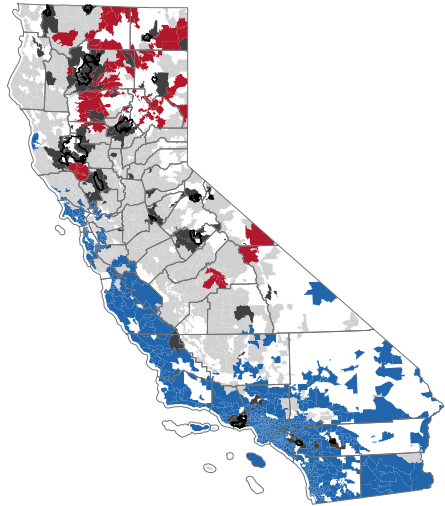
### Other Trips



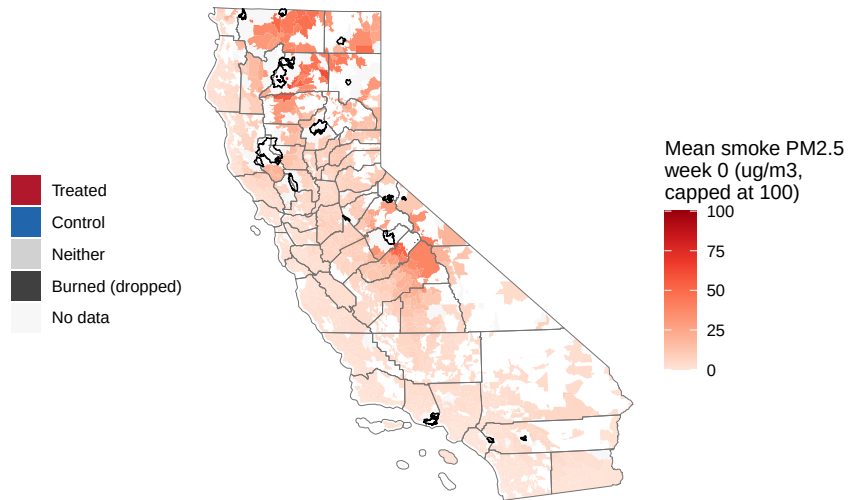
(b) Other trips

Figure A9: Treated and control ZIP codes in the event study, two onset weeks

**Cohort Jul 23: 68 treated, 849 controls**

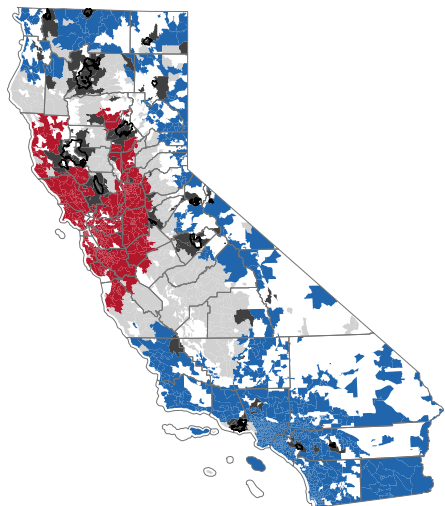


**Smoke, week of Jul 23**

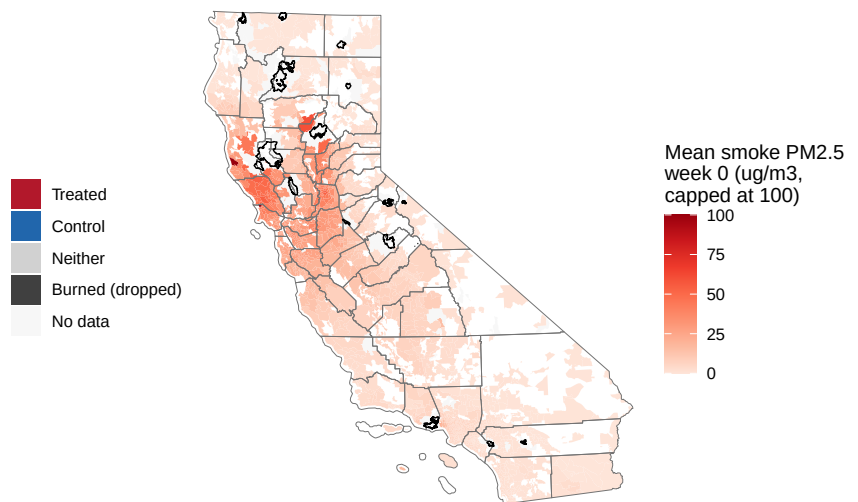


(a) Onset week of July 23, 2018

**Cohort Nov 05: 487 treated, 816 controls**



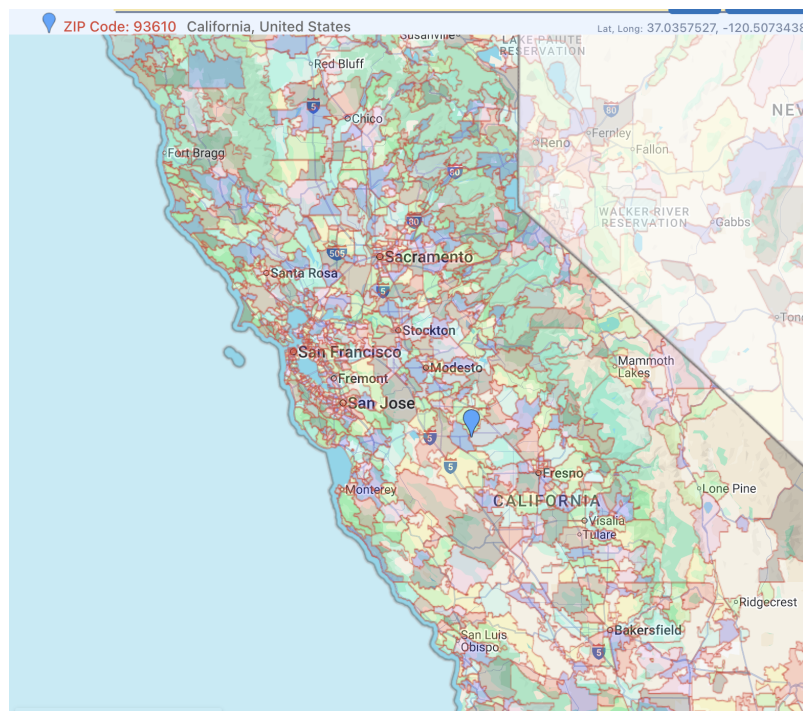
**Smoke, week of Nov 05**



(b) Onset week of November 5, 2018

*Notes:* Left panels: ZIP codes treated in the stacked event study of Figures 6 and 2 (smoke  $\text{PM}_{2.5}$  above  $35\mu\text{g}/\text{m}^3$  on at least two days of the onset week, after three weeks averaging below  $10\mu\text{g}/\text{m}^3$ ), control ZIP codes (weekly mean smoke below  $10\mu\text{g}/\text{m}^3$  in every week of the window), ZIP codes that are neither, and ZIP codes excluded because at least 1% of their area lies inside a 2018 fire perimeter. Right panels: mean smoke  $\text{PM}_{2.5}$  in the onset week, capped at  $100\mu\text{g}/\text{m}^3$ , with 2018 fire perimeters outlined.

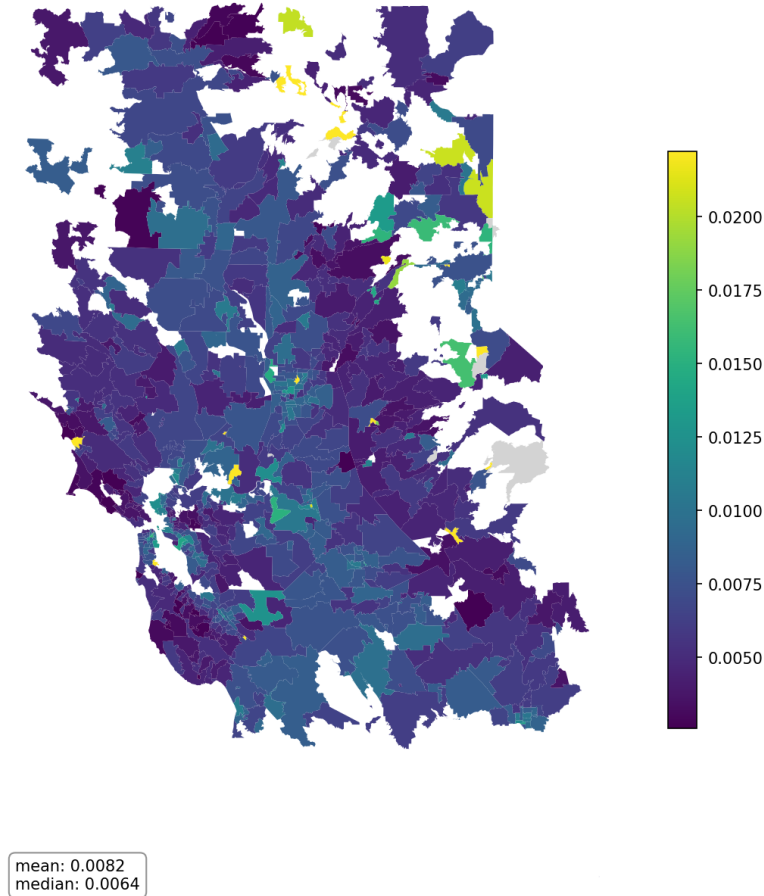
Figure A10: California ZIP Codes



Source: <https://www.randymajors.org/ZIPcodegmap>

Figure A11: Averaged Observed Devices

Daily-active devices per adult by home ZIP  
(NorCal bbox, pop=adult 18+; cumulative coverage is ~22x higher)



Over the time period, average number of devices observed daily scaled by ZIP adult population from the Census.

