Welfare Losses from Wildfire Smoke: Evidence from Daily Outdoor Recreation Data

Jacob Gellman, Margaret Walls, Matthew Wibbenmeyer*

15 April 2024

Abstract

Wildfire smoke pollution is growing in the western United States. Estimates of health impacts from smoke are numerous, but few revealed preference estimates of its damages exist. We study a setting where individuals are directly exposed to smoke and where avoidance behavior is measured with high frequency: outdoor recreation. We combine millions of administrative campground reservation records with satellite data on wildfire, smoke, and air pollution. The data allow us to model sequential recreation decisions under evolving information using a novel control function approach. We estimate wildfire smoke reduces welfare by $107 per person per trip. These damages are larger when campgrounds are affected by consecutive days of smoke. In total, 21.5 million outdoor recreation visits in the western United States are affected by smoke every year, with annual welfare losses of approximately $2.3 billion. These findings contribute to a growing body of evidence on the costs of wildfire smoke.

Keywords: Wildfire smoke, wildfire, air pollution, recreation
JEL codes: Q26, Q51, Q53, Q54

*Jacob Gellman is in the Department of Economics, University of Alaska Anchorage (jgellman@alaska.edu). Margaret Walls is at Resources for the Future, Washington, DC (walls@rff.org). Matthew Wibbenmeyer is at Resources for the Future, Washington, DC (wibbenmeyer@rff.org). We thank Andrew Plantinga, Olivier Deschênes, Kelsey Jack, Max Moritz, Steve Dundas, Jeff Englin, David J. Lewis, Molly Robertson, Roger von Haefen, and John Whitehead for comments and suggestions on this paper. We are also grateful to seminar participants in the UCSB Environmental Economics group, AERE 2021 Summer Conference, Western Economics Association 2022 Conference, 2023 Occasional Workshop, 2023 AERE OSWEET, University of Alaska Anchorage, and the University of Nevada, Reno. In addition, we thank Jude Bayham and coauthors, who generously shared the gridded air quality data. This research was funded by a United States Department of Agriculture National Institute of Food and Agriculture (NIFA) Agriculture and Food Research Initiative (AFRI) grant, award number 2020-67023-33258.

1 Introduction

Large wildfires have increased in frequency and severity in the western United States, and these trends are expected to continue as the climate warms (Abatzoglou and Williams 2016; Westerling 2016, 2018; Williams et al. 2019). This wildfire activity has generated considerable increases in smoke, which can cover large geographic areas and affect air quality hundreds of miles away. Wildfire smoke now accounts for up to half of particulate matter pollution in some areas of the western United States (Burke et al. 2021), and some literature has found that particulate emissions from smoke harm health more severely than those from other sources (Aguilera et al. 2021; Kochi et al. 2010). Empirical evidence on the negative effects of wildfire smoke is growing, with studies finding increased morbidity and mortality, negative effects on birth outcomes, reductions in academic performance, and impacts in labor markets (Borgschulte et al. 2022; Heft-Neal et al. 2023; Kochi et al. 2010; McCoy and Zhao 2021; Miller et al. 2021; Reid et al. 2016; Wen and Burke 2022).

Few estimates of the welfare costs of smoke exist, however, and those that do are based on stated preference or survey-based methods (Jones 2017; Richardson et al. 2012; Richardson et al. 2013) or were calculated by applying a value of statistical life (VSL) to changes in mortality (Miller et al. 2021). Welfare estimates derived from revealed preference methods are limited, perhaps because smoke is challenging to study in a revealed preference setting. In hedonic property value studies, it may be difficult to distinguish effects of long-run increases in wildfire smoke from other unobserved region-level changes to housing markets. In travel cost recreation studies, data at a high temporal resolution are needed to compare changes in behavior in response to transient smoke conditions.

We use high-frequency data on outdoor recreation in a travel cost framework to provide the first revealed preference estimates of welfare damages from wildfire smoke. Outdoor recreation is advantageous for studying responses to smoke for several reasons. First, exposure to smoke is high for recreationists. Wildfire season and peak outdoor recreation season tend to coincide, with more than 1 million national park visitor-days per year taking place during hazardous smoke conditions (Gellman et al. 2022). Second, visitors to recreation sites spend large amounts of time outdoors and tend to engage in vigorous activities, such as hiking, which may exacerbate the effects of exposure (Korrick et al. 1998; Richardson et al. 2012). Smoke can also reduce the visibility and amenity value from recreation trips. Third, structural modeling of outdoor recreation decisions using the travel cost framework can naturally yield welfare cost estimates (Parsons 2017; Lupi et al. 2020).

Our approach uses millions of administrative campground reservation records from nearly 1,000 federally managed campgrounds, which we combine with daily data on wildfire, smoke, and air pollution. The data include information on campsite reservations, which are typically made well in advance, and cancellations of planned trips. The cancellation data allow us to account for the short-term and transient nature of wildfire smoke by focusing on decisions made both before and after visitors have knowledge of air quality conditions at the site.

We model visitors' reservation and cancellation decisions as a two-stage discrete choice. In the first stage, a visitor chooses to reserve based on expected site conditions; in the second stage, they decide whether to cancel based on realized site conditions. A key feature of this setting is that reservations and cancellations are made in sequence; a visitor can only cancel if they have a reservation, meaning they have already demonstrated a preference for the site. This induces sample selection in the population holding reservations; if unaccounted for, this selection would bias welfare estimates of wildfire smoke. To correct for this bias, we develop a novel control function approach to link preferences across choices (Wooldridge 2015). The control function uses estimated preferences from the first-stage reservation decision to remove selection bias in the second-stage cancellation decision. We demonstrate the effectiveness of this approach through numerical simulations.

We find that wildfire smoke imposes welfare costs of $107 per person per recreation trip. Without accounting for sample selection using the control function, the analysis would have implied damages of $154 per person per trip, overstating welfare impacts by 44 percent. Damage estimates vary by the number of smoke-affected days in the week leading up to arrival. When a campground is affected by smoke on only one day in the week of arrival, estimated damages are as low as $32 per person per trip; but when it is affected on all seven days, losses are as high as $432 per person per trip. The heterogeneity in our estimates is consistent with greater perceived likelihood of experiencing smoke during weeks with more frequent smoke, more severe smoke impacts, or potentially both. In the appendix, we report additional heterogeneity and robustness checks, including a placebo for smoke, a test of the influence of uncancelled no-shows, and heterogeneous responses by type of recreation site.

The overall magnitude of welfare losses from wildfire smoke is large. We find that in the western United States, an average of 21.5 million outdoor recreation visits per year (approximately 4.2 percent of all annual visits), which includes both camping and non-camping trips, are affected by wildfire smoke on lands managed by the National Park Service, US Forest Service, Bureau of Land Management, US Army Corps of Engineers, and state park systems. Applying the empirical welfare estimate of $107 per person per trip, this figure implies welfare losses of roughly $2.3 billion per year.

Our study adds to the growing literature on wildfire and wildfire smoke. Several studies have examined short-run consequences of smoke, finding that it increases emergency room visits (Heft-Neal et al. 2023) and mortality (Miller et al. 2021), decreases labor earnings (Borgschulte et al. 2022), and worsens expressed sentiment on social media (Burke et al. 2022; Loureiro et al. 2022). Some researchers have used survey evidence to estimate welfare impacts of smoke. Based on US life satisfaction survey data, for example, Jones (2017) estimates a willingness to pay (WTP) of $373 to avoid one wildfire smoke day per six-month period. Richardson et al. (2012) employ a defensive expenditure approach using survey data on air purifier purchases to estimate a WTP to reduce one smoke-induced health symptom day of $84 (in 2009 dollars). Using a stated preference contingent valuation survey, Richardson et al. (2013) estimate a WTP to avoid one smoke-induced symptom day of $95, which they compare to an $87 WTP estimate based on a cost-of-illness approach (both figures in 2009 dollars).

We measure the welfare damages of smoke exposure using cancellations of planned recreation trips; our estimate of $107 per trip roughly translates to $38 per day. This value is in line with estimates from Richardson et al. (2012, 2013) but considerably lower than the estimate from Jones (2017). Compared to our aggregate welfare estimate of $2.3 billion per year, Miller et al. (2021) value damages from mortality among elderly Medicare recipients in the United States at $6 to 170 billion, depending on VSL assumptions. Borgschulte et al. (2022) find annual lost labor earnings of $125 billion per year due to wildfire smoke.

Measuring the cost of wildfire smoke is crucial to informing public policy. The federal government spent an average of $2.8 billion per year on fire suppression between 2017 and 2021, and the State of California spent an average of $900 million per year from 2018 to 2022.[1,2] California has proposed spending $1.2 billion over fiscal years 2022-23 and 2023-24 on wildfire mitigation measures, including vegetation management, prescribed burns, and home hardening.[3] These activities are consistent with the state's recently declared goal to treat 1 million acres of hazardous fuels per year.[4] Together, the 2021 Infrastructure Investment and Jobs Act (P.L. 117-58) and the Inflation Reduction Act of 2022 (P.L. 117-169) provided $24 billion for federal wildfire programs (Goldman et al. 2022). Understanding the welfare costs of wildfire and smoke is critical to assessing the benefit of these public expenditures.

We also contribute to the recreation demand literature by using novel methods and data sources. We value a transient environmental disamenity in a setting where users make decisions under evolving sets of information. The two-stage choice structure that links preferences across decisions is informed by literature on sample selection correction in nonlinear models (Greene 2012; Terza 2009) and recreation contexts (Cameron and DeShazo 2013; Cameron and Kolstoe 2022; Kolstoe and Cameron 2017; Lewis et al. 2019). Our framework could be used to model sample selection or sequential choices in other non-linear or discrete-choice settings. It could also be applied to recreation studies valuing other short-lived environmental amenities, such as temperature and precipitation extremes or acute pollution events. In addition to the modeling, our use of administrative data contributes to a recent literature using new, large, or innovative data to study recreation across broad regions (Cameron and Kolstoe 2022; Dundas and von Haefen 2020; English et al. 2018; Parthum and Christensen 2022).

The remainder of this paper is organized as follows. In Section 2, we describe the data sources for the study, including recreation, smoke, fire, and pollution data. We also discuss

[1]National Interagency Fire Center. Suppression Costs. `https://www.nifc.gov/fire-information/statistics/suppression-costs`.

[2]California Department of Forestry and Fire Protection. Suppression Costs. `https://www.fire.ca.gov/stats-events`.

[3]California Legislative Analyst's Office. The 2022-23 Budget Wildfire and Forest Resilience Package. `https://lao.ca.gov/Publications/Report/4495`.

[4]Agreement for shared stewardship of California's forest and rangelands between the State of California and the USDA Forest Service, Pacific Southwest Region. `https://www.gov.ca.gov/wp-content/uploads/2020/08/8.12.20-CA-Shared-Stewardship-MOU.pdf`.

several descriptive features of the data. Section 3 describes the modeling approach, including a conceptual framework and a description of the estimating dataset. In Section 4, we turn to estimation, describing various sets of results. Section 5 appraises the total annual welfare damages of wildfire smoke in the western United States. Section 6 concludes.

2 Data

We combine data on recreation behavior, travel costs, wildfire smoke, air pollution, wildfire activity, and weather into three main datasets. The first is a daily panel of federally managed campgrounds in the western United States over the period 2010 to 2017. This panel includes climate normals and local daily measurements of smoke, wildfire activity, pollution, and weather. The second dataset is a record of individual-level reservations for campgrounds, which we link to the daily campground panel to show site conditions for users' reservation dates. The last dataset aggregates the individual users to measure daily reservation activity by distance from each campground.

2.1 Recreation

We obtained data on campground use from Recreation.gov,[5] the primary web portal for reserving sites at federally managed campgrounds, including those managed by the National Park Service, Bureau of Land Management, US Forest Service, US Army Corps of Engineers, and Bureau of Reclamation. Figure A1 in Appendix A displays the web interface as a user would experience it. The website gives users information about campground amenities, prices, availability, and nearby points of interest.

The raw data include more than 90 million transactions from more than 7 million unique users. We limit attention to campgrounds in the eleven western states, during the months of May through September, and for the years 2010 to 2017, which leaves more than 16 million transactions from 2 million unique users at 999 campgrounds. Our analysis is primarily concerned with overnight camping and excludes, for instance, large-group or equestrian facilities.

[5] Recreation.gov. `https://www.recreation.gov`.

The data give detailed information on reservations, walk-ins, cancellations, no-shows, transaction dates, payments, refunds, zip code of origin, group size, user identifiers, and other information. For every transaction in an order, such as a payment or cancellation, the exact time is known. For the 999 campgrounds, 84 percent of transactions were made online, 9 percent over the phone, and 7 percent on-site (such as walk-ins or early checkouts).

2.2 Travel costs

We calculate travel costs based on the distance and travel time between origin zip codes and destination campgrounds with GraphHopper, an open source routing engine that uses Djikstra's algorithm and OpenStreetMap data.[6,7] We calculate nearly 5.4 million routes representing 5,379 origin points and 999 destinations. Our estimates reflect the fastest routes by car between each origin and destination. Optimal routes generally match routes identified by Google Maps during periods of low traffic. To identify the coordinates of each user's zip code, we match zip codes to Census Zip Code Tabulation Areas (ZCTAs) and find the centroid of each ZCTA.[8,9] Figure A2 in Appendix A displays an example route.

Following English et al. (2018), we calculate the per-person travel costs between ZCTA z and campground j as

$$c_{zjt} = \frac{p^D_{zt} D_{zj}}{n} + p^T_{zt} T_{zj}, \tag{1}$$

for travel distance D_{zj}, travel time T_{zj}, and group size n. The per-kilometer cost of traveling between ZCTA z and campground j is given by p^D_{zt} and includes gasoline costs, per-kilometer vehicle maintenance, and per-kilometer marginal vehicle depreciation. For gasoline costs, we use state- and year-specific averages of per-kilometer gasoline costs during summer months, based on per-gallon gasoline costs from the Energy Information Administration and nation-

[6]GraphHopper. `https://www.graphhopper.com`.

[7]GraphHopper GitHub. `https://github.com/crazycapivara/graphhopper-r`.

[8]Health Resources and Services Administration, John Snow, Inc., and American Academy of Family Physicians. Uniform Data System. `https://udsmapper.org/zip-code-to-zcta-crosswalk`.

[9]Because ZCTA centroids may not be located along roads, we snapped them to the nearest road using Census TIGER/Line shapefiles and used the nearest points along roads as origin points.

wide average fleet fuel economy.[10,11] We use per-kilometer vehicle maintenance costs and marginal depreciation from AAA data, as in English et al. (2018).[12,13] Lastly, we measure hourly costs of travel time p_{zt}^{T} as one-third of the average household income in ZCTA z divided by 2,080 hours worked per year (English et al. 2018). All numbers are inflation-adjusted to 2020 US dollars.

2.3 Smoke and air pollution

Our measure of smoke impacts combines satellite-derived wildfire smoke plume data and gridded ground-level $PM_{2.5}$ monitoring data. For each day, we record whether a campground was covered by smoke based on daily observations of smoke plumes from the National Oceanic and Atmospheric Administration (NOAA) Hazard Mapping System (Schroeder et al. 2008).[14] Each day, NOAA analysts manually trace the perimeters of smoke plumes using satellite photography, producing daily shapefiles. These data have been used in studies examining the effect of smoke on air pollution, health, labor markets, self-protective behavior, and crime (Borgschulte et al. 2022; Burke et al. 2021, 2022; Burkhardt et al. 2019; Gellman et al. 2022; Heft-Neal et al. 2022; Miller et al. 2021; Preisler et al. 2015).

One challenge presented by this dataset is that satellite photography does not reveal where in the air column a smoke plume is: it could be at the ground level or high in the atmosphere. If the latter, it may not reflect on-the-ground conditions. To address this challenge, we code an area as smoke-affected only if it is covered by a smoke plume and its ground-level $PM_{2.5}$ is above at least 1.64 standard deviations of the location-specific seasonal mean for non-smoke days, which represents the 95th percentile of a normal distribution (Burkhardt et al. 2019; Gellman et al. 2022).[15] Figure A3 in Appendix A displays an

[10] Energy Information Administration. Weekly Retail Gasoline and Diesel Prices. `https://www.eia.gov/dnav/pet/pet_pri_gnd_a_epmr_pte_dpgal_m.htm`.

[11] Bureau of Transportation Statistics. Average Fuel Efficiency of US Light Duty Vehicles. `https://www.bts.gov/content/average-fuel-efficiency-us-light-duty-vehicles`.

[12] For example: AAA. Your Driving Costs 2016. `https://publicaffairsresources.aaa.biz/wp-content/uploads/2016/03/2016-YDC-Brochure.pdf`.

[13] Marginal depreciation cost is calculated by taking the difference in "reduced depreciation" and "increased depreciation" between 10,000 and 20,000 miles, divided by 10,000 miles, as in English et al. (2018). Some literature has argued that depreciation functions as a fixed cost rather than a marginal cost, in which case inclusion would cause an overstatement of travel cost (Hang et al. 2016).

[14] NOAA. Hazard Mapping System. `https://www.ospo.noaa.gov/Products/land/hms.html`.

[15] A "season" is defined as fall, winter, spring, or summer.

example of this restriction using kriged $PM_{2.5}$ data from Burkhardt et al. (2019). The map shows that although many areas were covered by smoke, only some had air quality poor enough to be coded as smoke-affected.

2.4 Wildfire activity

At each campground, we measure the daily distance to an actively burning fire. We combine NASA MODIS fire detection points with the United States Geological Survey Monitoring Trends in Burn Severity (MTBS) fire perimeter dataset.[16,17] The MODIS detection points record 1 km centroids of fire activity at a daily resolution, including agricultural and prescribed fires (Giglio et al. 2016). The MTBS data map the final perimeters for US wildfires. Combining these data has two advantages. First, the use of known wildfire perimeters filters out any MODIS points not associated with a large wildfire. Second, the MODIS detection points limit attention only to the portion of a wildfire that was burning on a given day. We use a 1 km buffer around the fire's final perimeter and its start and containment dates to filter MODIS points. Figure A4 in Appendix A demonstrates an example of this process for the western United States.