Figure A12: Trip Detection Methodology

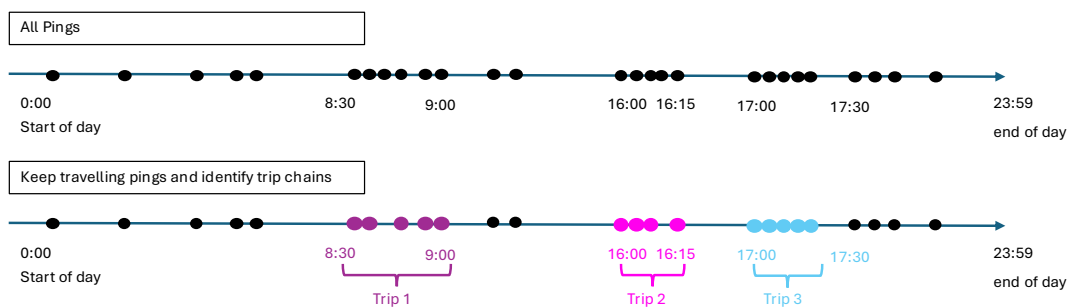


Figure A13: Origin Destination Methodology for Outdoor Trips: An Example

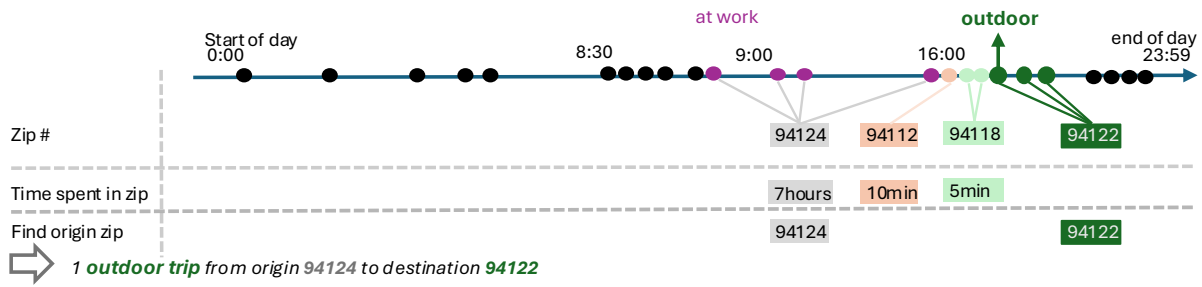
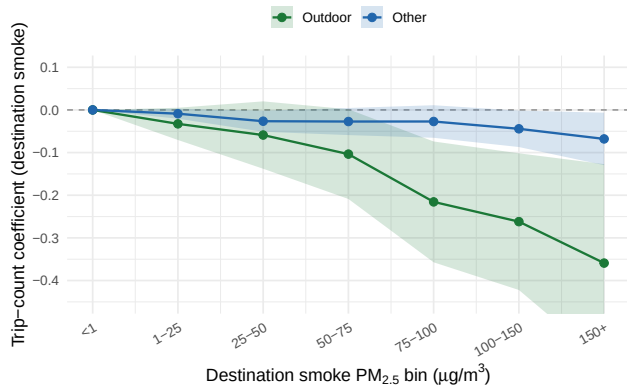
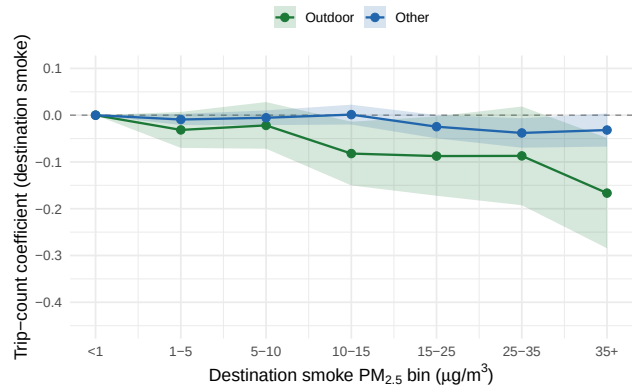


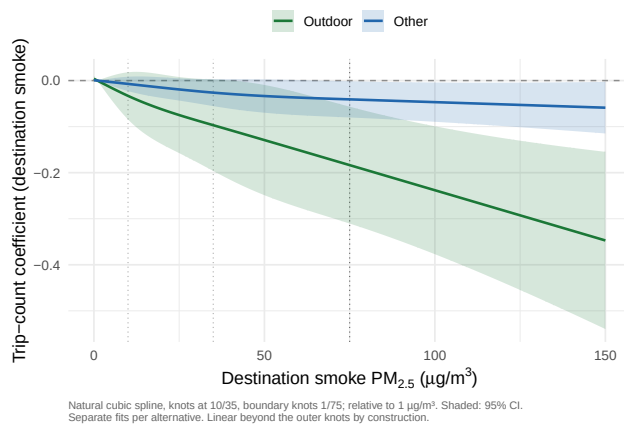
Figure A14: Varying the smoke partition



(a) Up to 35µg/m<sup>3</sup>

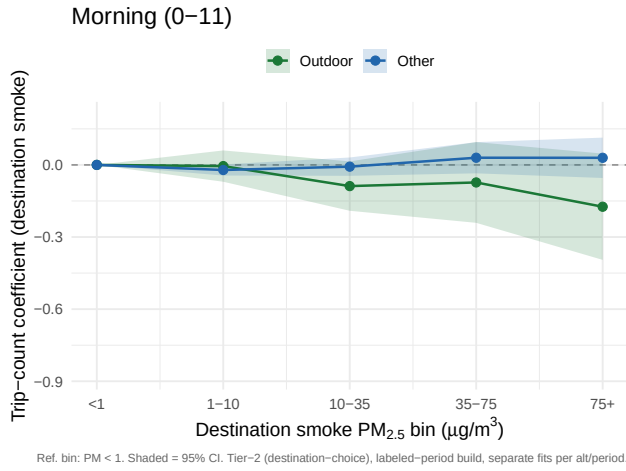


(b) Up to 150µg/m<sup>3</sup>

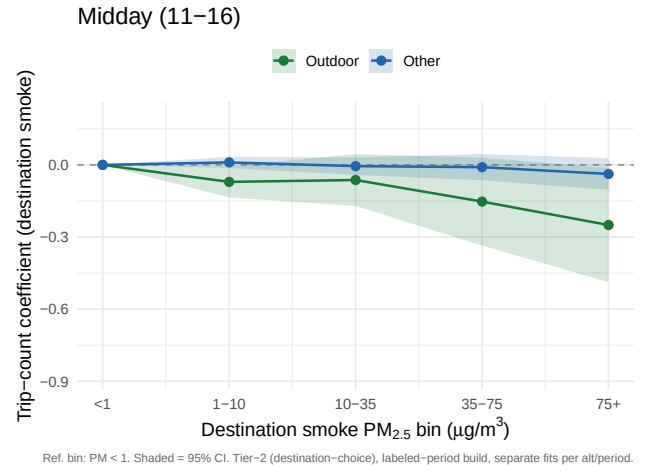


(c) Spline

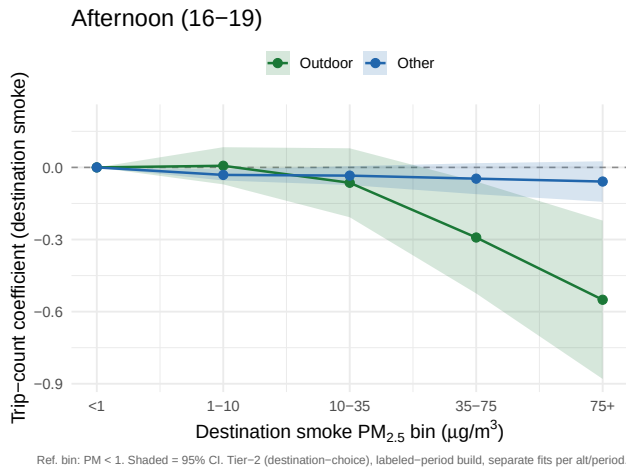
Figure A15: 4 period model - coefficients per period



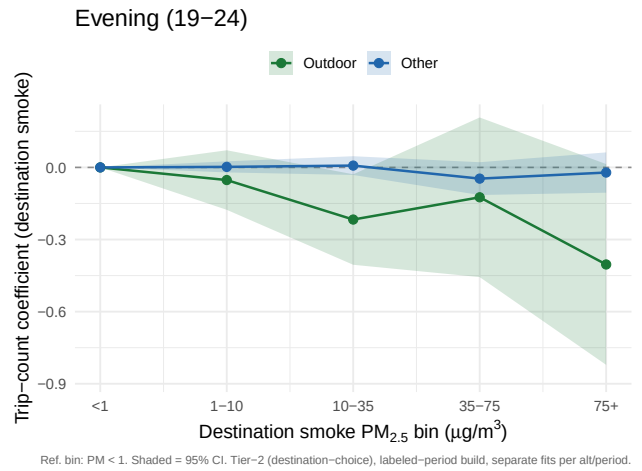
(a) Morning



(b) Midday

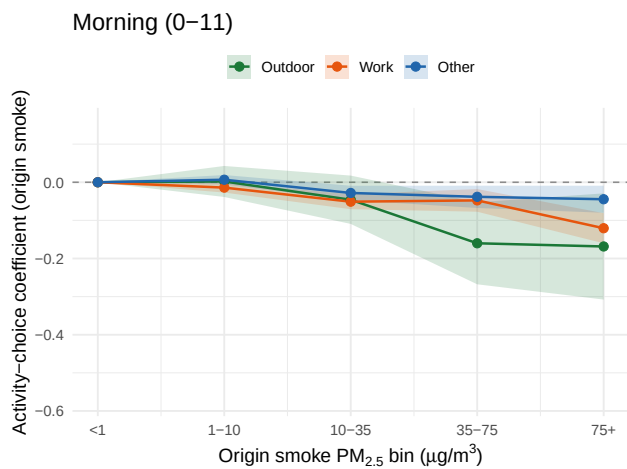


(c) Afternoon

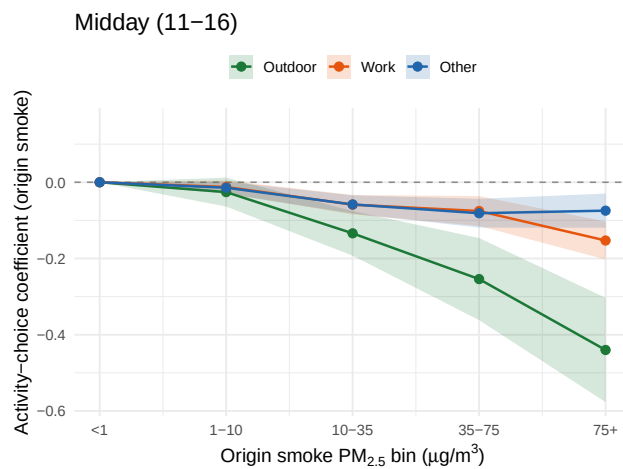


(d) Evening

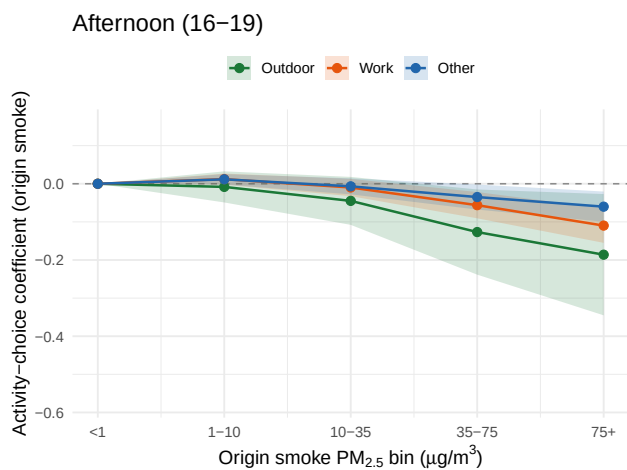
Figure A16: 4 period model - coefficients per period, tier 1