2.5 Temperature and precipitation

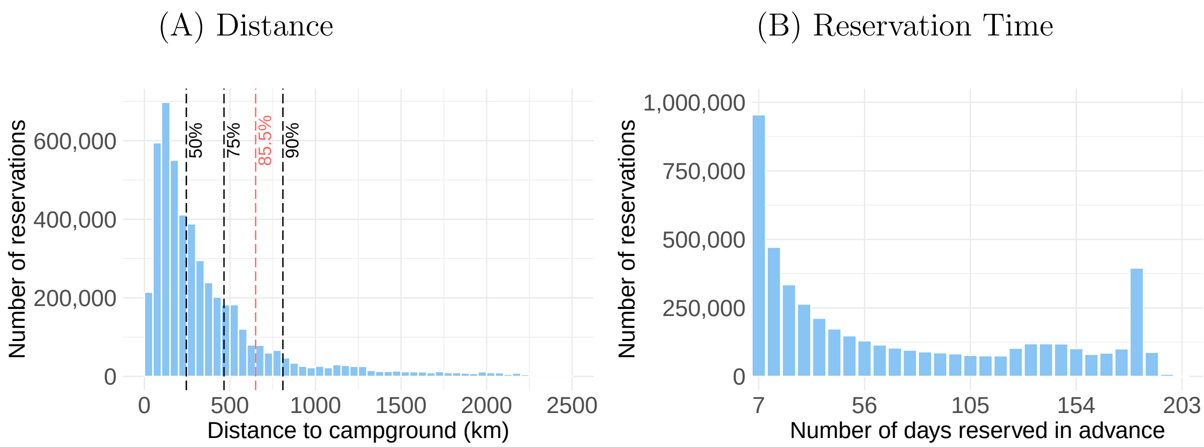
To control for weather conditions, we gather daily precipitation (mm) and maximum and minimum temperature (°C) for every campground. These data are published at a 4 km resolution by the PRISM Climate Group at Oregon State University.[18] In addition, at each campground, we record 30-year climate normals that reflect average conditions over the period 1980 to 2010.

[16]NASA. Earthdata. `https://earthdata.nasa.gov`.

[17]USGS. Monitoring Trends in Burn Severity. `https://www.mtbs.gov`.

[18]Northwest Alliance for Computational Science and Engineering, Oregon State University. PRISM Climate Data. `https://www.prism.oregonstate.edu`.

Figure 1: Driving Distance and Time between Reservation and Scheduled Arrival



Notes: Panel A shows the one-way driving distance of reservations from the destination campground, where the red line indicates a 650 km cutoff. Panel B displays the distribution of time reserved in advance of arrival date.

2.6 Descriptive features of the data

Of the 999 campgrounds in the analysis, 908 are managed by the US Forest Service, 50 by the National Park Service, 31 by the US Army Corps of Engineers, 5 by the Bureau of Land Management, and 5 by the Bureau of Reclamation. Figure A5 in Appendix A plots a map of the campgrounds in the dataset. Although the Forest Service manages most of the campgrounds, the most-visited locations tend to be in national parks.

For the main analysis, we restrict the set of potential reservers to residents living within one day's driving distance of a given campground. We set this restriction at 650 km (400 miles). English et al. (2018) report survey results showing that, beyond 500 miles of driving distance, a substantial portion of recreation visitors are likely to have flown to their destination, which adds additional complications to the calculation of travel cost. Exclusion of trips from far distances can also eliminate multipurpose trips. Panel A of Figure 1 shows that our 650 km restriction results in including more than 85 percent of reservations in the dataset. Half of our observed trips come from within 250 km (155 miles), and three-quarters come from within 450 km (280 miles). Appendix E reports the main results using alternative driving distance thresholds.

The timing of a reservation is also key for our setting. Wildfire smoke is a random event, meaning that visitors who reserve far in advance do not know whether their chosen campground will be smoke-affected during their visit. Panel B of Figure 1 shows that most visitors reserve far in advance, consistent with results in Walls et al. (2018). Although a plurality of visitors reserve within a week of arrival, a majority reserve early. In addition, a significant mass appears around six months in advance, which is the earliest that some popular destinations allow reservations. In the following section, we describe our modeling approach to study the cancellation decisions of visitors who reserved ahead of time.

3 Modeling approach

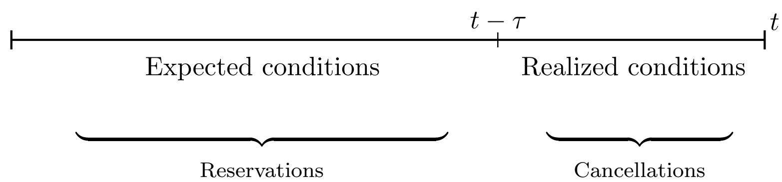
In this section, we model an individual's decision to visit a campground under smoke and non-smoke conditions. A key feature of this setting is that smoke is ephemeral; it is not a permanent feature of site quality but may impact air quality suddenly and without warning for several days or weeks. Given this feature, one estimation strategy might be to contrast reservation rates during smoke and non-smoke periods. However, because demand for campgrounds is high, campers must often make reservations far in advance, before smoke conditions are known. As campgrounds are often completely full shortly before a given date, rates of new reservations may not be responsive to smoke.[19]

We therefore consider the cancellation decisions of visitors who reserved ahead of time. The population that is eligible to cancel consists of those that already hold a reservation; thus, the sample is selected due to their demonstrated preference for the site. We account for selection bias induced by the reservation decision using a control function approach that explicitly models and estimates the "first-stage" reservation decision, in which visitors choose whether to reserve based on expected site conditions. In the second stage, close to the arrival date, they decide whether to cancel based on realized site conditions. Figure 2 illustrates the timing of these decisions, where t gives the arrival date and τ denotes a bandwidth

[19] For example, during May to September, 21 and 25 percent of Fridays and Saturdays, respectively, are completely booked as of one week in advance, and approximately 28 percent are completely unbooked; this low degree of variation in reservation rates would suppress identifying variation on smoke and non-smoke days.

sufficiently close to the arrival date. For the reservation decision we use a pooled zonal travel cost model to construct the control function for the second stage, trip-level model of cancellations. The control function uses preference parameters estimated in the first stage to remove selection bias in the second stage.

Figure 2: Timing of Decisions



3.1 Reservations

Define utility from visiting site j at time t as

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt} + \upsilon_{ijt}, \tag{2}$$

where V_{ijt} gives the observable portion of utility to household i from visiting campground j on date t, ε_{ijt} represents fixed unobserved preferences of household i for visiting site j at time t, and υ_{ijt} represents unobserved shocks to preferences that occur between time $t - \tau$ and time t. Let V_{ijt} be specified according to

$$V_{ijt} = \begin{cases} \delta c_{ijt} + \phi s_{jt} + X'_{jt}\gamma + \psi_{jt}, & j \in \{1, 2, ..., J\}; \\ 0, & j = 0, \end{cases} \tag{3}$$

where c_{ijt} gives the travel cost for person i to site j at time t, and s_{jt} represents smoke conditions at campground j on date t. The vector X_{jt} includes time-varying, campground-level conditions, such as precipitation, temperature, and proximity to an active wildfire. Additional variables include ψ_{jt}, which account for site- and time-specific fixed effects. Campground-related fixed effects capture location-specific unobservables, such as quality or desirability. Time fixed effects, such as for year, week-of-year, or day-of-week, account for seasonality and trends in preferences over time. The parameter of interest is the WTP to avoid smoke, which is found by taking the ratio of marginal disutility of smoke ϕ to the marginal disutility of expenditure δ, WTP $= \phi/\delta$.

When making a reservation at an early date, person i knows their fixed unobserved preference for the site ε_{ijt}, but not what site conditions or the shock υ_{ijt} will be. Therefore, they choose based on ε_{ijt} and an expectation of V_{ijt}; in expectation, υ_{ijt} is 0. Denoting the expectation of variable Y at the time of reservation as $\bar{Y}$, write the expectation of V_{ijt} as

$$\bar{V}_{ijt} = \begin{cases} \delta^R c_{ijt} + \phi^R \bar{s}_{jt} + \bar{X}'_{jt}\gamma^R + \psi^R_{jt}, & j \in \{1, 2, ..., J\}; \\ 0, & j = 0. \end{cases} \tag{4}$$

Thus, at the time of reservation, an individual's expected utility from visiting site j at time t is $\bar{V}_{ijt} + \varepsilon_{ijt}$.

The reservation decision is estimated using a pooled zonal travel cost model; because of the zonal structure, this implicitly assumes a binary choice set for a representative individual from each zone. Let $R_{ijt} = 1$ if an individual chooses to make a reservation to visit site j at time t. Individuals choose $R_{ijt} = 1$ if the expected utility of doing so exceeds the utility of their outside option, the observable portion of which we normalize to 0. Under the assumption that ε_{ijt} is distributed iid type I extreme value, the probability an individual makes a reservation is given by

$$\mathbb{P}(R_{ijt} = 1) = \mathbb{P}(\varepsilon_{i0t} - \varepsilon_{ijt} \leq \bar{V}_{ijt}) = \frac{\exp(\bar{V}_{ijt})}{1 + \exp(\bar{V}_{ijt})}. \tag{5}$$

For each campground j and day t, we sum the number of reservers and non-reservers in concentric zones z around a campground. The former (denoted N^1_{zjt}) is counted based on the reservations in the Recreation.gov dataset; for example, a reservation for four people is counted as four reservers. The non-reservers (N^0_{zjt}) are determined based on zip code-level populations within each concentric ring, less the number of people from each zip code that held a reservation to a different campground on that day. The unit of observation for the zonal estimation is a campground by day by 50 km distance bin, where each row of data reports the number of people choosing outcome variable $R_{ijt} \in \{0, 1\}$. Let ω^R denote the set of parameters $\{\delta^R, \phi^R, \gamma^R, \psi^R_{jt}\}$. Summing the individual log-likelihood function over all reservers and non-reservers in each zone and for each campground and date, the log-likelihood function is given by

$$\ell^R(\omega^R) = \sum_{z=1}^{Z}\sum_{j=0}^{J}\sum_{t=1}^{T} \; N_{zjt}^1 \log\left(\mathbb{P}(R_{i(z),j,t} = 1|\omega^R)\right) + N_{zjt}^0 \log\left(1 - \mathbb{P}(R_{i(z),j,t} = 1|\omega^R)\right), \quad (6)$$

for Z zones, T choice occasions, and J sites. Maximizing Equation 6 yields utility parameters for a representative individual i from zone z.