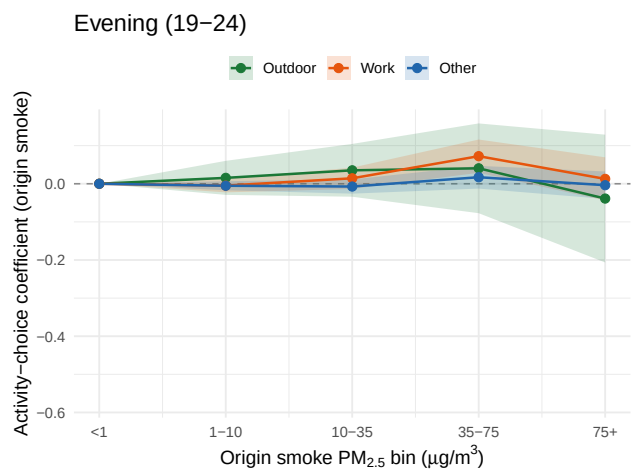
(a) Morning



(b) Midday

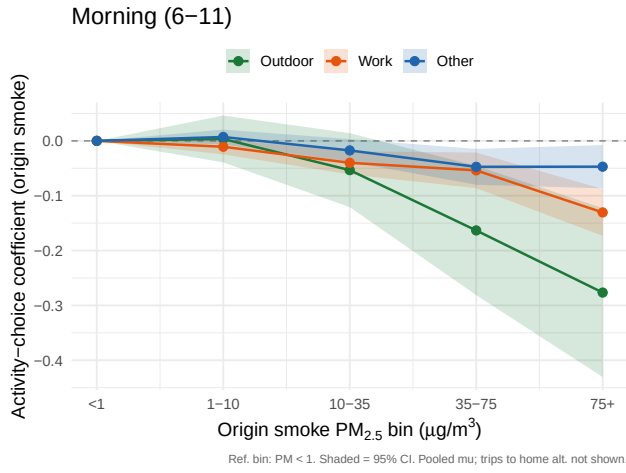


(c) Afternoon

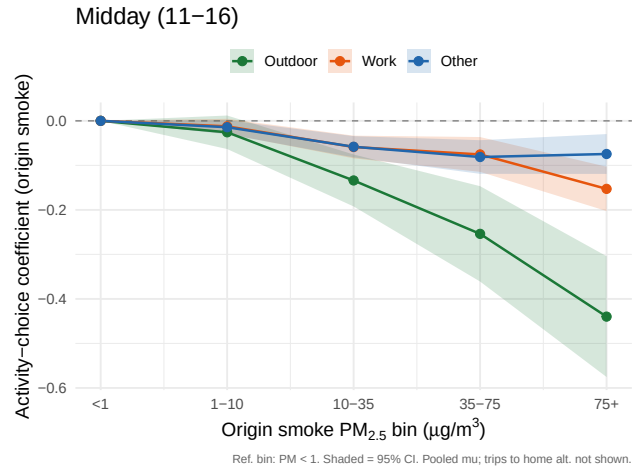


(d) Evening

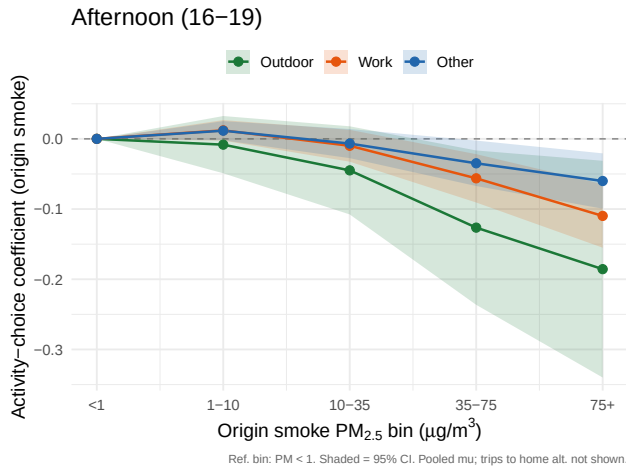
Figure A17: 5 period model - coefficients per period, tier 1 - Activity choice



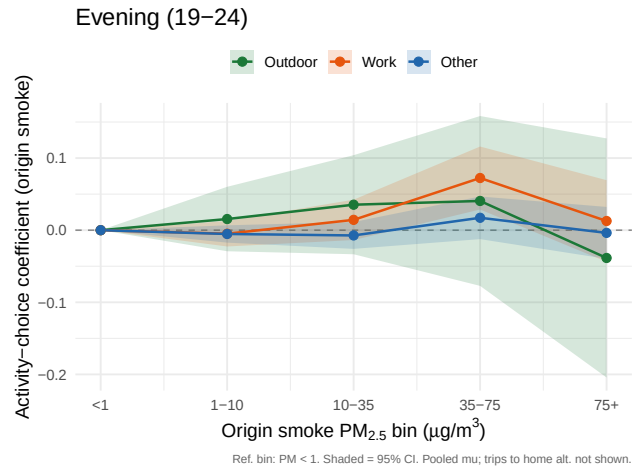
(a) Morning



(b) Midday

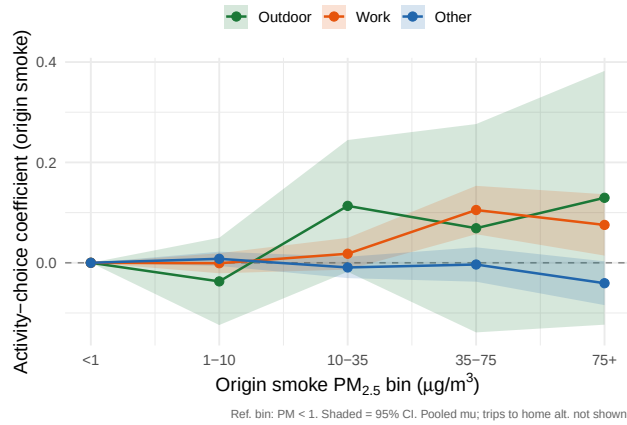


(c) Afternoon



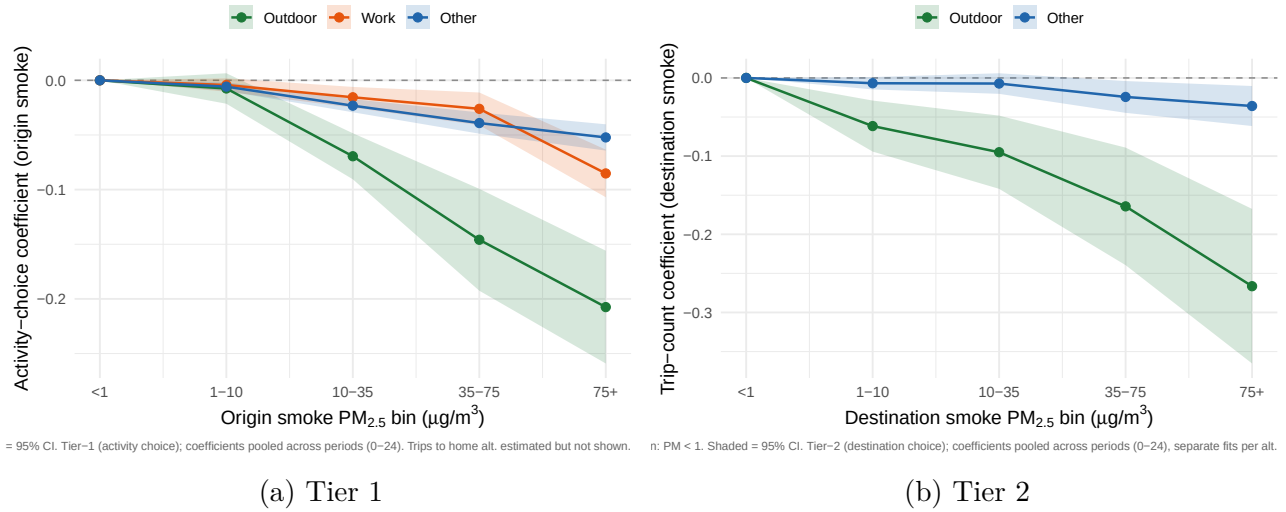
(d) Evening

**Night (0–6)**



(e) Night

Figure A18: Full sample - no workplace restriction



(a) Tier 1

(b) Tier 2

Figure A19: Probability of a Spare-the-Air alert by forecast ozone AQI bin (date-level). Bay area transitions sharply at AQI 100; Sacramento at AQI 126. Dot size is proportional to the number of dates per 10-AQI bin; dashed lines mark the region cutoffs.