We use a zonal approach for the reservation decision for several reasons. The primary purpose of this first-stage reservation estimation is to construct a control function that accounts for preferences in the cancellation estimation. A person can only cancel a trip if they held a reservation. Therefore, preferences from the reservation decision likely play a role in the cancellation. The zonal reservation model accounts for these preferences and provides computational advantages over a multinomial logit approach. In this setting, we have more than 1,200 arrival dates to define choice occasions, nearly 5 million reservations, nearly 5,400 zip codes to account for the non-reserving individuals, and 999 campgrounds to form the choice set. It would be infeasible to use all the data in a multinomial logit model of individual site choice. We could reduce the size of the dataset by, for example, restricting the study to a single region or year. However, smoke is temporally and spatially correlated within regions, meaning that we require multiple regions and years to provide necessary variation. The regional and temporal correlation of smoke also mean that site substitution is less likely to play a role in identifying the smoke parameter. Because we require regional and temporal variation, fixed effects are crucial to remove location- and time-specific unobservables across many heterogeneous sites. The zonal model accommodates a high number of fixed effects and is computationally less expensive than a contraction mapping method (Berry et al. 1995). This computational speed makes a difference when bootstrapping standard errors in the two-stage model.

3.2 Cancellations

For the second-stage cancellation decision, we model a binary choice at the level of the individual trip.[20] At time $t-\tau$, site conditions are realized and approximately known to the individual, resulting in random shock $e_{ijt} \equiv V_{ijt} - \bar{V}_{ijt} + \upsilon_{ijt}$ to the utility of site visitation. The shock e_{ijt} includes realization of the actual (V_{ijt}) as opposed to mean ($\bar{V}_{ijt}$) smoke and other weather conditions, as well as a random idiosyncratic and individual-specific taste shock (e.g., illness) represented by υ_{ijt}. Like ε_{ijt}, we assume υ_{ijt} is distributed type I extreme value. However, we allow its standard deviation to differ from the standard deviation of ε_{ijt} by a scale factor of $\frac{1}{\rho}$, letting $\upsilon_{ijt} \equiv \frac{1}{\rho}\eta_{ijt}$, where η_{ijt} is distributed standard type I extreme value.

Conditional on having a reservation, individual i follows through on their reservation if

$$\bar{V}_{ijt} + \varepsilon_{ijt} + e_{ijt} \geq \varepsilon_{i0t} + e_{i0t}. \tag{7}$$

Using the definition of e_{ijt}, and letting C_{ijt} be a binary variable equal to 1 if individual i cancels their reservation within τ days of arrival, the probability that i follows through is

$$\begin{aligned} \mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1) &= \mathbb{P}(V_{ijt} + \varepsilon_{ijt} + \upsilon_{ijt} \geq \varepsilon_{i0t} + \upsilon_{i0t}) \\ &= \mathbb{P}(\eta_{i0t} - \eta_{ijt} \leq \dot{V}_{ijt} - \rho(\varepsilon_{i0t} - \varepsilon_{ijt})), \end{aligned} \tag{8}$$

where the second line follows from the definition of υ_{ijt} and where $\dot{V}_{ijt} = \rho V_{ijt}$, a rescaling.

Equation 8 presents challenges for the econometrician. The variables ε_{i0t} and ε_{ijt} are unobserved. However, omitting these variables will bias parameter estimates because they are correlated with V_{ijt}, as only households with high taste for the site (ε_{ijt}) will have made a reservation. Specifically, ε_{ijt} is correlated with travel cost in the selected sample; for individuals with a high travel cost c_{ijt}, selection into the group of reservers implies a higher

[20] An alternative would have been to model a choice to cancel and rebook at another campground. However, very few users rebook after cancelling. In the estimating dataset, only 1.7 percent of all cancellations make a new reservation to a different campground for the same week of arrival as the initial reservation; in terms of later bookings, 12.5 percent of cancellations make a reservation for any new date within the next six months. Therefore, modeling cancellations as a binary decision is likely a reasonable representation of a visitor's choice.

ε_{ijt}. Without sample selection correction, this relationship downward biases estimates of the travel cost parameter δ in the cancellation decision and thus inflates estimates of WTP $=$ ϕ/δ. Section 3.4 explores this relationship using a numerical example. We show that the bias arises only when unobserved preferences from the first stage affect the second-stage decision ($\varepsilon_{i0t} - \varepsilon_{ijt} \neq 0$ in Equation 8) and when we can only observe the cancellation decision for the selected sample of reservers ($R_{ijt} = 1$).

3.3 Control function

To correct for this bias, we develop a control function approach (Wooldridge 2015). We begin by noting the conditional distribution of $(\varepsilon_{i0t} - \varepsilon_{ijt})$ in the selected sample of reservers. Let $f(\cdot)$ be the logistic density, $F(\cdot)$ the logistic distribution, and define $\tilde{\varepsilon}_{ijt} \equiv (\varepsilon_{i0t} - \varepsilon_{ijt})$. The conditional density of $\tilde{\varepsilon}_{ijt}$ is

$$\begin{aligned} f(\ \tilde{\varepsilon}_{ijt} \mid R_{ijt} = 1) &= f(\ \tilde{\varepsilon}_{ijt} \mid \tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt} - \bar{V}_{i0t}) \\ &= \frac{f(\tilde{\varepsilon}_{ijt}) \cdot \mathbb{1}\{\tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt} - \bar{V}_{i0t}\}}{F(\bar{V}_{ijt} - \bar{V}_{i0t})} \\ &= \frac{f(\tilde{\varepsilon}_{ijt}) \cdot \mathbb{1}\{\tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt}\}}{\mathbb{P}(R_{ijt} = 1)}, \end{aligned} \tag{9}$$

where the first line follows from the reservation condition in Equation 5, the second from the definition of a truncated density, and the third by noting that $V_{i0t} = 0$ and that $F(\bar{V}_{ijt}) = \mathbb{P}(R_{ijt} = 1)$.

An estimand for $\tilde{\varepsilon}_{ijt}$ is given by

$$
\begin{aligned}
\mathbb{E}[\ \tilde{\varepsilon}_{ijt} \mid \tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt}] &= \int_{-\infty}^{\infty} \tilde{\varepsilon}_{ijt}\ f(\tilde{\varepsilon}_{ijt} \mid \tilde{\varepsilon}_{ijt} \leq \bar{V}_{ijt})\ d\tilde{\varepsilon}_{ijt} \\
&= \frac{\int_{-\infty}^{\bar{V}_{ijt}} \tilde{\varepsilon}_{ijt}\ f(\tilde{\varepsilon}_{ijt})\ d\tilde{\varepsilon}_{ijt}}{\mathbb{P}(R_{ijt}=1)} \\
&= \frac{\bar{V}_{ijt} \cdot \frac{\exp(\bar{V}_{ijt})}{1+\exp(\bar{V}_{ijt})} - \log(1+\exp(\bar{V}_{ijt}))}{\mathbb{P}(R_{ijt}=1)} \\
&= \frac{\bar{V}_{ijt} \cdot \mathbb{P}(R_{ijt}=1) - I_{ijt}}{\mathbb{P}(R_{ijt}=1)} \\
&= \bar{V}_{ijt} - \frac{I_{ijt}}{\mathbb{P}(R_{ijt}=1)}.
\end{aligned} \tag{10}
$$

The first line follows from the definition of a conditional expectation, the second by substituting in Equation 9, the third by evaluating the definite integral, the fourth by substituting Equation 5 and defining $I_{ijt} \equiv \log(1+\exp(\bar{V}_{ijt}))$, and the last through simplification.

Equation 10 contains familiar terms. The $\bar{V}_{ijt}$ term gives the expected utility of the site choice from the reservation decision. The second term contains the inclusive value I_{ijt}, which is equivalent to the expected maximal utility a visitor could anticipate from holding the reservation, including the value of either the trip or the cancellation (Train 2009). The I_{ijt} term is scaled by the inverse of the probability that they would reserve at the site.

Including $\tilde{\varepsilon}_{ijt}$ as a control function in Equation 8 provides an estimate of the unobserved preferences of individual i and allows for unbiased estimation of the travel cost parameter in the cancellation problem. As travel cost is positively correlated with ε_{ijt}, we expect that it is negatively correlated with $\tilde{\varepsilon}_{ijt} \equiv (\varepsilon_{i0t} - \varepsilon_{ijt})$. We also expect a higher value of $\tilde{\varepsilon}_{ijt}$ to increase the likelihood of cancellation, as in Equation 8. In Section 3.4, we illustrate the bias correction of this control function through a numerical example.