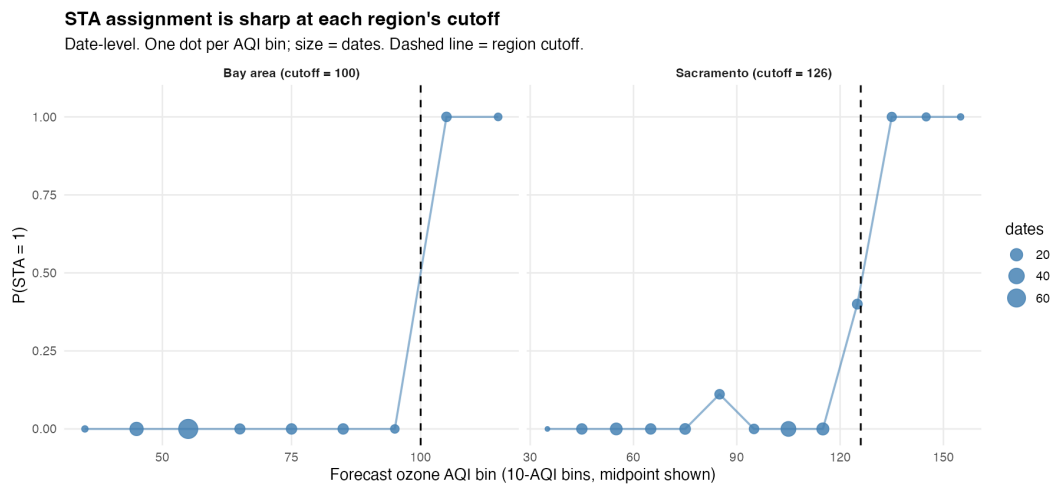
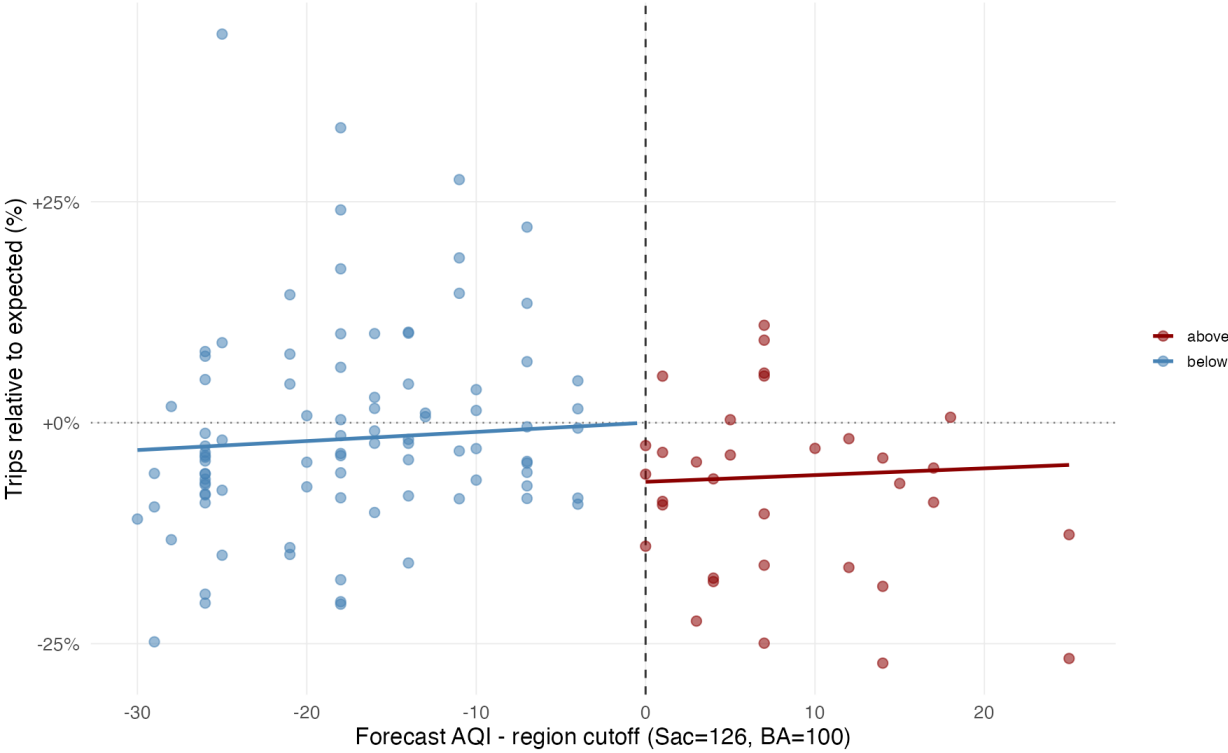


Figure A20: Coefficient on **above** (RD jump on trips at the STA cutoff) across ZIP-level terciles of median household income, share with a bachelor's degree or higher, and share Hispanic. Error bars are 95% CIs from zip-clustered standard errors. Estimated by FE Poisson on raw trips with  $|\text{dist}| \leq 20$  and zip + dow + month + week fixed effects.



## Tables

Table A1: Descriptive Statistics on Trips

|                                       | Mean  | Median | Min   | Max    |
|---------------------------------------|-------|--------|-------|--------|
| Trip travel time (min)                | 38.28 | 31.53  | 1.97  | 355.32 |
| Trip distance (road km)               | 23.23 | 17.16  | 0.02  | 297.72 |
| Predicted PM2.5 (per zip-day, ug/m3)  | 5.18  | 0.00   | 0.00  | 294.48 |
| Average temp at destination (Celsius) | 19.62 | 19.72  | -2.25 | 36.85  |
| Precipitation at destination (mm)     | 0.34  | 0.05   | 0.00  | 52.43  |

Table A2: Descriptive Statistics on Workplace Detection

|                                         | Devices | Share  |
|-----------------------------------------|---------|--------|
| Workplace identified                    | 98,054  | 36.5%  |
| No workplace identified                 | 170,921 | 63.5%  |
| Devices with a Northern California home | 268,975 | 100.0% |

*Notes:* Devices with a home in Northern California, observed on weekdays in September 2018. Workplaces are identified based on where devices are between 8am and 6pm on weekdays (work hours) only. The workplace is identified as the place where it spends the most time during work hours. If the device is mostly at home during work hours, workplace is identified as the second most visited place, which must be for at least 4 hours a day on at least 3 days. If no place matches this criteria, then the device does not have a workplace. This includes people who stay home during work hours and people who move between many places, such as delivery drivers.

Table A3: Observation density over air pollution intensity

| Smoke ( $\mu\text{g}/\text{m}^3$ ) | # pings |          | # hours with $\geq 1$ ping |        | N         |
|------------------------------------|---------|----------|----------------------------|--------|-----------|
|                                    | Mean    | SD       | Mean                       | SD     |           |
| < 1                                | 98.81   | (288.87) | 7.20                       | (5.41) | 3,361,334 |
| 1–10                               | 91.20   | (238.50) | 6.93                       | (5.29) | 1,527,697 |
| 10–35                              | 85.56   | (231.47) | 6.74                       | (5.23) | 454,248   |
| 35–75                              | 94.11   | (352.50) | 7.17                       | (5.65) | 276,078   |
| $\geq 75$                          | 94.27   | (310.58) | 7.22                       | (5.71) | 173,303   |
| All                                | 95.40   | (276.50) | 7.09                       | (5.39) | 5,792,660 |

Table A4: Summary Statistics for Device-Day Trip Sample

| Variable                                   | Value  |
|--------------------------------------------|--------|
| Daily average number of travelling devices | 21,201 |
| Daily average trips per travelling device  | 2.42   |

Table A5: Activity shares by period: smartphone data vs. 2017 NHTS

| Activity (%)       | Morning<br>(0–11) |      | Midday<br>(11–16) |      | Afternoon<br>(16–19) |      | Evening<br>(19–24) |      |
|--------------------|-------------------|------|-------------------|------|----------------------|------|--------------------|------|
|                    | Smartphone data   | NHTS | Smartphone data   | NHTS | Smartphone data      | NHTS | Smartphone data    | NHTS |
| Stayed home        | 18.0              | 21.4 | 8.6               | 14.5 | 20.6                 | 25.0 | 46.8               | 56.7 |
| Work               | 54.9              | 59.6 | 62.4              | 65.0 | 36.5                 | 46.4 | 10.0               | 13.3 |
| Outdoor recreation | 0.9               | 3.8  | 1.1               | 4.0  | 1.5                  | 5.8  | 1.3                | 7.2  |
| On their way home  | 0.0               | 0.2  | 2.9               | 4.7  | 10.9                 | 11.1 | 16.3               | 15.9 |
| Other              | 26.2              | 15.0 | 25.0              | 11.7 | 30.6                 | 11.6 | 25.6               | 6.9  |

*Notes:* Share of person-periods in each activity, in each of the four periods of the day (0–11h, 11–16h, 16–19h, 19–24h). Each period is assigned one activity, with priority work, stayed home, outdoor recreation, other. A period counts as *work* or *outdoor recreation* if the person is at that destination at any time during it, as *stayed home* only if the person is at home for the whole period, and as *on their way home* if it ends with a return home. Columns may not sum to 100 due to rounding.

*Smartphone data:* tier-1 estimation sample, California devices with an identified workplace, July–November 2018; a device-day enters if the device is observed in at least five distinct hours and comes within 500 m of home that day. Work and home are geofences (the device’s identified workplace; 80 m around home); outdoor recreation is a visit to a park.

*NHTS:* 25,546 person-days observations. Crosssection. 2017 National Household Travel Survey, workers in California households, travel days April 2016–May 2017, weighted by the person weight. Days with no trips are included and counted as stayed home. Each respondent’s day is reconstructed from the trip diary, and the same activity rules are applied. NHTS outdoor recreation covers all recreation and exercise trips (including gyms and entertainment venues), so it is an upper bound for our park-based measure.

Table A6: Week day versus week end trip behavior

| Activity                  | Morning<br>(0–11) | Midday<br>(11–16) | Afternoon<br>(16–19) | Evening<br>(19–24) |
|---------------------------|-------------------|-------------------|----------------------|--------------------|
| <i>Weekdays (Mon–Fri)</i> |                   |                   |                      |                    |
| Stayed home               | 18.0              | 8.6               | 20.6                 | 46.8               |
| Work                      | 54.9              | 62.4              | 36.5                 | 10.0               |
| Outdoor recreation        | 0.9               | 1.1               | 1.5                  | 1.3                |
| On their way home         | 0.0               | 2.9               | 10.9                 | 16.3               |
| Other                     | 26.2              | 25.0              | 30.6                 | 25.6               |
| <i>Weekends (Sat–Sun)</i> |                   |                   |                      |                    |
| Stayed home               | 40.4              | 26.0              | 34.6                 | 47.8               |
| Work                      | 13.3              | 14.2              | 9.1                  | 5.6                |
| Outdoor recreation        | 2.3               | 3.9               | 2.5                  | 1.4                |
| On their way home         | 0.0               | 6.2               | 8.8                  | 15.0               |
| Other                     | 44.1              | 49.6              | 45.0                 | 30.2               |

Table A7: Descriptive Statistics on Daily Trips

|                                     | Outdoor | Indoor    |
|-------------------------------------|---------|-----------|
| Total trips observed                | 141,664 | 1,544,772 |
| Distinct origin ZIPs                | 715     | 738       |
| Distinct destination ZIPs           | 874     | 1,152     |
| <i>Trips per origin ZIP per day</i> |         |           |
| Mean                                | 1.29    | 13.68     |
| Median                              | 0.00    | 6.00      |
| 90th percentile                     | 4       | 37        |
| Share of ZIP-days with 0 trips      | 51.3%   | 23.0%     |

Table A8: Descriptive Statistics on Daily ZIP device activity

|                                       | Stay home | Go outdoor | Leave (other) | Go to work |
|---------------------------------------|-----------|------------|---------------|------------|
| <i>Devices per origin ZIP per day</i> |           |            |               |            |
| Mean                                  | 2.99      | 1.27       | 14.85         | 18.83      |
| Median                                | 1         | 0          | 8             | 8          |
| 90th percentile                       | 8         | 4          | 38            | 53         |
| Share of ZIP-days with 0              | 35.2%     | 52.6%      | 10.5%         | 11.4%      |

Table A9: Travel time elasticities: comparison with the literature

|                                               | Setting                     | Trips and method                     | Workdays      | Days off | Pooled         |
|-----------------------------------------------|-----------------------------|--------------------------------------|---------------|----------|----------------|
| <i>This paper</i>                             |                             |                                      |               |          |                |
| Outdoor                                       | California, smartphone      | Destination choice, PPML             | −3.10         | −2.74    | −2.93          |
| Other                                         | California, smartphone      | Destination choice, PPML             | −3.61         | −3.45    | −3.55          |
| <i>Comparators</i>                            |                             |                                      |               |          |                |
| <a href="#">Miyauchi et al. (2025)</a>        | Tokyo, smartphone           | Daily itineraries, structural        | −2.92         | −1.99    |                |
| <a href="#">Davis et al. (2019)</a>           | New York, Yelp              | Restaurant visits, conditional logit |               |          | −0.94 to −1.95 |
| <a href="#">Kreindler and Miyauchi (2023)</a> | Dhaka / Colombo, cell phone | Commuting, PPML                      | −2.44 / −2.19 |          |                |

*Notes:* Elasticities of trip probability with respect to log travel time. Estimates for this paper come from PPML with origin  $\times$  day  $\times$  period and destination  $\times$  period fixed effects. Workdays are Monday–Friday excluding public holidays; days off are weekends and public holidays. [Miyauchi et al. \(2025\)](#): p. 2542 for workdays and non-workdays (Table 2). [Davis et al. \(2019\)](#): range of coefficients on log travel time from home and work across modes and demographic groups (Table 2, columns 4–6); conditional logit and PPML with origin fixed effects yield identical coefficients. [Kreindler and Miyauchi \(2023\)](#): PPML with origin and destination fixed effects (Table 1).