Estimation of the cancellation decision proceeds through the following two-stage process. First, we estimate the parameters of the reservation decision $\mathbb{P}(R_{ijt} = 1)$ by maximizing a zonal log-likelihood function as in Equation 6, for reservations made earlier than $t - \tau$ and using expected site conditions. Then, we use the parameters to create a fitted value

$\hat{\hat{\varepsilon}}_{ijt}$ for every observed reservation. We substitute them into the trip-level equation for the cancellation decision, where each row of data is a trip with a dependent variable $C_{ijt} \in \{0, 1\}$ indicating whether the user cancelled. In this second stage, the independent variables in V_{ijt} use realized rather than expected site conditions, as users approximately know the site conditions close to the arrival date. Let ω denote the set of parameters $\{\delta, \phi, \gamma, \psi_{jt}\}$. For individual i, the log-likelihood function for the cancellation decision is:

$$\ell^C(\omega, \rho) = \sum_{i=1}^{N} \sum_{j=0}^{J} \sum_{t=1}^{T} (1 - C_{ijt}) \log \big(\mathbb{P}(C_{ijt} = 0 | \omega, \rho, R_{ijt} = 1)\big) + \\ C_{ijt} \log \big(1 - \mathbb{P}(C_{ijt} = 0 | \omega, \rho, R_{ijt} = 1)\big). \tag{11}$$

Because of the two-stage estimation, we bootstrap the main estimates to obtain appropriate standard errors (Cameron and Miller 2015; Wooldridge 2015).

3.4 Numerical example

To illustrate the source of bias in naïve cancellation estimates, and the ability of our control function to correct for this bias, we test our estimator on simulated data. We simulate 10,000 draws with $N = 100,000$ users (individuals i) who make sequential reservation and cancellation decisions. We assign each user random travel costs, smoke conditions, and unobserved preferences ε_{ijt} and υ_{ijt}. We assert an arbitrary true WTP to avoid smoke of $\phi/\delta = 2$. Appendix B provides additional details.

Table 1 summarizes results from the simulation and illustrates two dimensions of the identification challenge. First, we test the role of sample selection, using the fact that in simulated data, we can observe the counterfactual cancellation decisions of users who never held a reservation. In Column 1, we estimate WTP based on the full sample of reservers and non-reservers. With no sample selection (i.e. when the sample is not limited to reservers), our estimate approaches 2, the true WTP. In Column 2, the sample is limited to reservers, but we model cancellation and reservation decisions as independent, conditional on observables. That is, unobserved preferences affecting reservation decisions do not affect cancellation decisions, and second-stage unobserved preferences are equal to $\frac{1}{\rho}\eta$ rather than $\varepsilon_{it} + \frac{1}{\rho}\eta$.

Under this assumption, our estimate is again nearly equal to the true WTP. It is only in Column 3, when the sample is selected and preferences in the first and second stages are not conditionally independent, that WTP is biased. In Appendix B, we discuss how this bias operates through correlation between preferences and travel cost in the selected sample, which attenuates estimates of the travel cost parameter. In Column 4, we maintain both assumptions from Column 3 but also introduce our control function for $\tilde{\varepsilon}_{ijt}$. Across Monte Carlo simulations, the control function corrects the bias and includes the true WTP in the confidence interval. For a full treatment, refer to Appendix B.

Table 1: Numerical Example for 10,000 Simulations of Cancellation Estimation, Bias, and Bias Correction from $\tilde{\varepsilon}_{ijt}$ Control Function

	(1)	(2)	(3)	(4)
WTP	2.01**	2.01**	6.70	2.00**
	(0.14)	(0.15)	(8.52)	(0.25)
Sample	All users	Reservers	Reservers	Reservers
Second-stage errors	$\varepsilon_{ijt} + \frac{1}{\rho}\eta_{ijt}$	$\frac{1}{\rho}\eta_{ijt}$	$\varepsilon_{ijt} + \frac{1}{\rho}\eta_{ijt}$	$\varepsilon_{ijt} + \frac{1}{\rho}\eta_{ijt}$
Control function	No	No	No	Yes

Notes: True willingness to pay (WTP) = 2. WTP values are means (with standard errors in parentheses) across estimates from M = 10,000 simulations with N = 100,000. * $p < 0.05$, ** $p < 0.01$.

4 Estimation

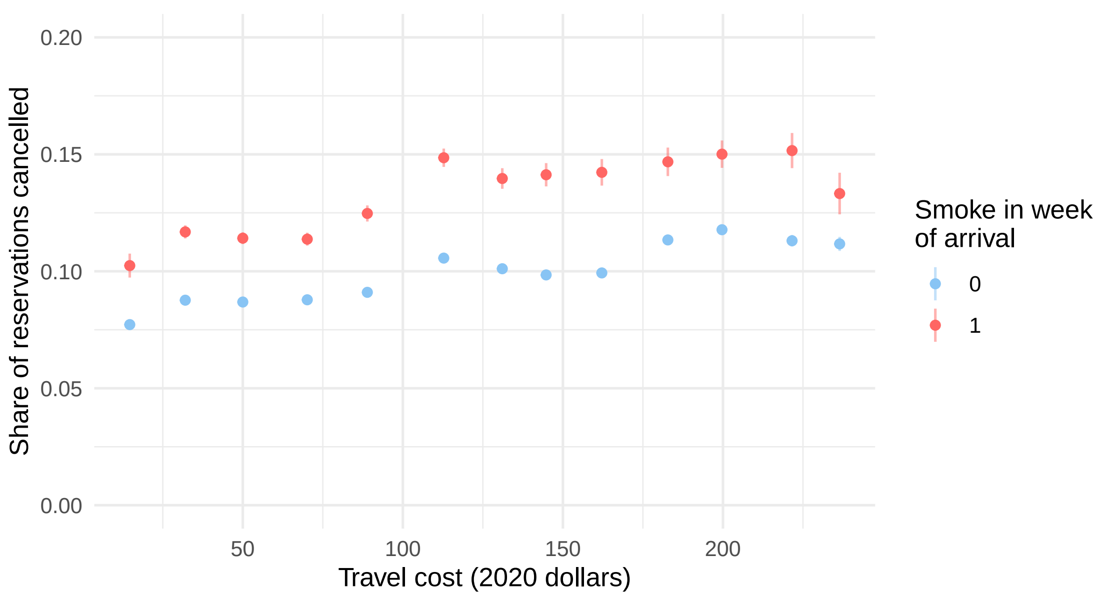
In this section, we estimate the welfare damages of wildfire smoke for outdoor recreation. As discussed in the previous section, the estimation follows a two-stage process that links reservations to cancellations close to arrival. Figure 2 shows the timing of decisions. We restrict the data to the set of users who booked more than a week ahead of time, or $\tau = 7$ in Figure 2, and who decided whether to cancel within a week of the arrival date. We therefore exclude reservations that were cancelled more than a week in advance. We also focus on trips scheduled for the months of May to September and over the years 2010 to 2017. Last, we limit attention to trips coming from within 650 km (400 miles), as described in Section 2.6.

These restrictions result in a sample of 2,723,940 reservations.[21]

4.1 Cancellations close to arrival

Figure 3 displays how the cancellation rate varies by travel cost and wildfire smoke conditions. The figure shows that users cancel their trips at higher rates during smoke conditions. This relationship does not appear to vary by travel cost, as the distance between the red and blue points is relatively constant across travel cost bins. Visually, the slope between cancellation rate and travel cost appears shallow. As explored in Section 3 and Appendix B, this shallow slope is likely due to positive correlation between travel cost and the unobserved preference parameter ε_{ijt} among the selected sample of reservers. Intuitively, if we were to observe someone reserve at site j despite a high travel cost, on average, they should have a higher preference ε_{ijt} for the site than someone with a similar travel cost who did not reserve, such that $\mathbb{E}[\varepsilon_{ijt} c_{ijt} | R_{ijt} = 1] > 0$. If ignored, we expect this correlation to depress the magnitude of the travel cost coefficient in the estimation of cancellations $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, which translates to a shallow slope in Figure 3.

Figure 3: Cancellation Rate Close to Arrival



[21] A "reservation" or "trip" is composed of multiple "transactions," which could include an initial booking, payment, check-in, cancellation, or refund.

Table 2: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ Within One Week, Uncorrected for Sample Selection

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.2195**	-0.2615**	-0.2346**	-0.2615**
	(0.0238)	(0.0283)	(0.0273)	(0.0215)
Travel cost (dollars)	-0.0024**	-0.0017**	-0.0017**	-0.0017**
	(0.0003)	(0.0001)	(0.0001)	(0.0001)
Inv. distance to wildfire (km^{-1})	-11.1276**	-12.0389**	-11.9174**	-7.8003**
	(0.9266)	(2.4288)	(2.4432)	(0.8291)
High temp. (degrees C)	0.0198**	0.0287**	0.0292**	0.0307**
	(0.0045)	(0.0023)	(0.0023)	(0.0022)
Low temp. (degrees C)	-0.0033	-0.0205**	-0.0214**	-0.0253**
	(0.0058)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0041**	-0.0058**	-0.0060**	-0.0057**
	(0.0011)	(0.0009)	(0.0009)	(0.0009)
N	2,723,830	2,692,468	2,692,468	2,689,216
WTP	91.1**	153.4**	137.35**	154.04**
	(12.36)	(21.06)	(19.85)	(15.43)
Campground x week-of-year FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Table 2 reports results for biased estimation of cancellations $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ using the trip-level maximum likelihood function of Equation 11. These estimates ignore the correlation between ε_{ijt} and travel cost among the set of users who chose to reserve. WTP is computed by taking the ratio between marginal disutility of smoke to that of expenditure (the smoke coefficient divided by the travel cost coefficient). Standard errors for WTP are computed using the delta method. In all estimations, the observations are weighted via frequency weights because a single reservation might represent, for example, two or eight visitors.

In Column 1, we display results without controlling for campground or seasonal fixed effects. Columns 2 through 4 add fixed effects. We include a campground-by-week fixed effect to control for unobserved location-specific site quality and seasonality, such as seasonal natural phenomena. We also account for differences in reservation rates based on the day of the week, as weekends see higher reservation activity. Lastly, we include various year fixed effects to account for time-related unobservables. Column 4 would imply that wildfire smoke causes \$154 in lost welfare per person per trip. This result is likely upward biased because WTP $= \phi/\delta$, and we expect the travel cost parameter δ to be attenuated.