Table A10: Tier 2 : Destination smoke coefficients with origin  $\times$  destination fixed effects

| Smoke PM <sub>2.5</sub>                            | Outdoor                | Other                |
|----------------------------------------------------|------------------------|----------------------|
| 1–10 $\mu\text{g}/\text{m}^3$                      | −0.0379*<br>(0.0172)   | −0.0104*<br>(0.0052) |
| 10–35 $\mu\text{g}/\text{m}^3$                     | −0.0932**<br>(0.0284)  | −0.0114<br>(0.0084)  |
| 35–75 $\mu\text{g}/\text{m}^3$                     | −0.1647***<br>(0.0470) | −0.0186<br>(0.0135)  |
| 75+ $\mu\text{g}/\text{m}^3$                       | −0.2962***<br>(0.0612) | −0.0254<br>(0.0161)  |
| log(travel cost)                                   | —                      | —                    |
| Observations                                       | 3,123,231              | 30,586,178           |
| Pseudo $R^2$                                       | 0.203                  | 0.309                |
| Origin $\times$ destination ZIP $\times$ period FE | Yes                    | Yes                  |
| Origin ZIP $\times$ day $\times$ period FE         | Yes                    | Yes                  |

Table A11: Tier 1 : Estimation Per Period - Coefficients for Outdoor and Work

| Smoke PM <sub>2.5</sub>                            | Morning (0–11)         | Midday (11–16)         | Afternoon (16–19)      | Evening (19–24)      |
|----------------------------------------------------|------------------------|------------------------|------------------------|----------------------|
| <i>Outdoor</i>                                     |                        |                        |                        |                      |
| 1–10 µg/m <sup>3</sup>                             | 0.0021<br>(0.0208)     | -0.0258<br>(0.0193)    | -0.0082<br>(0.0208)    | 0.0154<br>(0.0228)   |
| 10–35 µg/m <sup>3</sup>                            | -0.0459<br>(0.0324)    | -0.1340***<br>(0.0299) | -0.0449<br>(0.0321)    | 0.0352<br>(0.0353)   |
| 35–75 µg/m <sup>3</sup>                            | -0.1601**<br>(0.0549)  | -0.2540***<br>(0.0549) | -0.1267*<br>(0.0570)   | 0.0405<br>(0.0601)   |
| 75+ µg/m <sup>3</sup>                              | -0.1685*<br>(0.0711)   | -0.4400***<br>(0.0698) | -0.1861*<br>(0.0812)   | -0.0388<br>(0.0855)  |
| <i>Work</i>                                        |                        |                        |                        |                      |
| 1–10 µg/m <sup>3</sup>                             | -0.0140*<br>(0.0066)   | -0.0123<br>(0.0086)    | 0.0121<br>(0.0075)     | -0.0049<br>(0.0090)  |
| 10–35 µg/m <sup>3</sup>                            | -0.0510***<br>(0.0100) | -0.0588***<br>(0.0130) | -0.0098<br>(0.0116)    | 0.0143<br>(0.0143)   |
| 35–75 µg/m <sup>3</sup>                            | -0.0475**<br>(0.0152)  | -0.0755***<br>(0.0200) | -0.0562**<br>(0.0176)  | 0.0724**<br>(0.0223) |
| 75+ µg/m <sup>3</sup>                              | -0.1206***<br>(0.0202) | -0.1528***<br>(0.0254) | -0.1098***<br>(0.0231) | 0.0127<br>(0.0289)   |
| $\mu$ (V <sub>ot</sub> , pooled)                   |                        | 0.3235***              | (0.0745)               |                      |
| Observations                                       |                        | 1,743,419              |                        |                      |
| Pseudo $R^2$                                       |                        | 0.810                  |                        |                      |
| Alternative $\times$ day $\times$ period FE        | Yes                    | Yes                    | Yes                    | Yes                  |
| Origin ZIP $\times$ day $\times$ period FE         | Yes                    | Yes                    | Yes                    | Yes                  |
| Alternative $\times$ origin ZIP $\times$ period FE | Yes                    | Yes                    | Yes                    | Yes                  |

Table A12: Tier 1 : Estimation Per Period - Coefficients for Other and Home

| Smoke PM <sub>2.5</sub>                            | Morning (0–11)        | Midday (11–16)         | Afternoon (16–19)      | Evening (19–24)      |
|----------------------------------------------------|-----------------------|------------------------|------------------------|----------------------|
| <i>Other</i>                                       |                       |                        |                        |                      |
| 1–10 µg/m <sup>3</sup>                             | 0.0066<br>(0.0063)    | -0.0143<br>(0.0081)    | 0.0116<br>(0.0070)     | -0.0052<br>(0.0061)  |
| 10–35 µg/m <sup>3</sup>                            | -0.0281**<br>(0.0097) | -0.0582***<br>(0.0121) | -0.0068<br>(0.0107)    | -0.0073<br>(0.0095)  |
| 35–75 µg/m <sup>3</sup>                            | -0.0382*<br>(0.0149)  | -0.0813***<br>(0.0193) | -0.0349*<br>(0.0167)   | 0.0171<br>(0.0152)   |
| 75+ µg/m <sup>3</sup>                              | -0.0445*<br>(0.0181)  | -0.0745**<br>(0.0229)  | -0.0601**<br>(0.0202)  | -0.0038<br>(0.0184)  |
| <i>Trips home</i>                                  |                       |                        |                        |                      |
| 1–10 µg/m <sup>3</sup>                             | —                     | 0.0016<br>(0.0141)     | 0.0133<br>(0.0101)     | -0.0183*<br>(0.0078) |
| 10–35 µg/m <sup>3</sup>                            | —                     | 0.0057<br>(0.0215)     | -0.0164<br>(0.0152)    | -0.0190<br>(0.0120)  |
| 35–75 µg/m <sup>3</sup>                            | —                     | -0.0227<br>(0.0353)    | -0.1306***<br>(0.0230) | -0.0267<br>(0.0190)  |
| 75+ µg/m <sup>3</sup>                              | —                     | -0.0138<br>(0.0426)    | -0.1475***<br>(0.0298) | -0.0471<br>(0.0247)  |
| $\mu$ (V <sub>ot</sub> , pooled)                   |                       | 0.3235***              | (0.0745)               |                      |
| Observations                                       |                       | 1,743,419              |                        |                      |
| Pseudo $R^2$                                       |                       | 0.810                  |                        |                      |
| Alternative $\times$ day $\times$ period FE        | Yes                   | Yes                    | Yes                    | Yes                  |
| Origin ZIP $\times$ day $\times$ period FE         | Yes                   | Yes                    | Yes                    | Yes                  |
| Alternative $\times$ origin ZIP $\times$ period FE | Yes                   | Yes                    | Yes                    | Yes                  |

There are no coefficients for morning home trips because the model assumes that every person starts their days at home.

Table A13: Welfare loss by origin smoke concentration, pooled across periods

| Origin smoke<br>(µg/m <sup>3</sup> ) | Origin-<br>days | Mean loss<br>per person (\$) | Pop.-weighted<br>mean (\$) | Total<br>(\$M) |
|--------------------------------------|-----------------|------------------------------|----------------------------|----------------|
| < 1                                  | 62,233          | 0.015                        | 0.014                      | 6.31           |
| 1–10                                 | 23,758          | 0.448                        | 0.475                      | 79.78          |
| 10–35                                | 9,990           | 2.753                        | 3.036                      | 156.34         |
| 35–75                                | 4,944           | 3.717                        | 4.034                      | 130.14         |
| ≥75                                  | 2,574           | 6.578                        | 7.071                      | 132.88         |
| Total                                | 103,499         |                              |                            | 505.46         |

Table A14: Sensitivity of welfare to number of periods

| Specification               | Hours                                         | Total welfare<br>(\$M) | Per person,<br>smoky day (\$) | $\mu$         | $\mu$ 95% CI          |
|-----------------------------|-----------------------------------------------|------------------------|-------------------------------|---------------|-----------------------|
| Daily                       | one occasion/day                              | 265.6                  | 2.783                         | 0.4674        | [0.041, 0.894]        |
| 3 periods                   | 00–11   11–16   16–24                         | 675.9                  | 6.590                         | 0.2460        | [0.070, 0.422]        |
| 4 periods, version a        | 00–06   06–11   11–16   16–24                 | 552.6                  | 5.267                         | 0.2963        | [0.192, 0.400]        |
| <b>4 periods, version b</b> | <b>00–11   11–16   16–19   19–24</b>          | <b>505.5</b>           | <b>4.971</b>                  | <b>0.3235</b> | <b>[0.178, 0.469]</b> |
| 5 periods                   | 00–06   06–11   11–16   16–19   19–24         | 495.1                  | 4.756                         | 0.3250        | [0.227, 0.423]        |
| 6 periods                   | 00–06   06–08   08–11   11–16   16–19   19–24 | 469.2                  | 4.961                         | 0.2945        | [0.204, 0.385]        |

*Notes:* Annual welfare losses under alternative partitions of the day into choice occasions, with hours spanned given in the second column. The dissimilarity parameter is re-estimated under each partition; reported confidence intervals refer to  $\mu$ , not to the welfare totals. Per-person figures are per smoky day, defined as smoke above  $30 \mu\text{g}/\text{m}^3$ . Sample restricted to workers; home valuation excluded.

Table A15: Sensitivity of the welfare to numeraire choices

| Travel-cost assumption                         | Total welfare (\$M) | Change |
|------------------------------------------------|---------------------|--------|
| Baseline (VoT = $w/3$ , \$0.30/mi, round trip) | 505.5               | —      |
| VoT = $w/2$                                    | 637.6               | +26.1% |
| VoT = $w/4$                                    | 439.4               | -13.1% |
| \$0.20 per mile                                | 425.1               | -15.9% |
| \$0.40 per mile                                | 585.8               | +15.9% |

*Notes:* The table reports total annual welfare losses implied by alternative conventions for converting utility into dollars. The baseline values travel time at one-third of the individual hourly wage  $w$  following Gellman et al. (2025) and vehicle operating costs at \$0.30 per mile, both on a round-trip basis. VoT =  $w/2$  therefore denotes a *higher* value of time than baseline and  $w/4$  a lower one. Each row varies one component and holds the other at its baseline value. Estimated demand parameters are held fixed at their baseline values throughout: the rows rescale the monetary conversion and do not re-estimate travel-cost sensitivity. Because the conversion is linear in the travel-cost base, the entries are determined by the share of each component in that base (52% time, 48% distance) and are not independent of one another. The final column reports the change relative to the baseline total.

Table A16: Sensitivity of welfare to number of periods, holding  $\mu$  fixed

| Specification               | Hours                                         | Total welfare<br>(\$M) | Per person,<br>smoky day (\$) | $\mu$<br>imposed | $\mu$<br>fitted | $\mu$ 95% CI          |
|-----------------------------|-----------------------------------------------|------------------------|-------------------------------|------------------|-----------------|-----------------------|
| Daily                       | one occasion/day                              | 396.8                  | 4.192                         | 0.3047           | 0.4674          | [0.041, 0.894]        |
| 3 periods                   | 00–11   11–16   16–24                         | 553.1                  | 5.357                         | 0.3047           | 0.2460          | [0.070, 0.422]        |
| 4 periods, version a        | 00–06   06–11   11–16   16–24                 | 539.0                  | 5.120                         | 0.3047           | 0.2963          | [0.192, 0.400]        |
| <b>4 periods, version b</b> | <b>00–11   11–16   16–19   19–24</b>          | <b>532.9</b>           | <b>5.257</b>                  | <b>0.3047</b>    | <b>0.3235</b>   | <b>[0.178, 0.469]</b> |
| 5 periods                   | 00–06   06–11   11–16   16–19   19–24         | 522.6                  | 5.068                         | 0.3047           | 0.3250          | [0.227, 0.423]        |
| 6 periods                   | 00–06   06–08   08–11   11–16   16–19   19–24 | 456.0                  | 4.794                         | 0.3047           | 0.2945          | [0.204, 0.385]        |

*Notes:* The table reports total annual welfare losses under alternative partitions of the day into choice occasions, with the hours spanned by each period given in the second column. The dissimilarity parameter is imposed at  $\mu = 0.3047$  in all rows [SOURCE OF THIS VALUE], so that the partition is the only feature varying across specifications; the column headed *fitted* reports the value  $\mu$  takes when estimated freely under each partition, and the reported confidence intervals refer to that fitted value rather than to the welfare totals. Per-person figures are expressed per smoky day, defined as a day on which smoke exceeds  $30 \mu\text{g}/\text{m}^3$ . The population is restricted to workers. Home valuation is excluded throughout.

Table A17: Effect of Spare-the-Air Alerts on Outdoor Trips - Regression Discontinuity

|                               | trips                  |                      |                       |
|-------------------------------|------------------------|----------------------|-----------------------|
|                               | BW = 20<br>(1)         | BW = 25<br>(2)       | BW = 30<br>(3)        |
| Above alert threshold         | -0.0955***<br>(0.0363) | -0.0697*<br>(0.0370) | -0.0692**<br>(0.0343) |
| Forecast AQI – cutoff         | 0.0008<br>(0.0021)     | 0.0003<br>(0.0021)   | 0.0010<br>(0.0013)    |
| Mean temperature              | 0.0059<br>(0.0068)     | 0.0076<br>(0.0068)   | 0.0070<br>(0.0053)    |
| Precipitation                 | -0.0524<br>(0.0672)    | -0.0570<br>(0.0679)  | -0.1031<br>(0.0898)   |
| Above $\times$ (AQI – cutoff) | 0.0043<br>(0.0038)     | 0.0014<br>(0.0037)   | -0.0002<br>(0.0031)   |
| <i>Fit statistics</i>         |                        |                      |                       |
| Observations                  | 15,131                 | 16,647               | 21,536                |
| <i>Fixed effects</i>          |                        |                      |                       |
| zip-dow fixed effects         | Yes                    | Yes                  | Yes                   |
| week fixed effects            | Yes                    | Yes                  | Yes                   |

*Notes:* Poisson pseudo-maximum-likelihood at the ZIP-day level, pooling the Bay Area and Sacramento and restricted to  $|AQI_{rt}^f - c_r| \leq \text{BW}$ , where  $c_r$  is the basin's alert threshold ( $c_{\text{BA}} = 100$ ,  $c_{\text{Sac}} = 126$ ). Each trip is counted at the ZIP where it starts (origin of the trip). All columns include ZIP  $\times$  day-of-week and week fixed effects and control for mean temperature and precipitation.  $\exp(\hat{\beta}) - 1$  on *Above alert threshold* is the percentage change in expected trips at the threshold:  $-8.6$ ,  $-6.4$  and  $-6.7$  percent across the three columns. Standard errors, in parentheses, are clustered by basin  $\times$  date, the level at which alerts are assigned; the design rests on 33 treated basin-days. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table A18: Alternative Fixed Effects and Standard Errors - Regression Discontinuity Robustness

|                               | Baseline              |                        | No FE                  |                        | ZIP FE                 | ZIP, dow, wk FE       |
|-------------------------------|-----------------------|------------------------|------------------------|------------------------|------------------------|-----------------------|
|                               | (1)                   | (2)                    | (3)                    | (4)                    | (5)                    | (6)                   |
| Above alert threshold         | -0.0692**<br>(0.0343) | -0.0692***<br>(0.0093) | -0.0849<br>(0.0996)    | -0.0849***<br>(0.0234) | -0.1395*<br>(0.0767)   | -0.0714**<br>(0.0314) |
| Forecast AQI – cutoff         | 0.0010<br>(0.0013)    | 0.0010***<br>(0.0004)  | 0.0067<br>(0.0044)     | 0.0067***<br>(0.0008)  | 0.0060*<br>(0.0032)    | 0.0008<br>(0.0012)    |
| Above $\times$ (AQI – cutoff) | -0.0002<br>(0.0031)   | -0.0002<br>(0.0009)    | -0.0060<br>(0.0075)    | -0.0060***<br>(0.0011) | -0.0029<br>(0.0052)    | 0.0011<br>(0.0029)    |
| Mean temperature              | 0.0070<br>(0.0053)    | 0.0070***<br>(0.0024)  | -0.0897***<br>(0.0052) | -0.0897***<br>(0.0131) | -0.0390***<br>(0.0087) | 0.0057<br>(0.0049)    |
| Precipitation                 | -0.1031<br>(0.0898)   | -0.1031*<br>(0.0587)   | -2.4327***<br>(0.9424) | -2.4327<br>(1.4833)    | 0.6555*<br>(0.3698)    | -0.0803<br>(0.0681)   |
| ZIP $\times$ dow, week FE     | Yes                   | Yes                    |                        |                        | Yes                    |                       |
| ZIP FE                        |                       |                        |                        |                        |                        |                       |
| ZIP, dow, week FE             |                       |                        |                        |                        |                        | Yes                   |
| Clustering                    | region-date           | ZIP                    | region-date            | ZIP                    | region-date            | region-date           |
| Observations                  | 21,536                | 21,536                 | 22,231                 | 22,231                 | 22,188                 | 22,188                |

*Notes:* Bandwidth 30 throughout, Poisson pseudo-maximum-likelihood at the ZIP-day level, trips counted at the origin ZIP. Column (1) is the specification of Table A17. Fixed effects and the clustering level are stated at the foot of each column; all columns control for mean temperature and precipitation. Columns (1) and (2) are the same regression under two variance estimators, as are (3) and (4): clustering by ZIP permits arbitrary correlation within a ZIP over time but treats ZIPs as independent on the same day, which they are not, since an alert is issued for a whole basin for a whole day. It yields standard errors roughly a third the size. Columns (3) and (4) omit fixed effects and so compare across all ZIP-days. \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Table A19: Placebo Thresholds Below the Alert Cutoff - Regression Discontinuity

|                           | BW = 25 |               |       | BW = 30 |               |       |
|---------------------------|---------|---------------|-------|---------|---------------|-------|
| Cut point                 | Alerts  | $\hat{\beta}$ | $t$   | Alerts  | $\hat{\beta}$ | $t$   |
| <b>0 (true threshold)</b> | 33      | -0.0697       | -1.88 | 33      | -0.0692       | -2.02 |
| -25                       | 3       | 0.0053        | 0.18  | 14      | 0.0453        | 1.49  |
| -30                       | 0       | -0.0120       | -0.35 | 3       | 0.0171        | 0.54  |
| -35                       | 0       | -0.0243       | -0.75 | 0       | -0.0394       | -1.30 |
| -40                       | 0       | 0.0299        | 0.99  | 0       | 0.0089        | 0.31  |
| -45                       | 0       | -0.0007       | -0.03 | 0       | -0.0041       | -0.22 |
| -50                       | 0       | -0.0284       | -1.24 | 0       | -0.0257       | -1.21 |

*Notes:* Each row re-estimates the specification of Table A17 with the threshold shifted to the stated point on the AQI scale, so that *Above* switches on where no alert was in fact issued. *Alerts* counts real alert basin-days falling inside the window: a row is a clean null only where that count is zero, which holds from  $-30$  down at  $BW = 25$  and from  $-35$  down at  $BW = 30$ . The  $-25$  row at  $BW = 30$  still contains 14 of the 33 alert basin-days and so is not a null.  $t$  uses standard errors clustered by basin  $\times$  date. The true threshold produces the largest discontinuity in the grid, between 1.7 and 2.4 times the largest clean placebo, and it is the only cut point at which  $|t|$  exceeds 1.8. Placebo estimates take both signs.

## A.1 Appendix - Link between Alerts and Model

Summarizing what the computation delivers, Section 7 asks what the welfare cost of the season would have been if the awareness an alert induces had extended to days with moderate or low smoke. Answering this requires changing the smoke coefficients on those days, since awareness mediates the behavioral response to smoke. The regression discontinuity delivers a drop in trips while the counterfactual requires a scale factor on the smoke coefficients. Converting one into the other requires a rate: the change in trips that one unit of awareness produces. I denote that rate  $\bar{B}$  and compute it from the model. The division  $\beta_{RD}/\bar{B}$  removes the units of trips and leaves the awareness gain  $r$ , which is what I use to rescale coefficients in the counterfactual welfare scenario. The rest of this appendix derives  $\bar{B}$ .