To correct for the biased WTP in Table 2, we use the control function described in Equation 10, $\tilde{\varepsilon}_{ijt} = \bar{V}_{ijt} - \frac{I_{ijt}}{\mathbb{P}(R_{ijt}=1)}$. The first step is to estimate the probability of reservation earlier than one week based on expected site conditions and using a pooled zonal travel cost model. Then, we fit the parameters from the reservation estimation to form an estimate for $\tilde{\varepsilon}_{ijt}$. This estimate is used as a covariate in the trip-level estimation of cancellations $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$, after site conditions become approximately known to visitors.

We construct expected site conditions in the following way. For temperature and precipitation, we use climate normals from the PRISM data source, which represent average weather conditions for 1980 to 2010. Because travel cost is likely known to the individual ahead of time, we use the visitor's actual travel cost. For expected smoke and expected distance to fire, we use the average conditions over the past four years. For example, if a site was affected by smoke for one out of the past four years, we code expected smoke as 0.25.

Table 3 shows results from the first-stage reservation decision $\mathbb{P}(R_{ijt} = 1)$ implied by Equation 6. Users appear unexpectedly more likely to reserve at a campground with a higher expectation of wildfire smoke. Including more fixed effects generally decreases the magnitude and significance of the estimate, including moving the WTP closer to 0. Still, even with a high number of seasonal fixed effects, we may be unable to remove the correlation of seasonal variation in camping with wildfire smoke. Nevertheless, the primary purpose to estimate the likelihood of reservation $\mathbb{P}(R_{ijt} = 1)$ is as an input for the control function $\tilde{\varepsilon}_{ijt}$ in the estimation of $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$, so we should be unconcerned by the direction of the smoke expectation parameter.

After zonal estimation of $\mathbb{P}(R_{ijt} = 1)$ for early reservers, we use the parameter estimates to create fitted probabilities of reservation at the trip level. We expect that preferences and travel costs are correlated in the selected sample, $\mathbb{E}[c_{ijt}\varepsilon_{ijt}|R_{ijt} = 1] > 0$, so our control function should be inversely correlated with travel cost, $\mathbb{E}[c_{ijt}(\varepsilon_{i0t} - \varepsilon_{ijt})|R_{ijt} = 1] < 0$. Figure A6 in Appendix A illustrates this correlation empirically using the fitted values of

Table 3: $\mathbb{P}(R_{ijt} = 1)$ for Reservations Made Earlier Than One Week Based on Expected Site Conditions

	(1)	(2)	(3)	(4)
Smoke exp.	0.9260**	0.2513**	0.1032**	0.0822*
	(0.0036)	(0.0423)	(0.0363)	(0.0324)
Travel cost (dollars)	-0.0202**	-0.0244**	-0.0244**	-0.0244**
	(0.0000)	(0.0013)	(0.0013)	(0.0013)
Inv. distance to wildfire exp. (km^{-1})	39.6901**	6.0569**	6.1590**	6.7856**
	(0.0742)	(1.7451)	(1.5707)	(1.4654)
High temp. exp. (degrees C)	0.0191**	0.0597**	0.0611**	0.0588**
	(0.0001)	(0.0130)	(0.0131)	(0.0126)
Low temp. exp. (degrees C)	-0.0191**	-0.0818**	-0.0835**	-0.0812**
	(0.0001)	(0.0153)	(0.0153)	(0.0148)
Precip. exp. in week of arrival (mm)	-0.0126**	0.0071**	0.0066*	0.0067*
	(0.0001)	(0.0027)	(0.0027)	(0.0027)
N	15,209,187	12,668,366	12,668,366	12,298,572
WTP	-45.93**	-10.31**	-4.23**	-3.37*
	(0.18)	(1.72)	(1.45)	(1.31)
Campground x week-of-year FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

$\tilde{\varepsilon}_{ijt}$ and the travel cost for the sample of reservers. This empirical result is consistent with the prediction of our theory and numerical exercise in Section 3 and Appendix B.

Table 4 reports the main trip-level results for the cancellation estimation $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ using the bias-correcting control function $\tilde{\varepsilon}_{ijt}$. The coefficient for $\tilde{\varepsilon}_{ijt}$ is significant, suggesting that preferences at the time of reservation are an important determinant of the cancellation decision. In addition, comparing to Table 2, the travel cost coefficient was the only parameter to change when including $\tilde{\varepsilon}_{ijt}$, which is consistent with sample selection bias operating through correlation with travel cost. Overall, the WTP estimates are reduced to \$107 per person per trip of lost utility. By comparison, the biased results in Table 2 were \$154 per person per trip, which is 44 percent higher.

Wooldridge (2015) recommends bootstrapping standard errors for control functions be-

Table 4: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ Within One Week, Corrected for Sample Selection

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.2175**	-0.2708**	-0.2438**	-0.2603**
	(0.0247)	(0.0238)	(0.0221)	(0.0218)
Travel cost (dollars)	-0.0026**	-0.0024**	-0.0024**	-0.0025**
	(0.0004)	(0.0003)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-11.1017**	-10.8883**	-10.7067**	-7.8141**
	(0.8580)	(1.4280)	(1.4288)	(0.7920)
High temp. (degrees C)	0.0202**	0.0284**	0.0289**	0.0306**
	(0.0043)	(0.0024)	(0.0024)	(0.0023)
Low temp. (degrees C)	-0.0037	-0.0204**	-0.0214**	-0.0252**
	(0.0052)	(0.0026)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0041**	-0.0058**	-0.0060**	-0.0057**
	(0.0010)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0112	-0.0356**	-0.0366**	-0.0385**
	(0.0284)	(0.0106)	(0.0105)	(0.0106)
N	2,723,034	2,691,655	2,691,655	2,688,739
WTP	85.23**	113.91**	101.50**	107.14**
	(17.82)	(18.48)	(16.50)	(16.33)
Campground x week-of-year FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Bootstrapped std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

cause of the two-stage estimation process. For the main estimates in Table 4, we follow the clustered bootstrapping process of Cameron and Miller (2015), drawing with replacement at the campground level for 400 bootstraps. In Appendix C, we report results from Shapiro-Wilk tests for normality, failing to reject the null hypothesis that the bootstrapped smoke coefficients and travel cost coefficients are normally distributed. These tests suggest that 400 bootstraps are adequate for the analysis.

4.2 Relationship of damages to smoke duration

In this section, we investigate how estimated welfare losses vary according to one measure of the severity of smoke impacts: the number of smoke days in the week of arrival. We hypothesize that welfare impacts could vary with this measure for two reasons. First, it is likely that sites experiencing a greater number of smoke days in close succession also experience a larger degradation in air quality. Second, additional smoke days in the week of arrival communicate a higher likelihood of smoke on the actual arrival date.[22]

To investigate heterogeneity in estimated welfare impacts, we respecify $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ to allow for differential effects based on the number of smoke-affected days in the week of arrival. Figure 4 plots the resulting WTP estimates.[23] Damages monotonically increase in the number of smoke days, giving confidence that the estimates reflect damages from smoke. When a campground was affected by smoke on all seven days in the week of arrival, we find welfare damages of \$432 per person per trip. Estimated welfare impacts are likely greater during more smoke-affected weeks due to worse air quality conditions or because the perceived likelihood of smoke during the visit is greater.

4.3 Placebo test for smoke

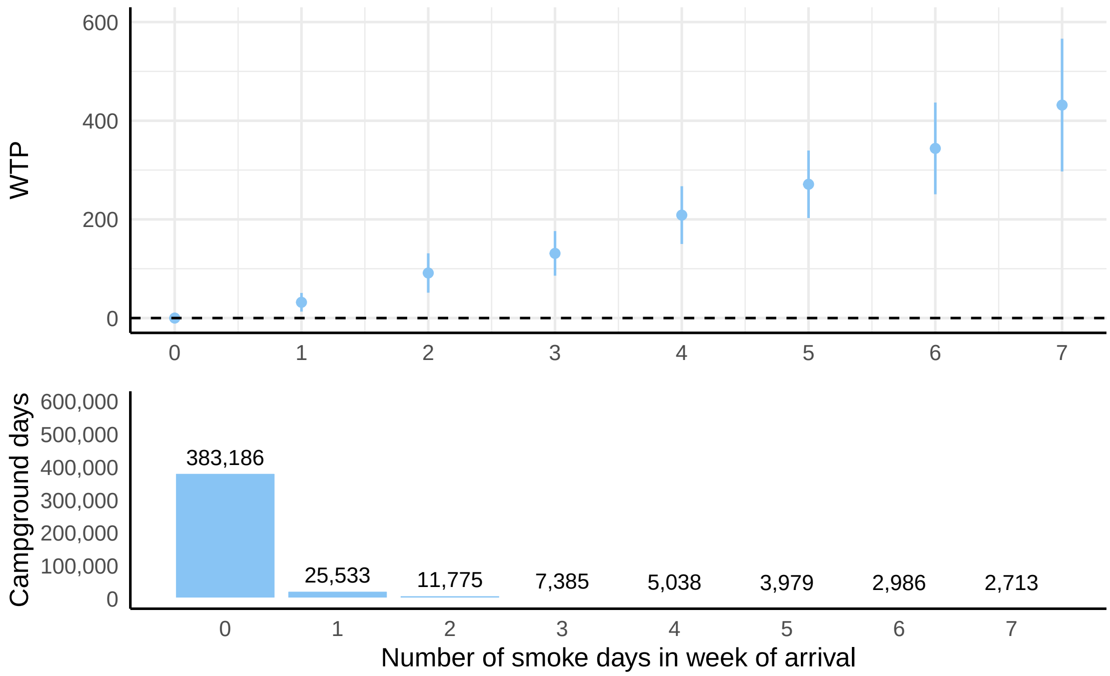
As a robustness check, we devise a placebo test to check whether the smoke coefficient actually measures responses to smoke. The placebo considers the responses of visitors whose campground was not affected by smoke until one or two weeks after their arrival. If visitors are truly averting recreation due to smoke, then we should see no response to these placebos. Of the 2.38 million reservations without smoke in the week of arrival, more than 375,000 placebo reservations are created.