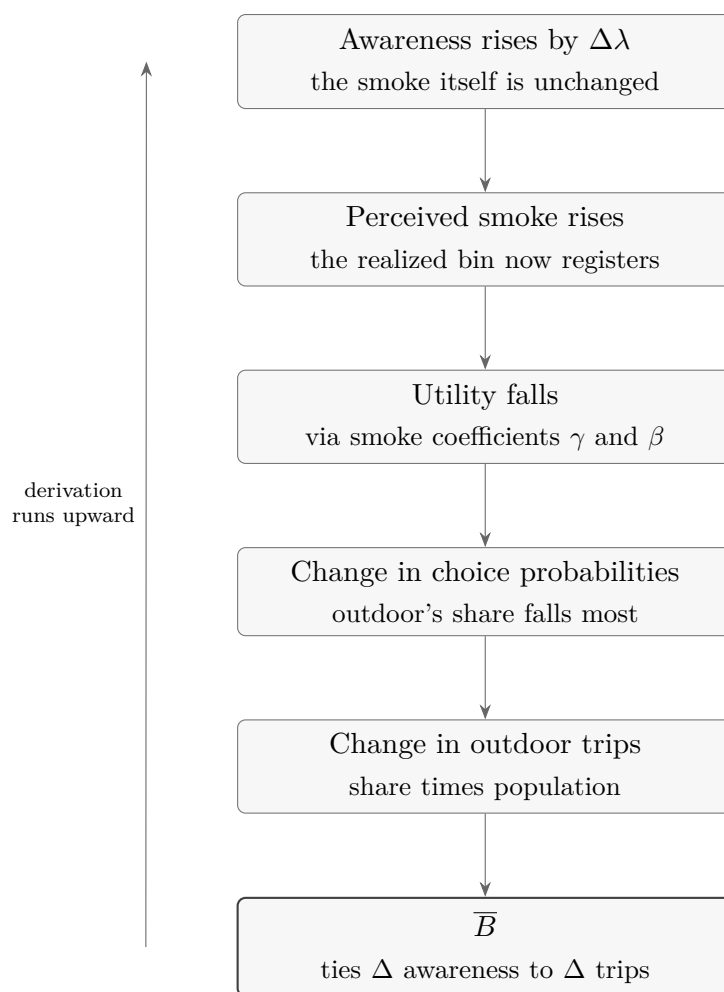
Figure A21 illustrates how awareness affects trips in the model. The alert raises awareness of existing smoke by  $\Delta\lambda$ , so that perceived smoke increases. The estimated coefficients  $\gamma_a$  and  $\beta_a$ , which are responses to perceived smoke, price that increase into the utility of each activity  $a$ . Utility falls by different amounts across activities. Because choice probabilities are shares, those differences shift the probability of choosing an outdoor activity, and that probability times population is outdoor trips.

The share of the awareness gain  $\Delta\lambda$  is unobserved, so the derivation of  $\bar{B}$  works backward (see Figure A21). Steps 1-3 link the observed jump in trips to the model. Step 1 writes outdoor trips as population times a choice probability. Step 2 differences this across the threshold, and everything continuous there drops out, leaving changes in the utility of each activity. Step 3 linearizes that expression. Steps 4-5 writes the utility change as dependent on change in awareness. The alert leaves air quality unchanged, so the only thing that moves at the threshold is what individuals believe about smoke, and beliefs enter utility through the smoke terms. Each utility change is therefore a change in beliefs priced by smoke coefficients, equal to  $\Delta\lambda$  times the disutility of being exposed to that additional smoke ( $K_a$ ). Finally, in step 6, I use the expression from steps 4-5 to factor  $\Delta\lambda$  out of every term in step 3. What remains is a weighted sum of the  $K_a$ , computable from the estimated coefficients and the fitted choice probabilities, which I call  $\bar{B}$ . I obtain that  $\beta_{RD} = \Delta\lambda\bar{B}$ .

**STEP 1.** The starting point of this exercise is to express the RD outcome, outdoor trips, as the model generates it. Expected outdoor trips from origin  $o$  on day  $t$  are the product of the population at  $o$  and the probability of choosing the outdoor activity

$$\mathbb{E}[y_{ot}] = N_{ot} P_{outdoor|o,t}.$$

Figure A21: Construction of  $\overline{B}$ . An alert raises awareness without changing air quality, so more of the smoke already present is perceived. The smoke coefficients price that perception into the utility of each activity, and the choice probabilities convert the resulting utility drops into a change in outdoor trips. The appendix derives this chain in reverse, from the measured jump in trips back to the awareness gain that produced it.



Write  $W_{ot}^a$  for the deterministic utility of activity  $a$  in the upper tier, evaluated at beliefs

$$W_{ot}^a = \sum_{b \geq 1} \gamma_a^b \pi_{ot}^b + \mu \mathcal{V}_{ot}^a$$

where  $\pi_{ot}^b$  is the belief that smoke  $o$  falls in bin  $b$  and  $\mathcal{V}_{ot}^a$  the inclusive value over the destinations reachable under activity  $a$ , through which destination beliefs enter. The choice probability reads

$$P_{outdoor|o,t} = \frac{\exp(W_{ot}^{outdoor})}{\sum_{a'} \exp(W_{ot}^{a'})}.$$

Taking logs

$$\log \mathbb{E}[y_{ot}] = \log N_{ot} + W_{ot}^{outdoor} - \log \sum_{a'} \exp(W_{ot}^{a'}).$$

The population  $N_{ot}$  counts devices by home location and is fixed at the daily frequency.

**STEP 2.** Consider two days around the alert threshold, one right below and one right above, where the alert switches on. In this model, the alert operates through beliefs alone, so are beliefs  $\pi_{ot}^b$  and  $\pi_{dt}^b$  and through them the utility  $W_{ot}^a$ . Differencing  $\log \mathbb{E}[y_{ot}]$ , all terms that are continuous through the threshold ( $N_{ot}$ ) drop out:

$$\Delta \log \mathbb{E}[y_{ot}] = \Delta W_{ot}^{outdoor} - \Delta \log \sum_{a'} \exp(W_{ot}^{a'}). \quad (\text{A1})$$

The left-hand side is the jump in log outdoor trips at the alert threshold. The regression discontinuity measures this jump and finds a drop of 9 to 14 percent.

**STEP 3.** The jump in trips is not linear in utility because of the term  $\log \sum_{a'} \exp(W_{ot}^{a'})$  in (A1). This step linearizes it, giving the rate at which utility drops translates into trip drops.

The second term on the right-hand side is the jump in the log-sum across activities. Write the log-sum as a function of the four activity utilities,

$$f(W_{ot}^{home}, W_{ot}^{work}, W_{ot}^{outdoor}, W_{ot}^{inside}) = \log \sum_{a'} \exp(W_{ot}^{a'}).$$

The alert shifts every argument of  $f$ : each activity's utility jumps by  $\Delta W_{ot}^a$ . To first order, the jump in  $f$  is the sum of its partial derivatives times the jumps in each argument,

$$\Delta f \approx \frac{\partial f}{\partial W_{ot}^{home}} \Delta W_{ot}^{home} + \frac{\partial f}{\partial W_{ot}^{work}} \Delta W_{ot}^{work} + \frac{\partial f}{\partial W_{ot}^{outdoor}} \Delta W_{ot}^{outdoor} + \frac{\partial f}{\partial W_{ot}^{inside}} \Delta W_{ot}^{inside} = \sum_a \frac{\partial f}{\partial W_{ot}^a} \Delta W_{ot}^a.$$

Each partial derivative is a choice probability. Differentiating  $f$  with respect to the utility of activity  $a$ , the chain rule gives one over the sum times the derivative of the sum, and only the  $a$  term of the sum involves  $W_{ot}^a$

$$\frac{\partial f}{\partial W_{ot}^a} = \frac{\exp(W_{ot}^a)}{\sum_{a'} \exp(W_{ot}^{a'})} = P_{a|o,t}.$$

Substituting the probabilities for the derivatives

$$\Delta \log \sum_a \exp(W_{ot}^a) \approx \sum_a P_{a|o,t} \Delta W_{ot}^a.$$

Substituting into (A1) and grouping the outdoor term

$$\Delta \log \mathbb{E}[y_{ot}] \approx (1 - P_{outdoor|o,t}) \Delta W_{ot}^{outdoor} - \sum_{a \neq outdoor} P_{a|o,t} \Delta W_{ot}^a. \quad (\text{A2})$$

The jump in trips is now a weighted sum of the utility jumps, with weights that are fitted choice probabilities. The jumps themselves are not yet computable. Step 4 expresses them in terms of the awareness gain.

**STEP 4.** It remains to compute the utility jumps  $\Delta W_{ot}^a$ . Recall the upper-tier utility,

$$W_{ot}^a = \sum_{b \geq 1} \gamma_a^b \pi_{ot}^b + \mu \mathcal{V}_{ot}^a.$$

The coefficients  $\gamma_a^b$  and  $\mu$  do not change at the threshold; only the beliefs jump, directly in the first term and through destination beliefs inside the inclusive value. Differencing term by term,

$$\Delta W_{ot}^a = \sum_{b \geq 1} \gamma_a^b \Delta \pi_{ot}^b + \mu \Delta \mathcal{V}_{ot}^a. \quad (\text{A3})$$

I impose a two-point structure on beliefs, governed by a single parameter  $\lambda$ , the awareness share. On a day with smoke in bin  $b^*$ , an individual believes with probability  $\lambda$  that smoke is in  $b^*$  and with probability  $1 - \lambda$  that the air is clean,

$$\pi^b = \begin{cases} \lambda & \text{if } b = b^* \\ 1 - \lambda & \text{if } b = 0 \\ 0 & \text{otherwise.} \end{cases}.$$

The same structure holds at the origin and at each destination, with  $b^*$  the bin realized at the relevant location. A single  $\lambda$  governs both tiers: beliefs about destination smoke are formed from the information available at the origin, and an alert is a basin-wide broadcast about the day rather than a destination-specific signal.

An alert raises  $\lambda$  by  $\Delta\lambda$ . The clean bin is the omitted category and carries no coefficient, so only the realized bin enters utility, and for  $b \geq 1$  the belief jump is

$$\Delta\pi^b = \begin{cases} \Delta\lambda & \text{if } b = b^* \\ 0 & \text{otherwise.} \end{cases}.$$

Multiplying by coefficients and summing over bins, the indicator is one for the realized bin and zero for every other, so only one term of the sum survives,

$$\sum_{b \geq 1} \gamma_a^b \Delta\pi_{ot}^b = \Delta\lambda \gamma_a^{b_{ot}}$$

where  $b_{ot}$  is the realized smoke bin at the origin. The first term of (A3) is now  $\Delta\lambda$  multiplying a single estimated coefficient. Step 5 does the same for the second term, the inclusive value, which carries the destination beliefs with the same logic.

Both terms of (A3) are now resolved. Each activity's utility jump is the same unknown  $\Delta\lambda$  multiplying  $K_a$ , a term built from the estimated coefficients and fitted shares. Step 6 substitutes this into (A2) and solves for  $\Delta\lambda$ .

**STEP 5.** The inclusive value is itself a log-sum, over destinations rather than activities. The same expansion as above applies one tier down: its first-order jump is the sum of the destination-utility jumps, weighted by the destination choice probabilities,

$$\Delta\mathcal{V}_{ot}^a \approx \sum_{d \in D_a} P_{d|a,o,t} \Delta V_{odt}^a.$$

Within each destination's utility, only the smoke beliefs jump, and the same collapse applies with  $\beta_a$  and the destination's realized bin,

$$\Delta V_{odt}^a = \sum_{b \geq 1} \beta_a^b \Delta\pi_{dt}^b = \Delta\lambda \beta_a^{b_{dt}}$$

where  $\beta_a^{b_{dt}}$  denotes the coefficient on whichever bin the smoke at  $d$  actually falls in, and is zero when  $d$  is clean.

Substituting, and pulling  $\Delta\lambda$  out of the sum,

$$\Delta\mathcal{V}_{ot}^a \approx \Delta\lambda \sum_{d \in D_a} P_{d|a,o,t} \beta_a^{b_{dt}}.$$

Substituting the two pieces back into (A3) the decomposition of  $\Delta W_{ot}^a$ ,

$$\Delta W_{ot}^a = \Delta\lambda \gamma_a^{b_{ot}} + \mu \Delta\lambda \sum_{d \in D_a} P_{d|a,o,t} \beta_a^{b_{dt}} = \Delta\lambda \left[ \gamma_a^{b_{ot}} + \mu \sum_{d \in D_a} P_{d|a,o,t} \beta_a^{b_{dt}} \right].$$

Each utility jump is therefore  $\Delta\lambda$  times a computable term, which I denote  $K_a$ ,

$$\Delta W_{ot}^a = \Delta\lambda K_a, \quad K_a = \gamma_a^{b_{ot}} + \mu \sum_{d \in D_a} P_{d|a,o,t} \beta_a^{b_{dt}}. \quad (\text{A4})$$

**STEP 6.** The remaining task is the inversion: express the jump in trips as the awareness gain times a computable factor, and solve for the gain. Substituting  $\Delta W_{ot}^a = \Delta\lambda K_a$  into the collected expression for the jump in trips (A2),  $\Delta\lambda$  factors out of every term,

$$\Delta \log \mathbb{E}[y_{ot}] \approx \Delta\lambda \left[ \left(1 - P_{outdoor|o,t}\right) K_{outdoor} - \sum_{a \neq outdoor} P_{a|o,t} K_a \right] \equiv \Delta\lambda B_{ot} \quad (\text{A5})$$

which defines the factor  $B_{ot}$ . It varies across origins and days through the realized bins and the fitted shares. Averaging over the threshold days entering the regression discontinuity,

$$\beta_{RD} = \Delta\lambda \bar{B}. \quad (\text{A6})$$

The goal is to solve  $\beta_{RD} = \Delta\lambda \bar{B}(\gamma, \beta)$  for  $\Delta\lambda$ , the increase in awareness an alert introduces. Note that this step assumes the awareness gain  $\Delta\lambda$  is the same across ZIP-days with different pollution levels. Two obstacles:  $\bar{B}(\gamma, \beta)$  contains the true coefficients  $\gamma$  and  $\beta$ , which are not observed. I showed that estimated smoke coefficients blend true responses with beliefs: with the modeling assumptions, beliefs place weight  $\lambda_0$  on the realized bin, so the smoke covariate driving behavior is  $\lambda_0$  times the indicator, and estimation recovers  $\lambda_0$  times the true coefficient:

$$\hat{\gamma}_a = \lambda_0 \gamma_a, \quad \hat{\beta}_a = \lambda_0 \beta_a.$$

Evaluate the bracket at these estimated coefficients. This object,  $\bar{B}(\hat{\gamma}, \hat{\beta})$ , is computable: coefficients, realized bins, and fitted shares are all available. The bracket is a weighted sum of coefficients, so

$$\bar{B}(\hat{\gamma}, \hat{\beta}) = \bar{B}(\lambda_0 \gamma, \lambda_0 \beta) = \lambda_0 \bar{B}(\gamma, \beta). \quad (\text{A7})$$

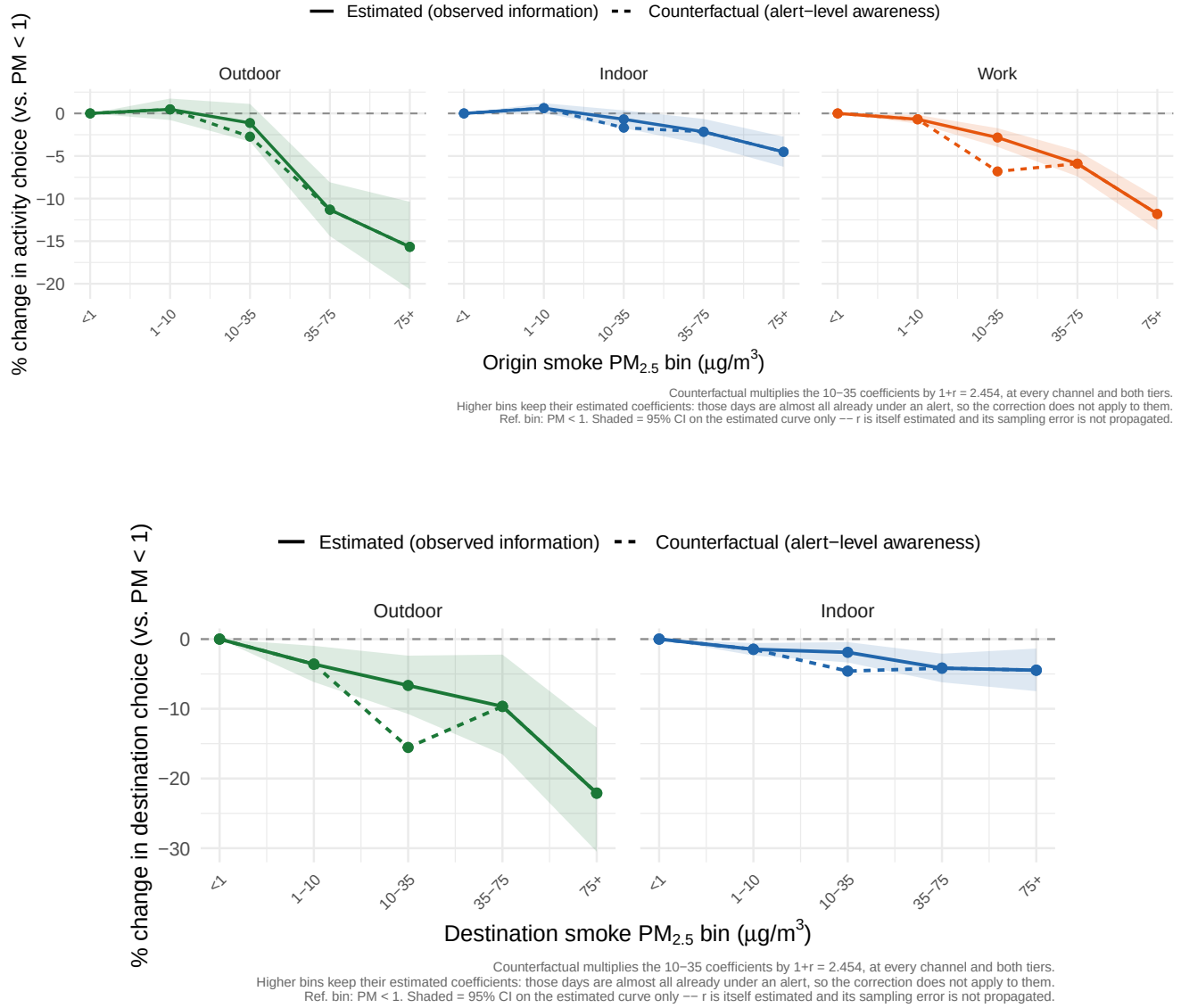
Now divide equation (A6) by (A7)

$$\frac{\beta_{RD}}{\bar{B}(\hat{\gamma}, \hat{\beta})} = \frac{\Delta\lambda \bar{B}(\gamma, \beta)}{\lambda_0 \bar{B}(\gamma, \beta)} = \frac{\Delta\lambda}{\lambda_0} \equiv r. \quad (\text{A8})$$

$\bar{B}(\gamma, \beta)$  cancels. The left-hand side is computable:  $\beta_{RD} = -0.069$  from the regression discontinuity, and  $\bar{B}(\hat{\gamma}, \hat{\beta}) = -0.047$  from the estimated coefficients and observed shares, averaged over smoky days within the same bandwidth. The right-hand side defines  $r$ , the awareness gain relative to its baseline level. The gain itself is not identified:  $\Delta\lambda$  and  $\lambda_0$  enter only through their ratio.

The result is one number,  $r$ : an alert raises awareness by a proportion  $r$  of its baseline level. On days without an alert, behavior runs on the operative coefficients  $\hat{\gamma}_a = \lambda_0 \gamma_a$  and  $\hat{\beta}_a = \lambda_0 \beta_a$ , which are exactly the estimated ones. Raising awareness by  $\Delta\lambda$  multiplies them by  $(\lambda_0 + \Delta\lambda)/\lambda_0 = 1 + r$ . The counterfactual therefore multiplies the estimated smoke coefficients at both tiers by  $1 + r$  on smoky days without an alert and recomputes welfare. Figure A22 plots the estimated dose-response alongside the counterfactual obtained by scaling the 10–35 coefficients by  $1 + r$ .

Figure A22: Smoke dose-response under observed and alert-level awareness



*Notes:* Percent change in choice probability relative to the reference bin,  $\text{PM}_{2.5} < 1 \mu\text{g}/\text{m}^3$ . The upper panel gives tier-1 activity choice by origin smoke, the lower panel tier-2 destination choice by destination smoke. Solid markers give the estimated dose-response under observed information. Hollow markers give the counterfactual in which the  $10-35 \mu\text{g}/\text{m}^3$  coefficients are multiplied by  $1 + r = 2.45$ . Higher bins retain their estimated coefficients, since days in those bins are almost all already under an alert. Shaded bands are 95% confidence intervals on the estimated curve only. The sampling error in  $r$  is not propagated.