Table 5 displays results from the placebo test. Across the main specifications, we find null responses to the two smoke placebos. Comparing to Table 4, most coefficients remain the same for this placebo test. This exercise should add confidence that individuals are actually responding to smoke in the main estimation.

[22] Table A1 in Appendix A reports results from a regression of $\mathbb{1}\{\text{campground is smoke-affected}\}$ on indicator variables for one, two, ..., seven days of smoke in the week before arrival. Each additional smoke day increases the probability of smoke on the actual day of arrival. For instance, compared to a week with no smoke days, a week with two days raises the probability by 0.301; a week with six days raises it by 0.739.

[23] Full results are included in Table A2 in Appendix A.

Figure 4: Greater Welfare Damages for Weeks That Were More Smoke-Affected, Consistent with Either More Severe Events or Increased Certainty of Smoke Conditions



4.4 Additional results and robustness checks

We report several additional results and robustness checks in the appendices. The first analysis assesses the role of no-shows in the estimation of cancellation probability. A second exercise varies the distance threshold that defines the sample restriction. Similarly, we test alternative temporal thresholds for τ when considering cancellations within τ days of arrival. Lastly, we show how results vary by the popularity of the recreation destination.

One concern when studying cancellations is whether an individual will simply not show up without formally cancelling. For most campgrounds, we do not observe whether an individual checks in. However, campers have an incentive to cancel their reservation. For cancellations made more than 24 hours before the arrival date, visitors are reimbursed for the full cost less a \$10 cancellation fee; when cancelling within 24 hours of arrival, they are reimbursed for the full trip less the \$10 fee and the price of the first night's stay. Still, we explore this question in Appendix D. For a small subset of campgrounds, we are able to observe no-shows. In this sample, we test recoding no-shows as arrivals, which would mirror

Table 5: Placebo Test for $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ Using Smoke Long After Arrival

	(1)	(2)	(3)	(4)
Smoke in week after arrival	0.0874**	0.0248	0.0226	0.0066
	(0.0190)	(0.0163)	(0.0155)	(0.0159)
Smoke two weeks after arrival	0.0783**	0.0041	0.0034	-0.0042
	(0.0231)	(0.0147)	(0.0146)	(0.0156)
Travel cost (dollars)	-0.0027**	-0.0025**	-0.0025**	-0.0025**
	(0.0004)	(0.0003)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-8.5921**	-7.3229**	-7.2413**	-5.2047**
	(0.8373)	(0.8998)	(0.8823)	(0.7795)
High temp. (degrees C)	0.0201**	0.0294**	0.0285**	0.0303**
	(0.0043)	(0.0022)	(0.0022)	(0.0022)
Low temp. (degrees C)	-0.0010	-0.0184**	-0.0184**	-0.0221**
	(0.0057)	(0.0026)	(0.0026)	(0.0025)
Precip. in week of arrival (mm)	-0.0039**	-0.0062**	-0.0062**	-0.0058**
	(0.0011)	(0.0009)	(0.0009)	(0.0010)
$\tilde{\varepsilon}_{ijt}$	-0.0089	-0.0332**	-0.0337**	-0.0352**
	(0.0262)	(0.0123)	(0.0122)	(0.0124)
N	2,379,842	2,344,620	2,344,620	2,340,894
Campground x week-of-year FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

what the data would report at campgrounds that do not record no-shows. This recoding does not change the estimates for the smoke or travel cost coefficients, suggesting unreported no-shows should not affect the results of the main estimates. For a discussion of this issue, see Appendix D.

We also explore alternative distance thresholds for the sample restriction. In the main results, we limit attention to reservations made within 650 km of one-way driving distance, or approximately 400 miles. Figure 1 shows that this distance restriction includes more than 85 percent of all reservations. Appendix E reports how estimates vary with this threshold. Increasing the distance threshold attenuates the parameter estimate for travel cost, which is an input to welfare calculation. This difference may be due to the inclusion of visitors traveling from greater distances, some of whom could be on multipurpose trips and were therefore less likely to cancel. As a result, the estimated welfare damages increase as the distance threshold is relaxed. For more information, see Appendix E.

In addition, we vary the threshold value for τ in Figure 2. The main estimation considers cancellation decisions within $\tau = 7$ days of arrival, but Appendix F reports results for alternative thresholds of 3, 5, and 9 days. As in the main estimation, the variable of interest is a binary variable equal to 1 for a smoke-affected day within τ days of arrival. Welfare estimates are larger for shorter bandwidths of τ. For $\tau = 3$, we find estimates of \$137 per person per trip; for $\tau = 9$, we find estimates of \$92 per person per trip. These results point to a similar mechanism as in Section 4.2. During shorter time windows, the occurrence of smoke communicates a greater likelihood of smoke on the actual arrival date. An information effect is consistent with greater smoke avoidance closer to the arrival date.

Lastly, we assess which types of campgrounds drive the parameter estimates. Even with a high number of fixed effects, visitors could respond differentially to disamenities at highly valued destinations, such as Glacier National Park, versus small, primarily local US Forest Service campgrounds. Appendix G reports heterogeneous results by campground popularity, where popularity is defined by annual visitation. We find that visitors are less responsive to both smoke and travel cost at the most popular destinations. Visitors may be more tolerant of environmental disamenities at highly valued destinations. Across most specifications, welfare damages are highest for destinations in the middle quartiles of popularity.

5 Total welfare losses

In the preceding sections, we estimated per-trip damages of wildfire smoke. We now turn to an appraisal of the total annual welfare damages for recreation. We combine the camping data from Recreation.gov with overall visitation data from federal and state agencies to determine the total number of outdoor visits in the West that are affected by smoke each year. As a back-of-the-envelope calculation, we multiply total smoke-affected visitation by the empirical per-trip welfare estimate to approximate the total annual welfare loss due to smoke. One limitation of this analysis is that the estimates are derived from camping activity, which may not be representative of losses to other forms of recreation, such as angling, swimming, or daytime visits. Still, this figure approximates the relative magnitude of total annual smoke damages for recreation in the western United States.

We find that across federal and state lands, an average of 21.5 million outdoor recreation visits per year are affected by wildfire smoke. Multiplying by a per-trip damage of \$107 per person, this result implies more than \$2.3 billion of welfare losses each year. This back-of-the-envelope estimate represents the lost welfare to inframarginal visitors and does not include the value of lost trips or substitution. Therefore, the estimate is conservative.

To arrive at this number, we use total visitation numbers from the National Park Service,[24] US Forest Service,[25] Bureau of Land Management,[26] US Army Corps of Engineers,[27] and National Association of State Park Directors (Smith et al. 2019) for 2008 to 2017. These data sources have varying levels of spatial and temporal granularity. For each data source, we use the Recreation.gov data to determine, at the relevant spatial and temporal scale, the proportion of total visits at each agency that were affected by smoke. For more information on the estimation of smoke-affected visitation, see Appendix H.

Table 6 displays estimates of total visitation, smoke-affected visitation, and total welfare losses by agency. One key point is the high overall level of outdoor recreation, with more than 511 million annual visits to state and federal lands in the western United States. In addition, a high proportion of these visits are affected by smoke. We estimate that approximately 21.5 million visits per year are smoke-affected, or 4.2 percent. When multiplied by the per-trip estimate of \$107, we find total annual welfare losses of approximately \$2.3 billion. Nearly half of these damages occur at state parks, which see larger visitation compared to federal agencies. Of any agency, the US Army Corps of Engineers saw the highest proportion of its visitors affected by smoke. This is likely because much of that agency's visitation (nearly 40 percent) occurs at lakes and reservoirs in the Pacific Northwest, a region that has seen particularly high wildfire smoke impacts relative to other regions (Burke et al. 2021; Gellman et al. 2022; Miller et al. 2021).

[24]National Park Service. Annual Summary Report. `https://irma.nps.gov/STATS`.

[25]US Forest Service. National Visitor Use Monitoring Program. `https://www.fs.usda.gov/about-agency/nvum`.

[26]Bureau of Land Management. Public Land Statistics. `https://www.blm.gov/about/data/public-land-statistics`.

[27]US Army Corps of Engineers. Value to the Nation. `https://www.iwr.usace.army.mil/Missions/Value-to-the-Nation`.

Welfare losses vary by region. Some states saw high smoke damages due to high baseline levels of visitation, and damages in other regions were driven by a high proportion of smoke-affected visits. Panel B of Table 6 reports losses by state. For states such as California and Colorado, damages are large due to high visitation. States such as Oregon and Washington saw both relatively high visitation and a high share of smoke-impacted visits. At the other end of the spectrum, states in the Southwest, such as Arizona, Nevada, and Utah, saw high visitation but a low proportion of smoke-affected visits. In Northern Rocky Mountain states, such as Idaho, Montana, and Wyoming, damages are driven by a high share of smoke-affected days despite lower total visitation. Figure A7 in Appendix A maps the results to contrast proportional versus total impacts. As a whole, these findings show the high cost of wildfire smoke for outdoor recreation in the western United States.

Table 6: Smoke-affected recreation visits and welfare losses for the western US, 2008 to 2017.

	Total visits/year (millions)	Smoke-affected visits/year (millions)	Proportion visits smoke-affected	Welfare loss/year (millions)
Panel A: Agency				
National Park Service	102.6	2.3	0.023	$248.1
US Forest Service	108.0	4.8	0.044	$511.4
Bureau of Land Management	59.8	2.5	0.042	$267.3
US Army Corps of Engineers	46.4	2.4	0.051	$251.7
State Parks	194.5	9.6	0.049	$1,022.1
Total	511.4	21.5	0.042	$2,300.6
Panel B: State				
Arizona	33.0	0.4	0.013	$46.3
California	162.7	6.1	0.037	$649.6
Colorado	55.3	2.1	0.037	$220.9
Idaho	19.6	1.3	0.065	$136.8
Montana	18.3	1.1	0.062	$121.8
Nevada	20.2	0.3	0.014	$29.8
New Mexico	13.7	0.6	0.041	$60.9
Oregon	69.9	4.4	0.062	$466.2
Utah	32.1	0.8	0.025	$84.8
Washington	64.8	3.3	0.051	$351.9
Wyoming	21.9	1.2	0.056	$131.6
Total	511.4	21.5	0.042	$2,300.6

Notes: Welfare losses computed by multiplying $107 per trip by smoke-affected visits.

6 Conclusion

This study provides the first revealed preference welfare estimates of the damage of wildfire smoke for outdoor recreation. Using high-frequency data on campground reservations, wildfire, smoke, and air pollution, we study avoidance behavior at federally managed lands in the western United States. We estimate that wildfire smoke causes welfare losses of $107 per person per trip. These damages increase at an increasing rate when campgrounds are affected by consecutive days of smoke, consistent with either greater smoke impacts or greater perceived likelihood of a smoke-affected visit. Combining these results with federal and state data on total visitation, we estimate that 21.5 million outdoor recreation visits per year are affected by smoke, with associated welfare losses of $2.3 billion.

This paper provides several contributions to the literature. First, we use novel methods and data. We value a temporary environmental bad, wildfire smoke, in a context where visitors face changing sets of information. We develop a two-stage decision structure that links preferences using a control function. This model draws on work from economists concerned with sample selection in nonlinear models (Greene 2012; Terza 2009) and researchers confronting sample selection in recreation settings (Cameron and DeShazo 2013; Cameron and Kolstoe 2022; Kolstoe and Cameron 2017; Lewis et al. 2019). The framework we develop could be used in other studies facing sample selection or sequential choices. Or, it could be used in applications of cancellations for intended visits (Whitehead et al. 2018). The model also relates strongly to prior work investigating environmental valuation when visitors have imperfect information about site quality (Leggett 2002); here, we explicitly model individuals' perceptions of likely environmental quality at the time of reservation. Lastly, our use of administrative data complements recent literature using large or innovative datasets to study recreation across multistate regions (Cameron and Kolstoe 2022; Dundas and von Haefen 2020; English et al. 2018; Parthum and Christensen 2022).

We also add to the literature on the costs of wildfire smoke. To contextualize the results of this study, we compare to several other studies on the costs of wildfire smoke. Most of these studies have used survey methods or healthcare costs, or valued changes in mortality using VSL. Richardson et al. (2012) report results from a survey following a large wildfire in Los Angeles County. They asked respondents about avoidance behavior during this fire, including expenditures on air purifiers, as well as health outcomes and risk perceptions. They derive WTP to avert one wildfire-induced symptom day of $84 in 2009 dollars. We estimate WTP to avoid an exposure day rather than a symptom day. Taking the empirical estimate of $107 per trip, this translates to approximately $38 per day based on an average trip of 2.84 days.

We can also compare the aggregated, total welfare results to the literature. We estimate welfare losses of approximately \$2.3 billion per year for recreation in the western United States. Miller et al. (2021) combined a VSL with estimates of mortality among elderly Medicare recipients due to wildfire smoke. They found between \$6 and \$170 billion in annual damages, in 2021 dollars. These results mainly vary due to assumptions on remaining years of life, as their sample is comprised largely of elderly individuals. When assuming that those who die from wildfire smoke would have lived an additional 3.5 years, they arrive at a lower bound of \$6 billion. Borgschulte et al. (2022) found annual lost labor earnings of \$125 billion per year, in 2018 dollars, due to wildfire smoke. Several other studies have found costs of wildfire smoke for test scores, crime, and hospital visits (Burkhardt et al. 2019; Heft-Neal et al. 2023; Wen and Burke 2022).

Estimating these costs can inform public policy. The federal government spends an average of \$2.8 billion per year on fire suppression, and the State of California spends \$900 million per year.[28,29] Wildfires destroy thousands of structures per year, which has cost tens of billions of dollars in recent years (Baylis and Boomhower 2022, 2023; Buechi et al. 2021). Both states and the federal government have pledged to increase fuel treatment projects to mitigate the risk of fire ignition and spread. California has jointly declared a goal with the US Forest Service to treat more than 1 million acres of hazardous vegetation per year[30] and proposed to spend \$1.2 billion across fiscal years 2022-23 and 2023-24 for fire mitigation activities, such as vegetation management and home hardening.[31] Understanding the cost of wildfires is crucial to assess the benefit of these public policies. Our study contributes to a growing understanding of the costs of wildfire smoke.

[28] National Interagency Fire Center. Suppression Costs. https://www.nifc.gov/fire-information/statistics/suppression-costs.

[29] California Department of Forestry and Fire Protection. Suppression Costs. https://www.fire.ca.gov/stats-events.

[30] Agreement for shared stewardship of California's forest and rangelands between the State of California and the USDA Forest Service, Pacific Southwest Region. https://www.gov.ca.gov/wp-content/uploads/2020/08/8.12.20-CA-Shared-Stewardship-MOU.pdf.

[31] California Legislative Analyst's Office. The 2022-23 Budget Wildfire and Forest Resilience Package. https://lao.ca.gov/Publications/Report/4495.

7 References

Abatzoglou, J.T., and A.P. Williams. 2016. Impact of Anthropogenic Climate Change on Wildfire Across Western US Forests. *Proceedings of the National Academy of Sciences* 113(42): 11770-75.

Aguilera, R., T. Corringham, A. Gershunov, and T. Benmarhnia. 2021. Wildfire Smoke Impacts Respiratory Health More Than Fine Particles from Other Sources: Observational Evidence from Southern California. *Nature Communications* 12(1): 1-8.

Baylis, P., and J. Boomhower. 2022. *Mandated vs. Voluntary Adaptation to Natural Disasters: The Case of US Wildfires* (No. w29621). National Bureau of Economic Research.

Baylis, P., and J. Boomhower. 2023. The Economic Incidence of Wildfire Suppression in the United States. *American Economic Journal: Applied Economics* 15(1): 442-73.

Berry, S., J. Levinsohn, and A. Pakes. 1994. Estimating Discrete-Choice Models of Product Differentiation. *Econometrica* 63(4), 841-890.

Borgschulte, M., D. Molitor, and E.Y. Zou. 2022. Air Pollution and the Labor Market: Evidence from Wildfire Smoke. *Review of Economics and Statistics*: 1-46.

Buechi, H., P. Weber, S. Heard, D. Cameron, and A.J. Plantinga. 2021. Long-Term Trends in Wildfire Damages in California. *International Journal of Wildland Fire* 30(10): 757-62.

Burke, M., A. Driscoll, S. Heft-Neal, J. Xue, J. Burney, and M. Wara. 2021. The Changing Risk and Burden of Wildfire in the United States. *Proceedings of the National Academy of Sciences*: 118(2).

Burke, M., S. Heft-Neal, J. Li, A. Driscoll, P.W Baylis, M. Stigler, J. Weill, and C. Gould. 2022. Exposures and Behavioral Responses to Wildfire Smoke. *Nature Human Behaviour* 6(10): 1351-61.

Burkhardt, J., J. Bayham, A. Wilson, E. Carter, J.D. Berman, K. O'Dell, B. Ford, E.V. Fischer, and J.R. Pierce. 2019. The Effect of Pollution on Crime: Evidence from Data on Particulate Matter and Ozone. *Journal of Environmental Economics and Management* 98: 102267.

Cameron, A.C., and D.L. Miller. 2015. A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources* 50(2): 317-72.

Cameron, T.A., and J.R. DeShazo. 2013. Demand for Health Risk Reductions. *Journal of Environmental Economics and Management* 65(1): 87-109.

Cameron, T.A., and S.H. Kolstoe. 2022. Using Auxiliary Population Samples for Sample-Selection Correction in Models Based on Crowd-Sourced Volunteered Geographic Information. *Land Economics* 98(1): 1-21.

Dundas, S.J., and R.H. von Haefen. 2020. 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): 209-42.

English, E., R.H. von Haefen, J. Herriges, C. Leggett, F. Lupi, K. McConnell, M. Welsh, A. Domanski, and N. Meade. 2018. Estimating the Value of Lost Recreation Days from the Deepwater Horizon Oil Spill. *Journal of Environmental Economics and Management* 91: 26-45.

Gellman, J., M. Walls, and M. Wibbenmeyer. 2022. Wildfire, Smoke, and Outdoor Recreation in the Western United States. *Forest Policy and Economics* 134: 102619.

Giglio, L., W. Schroeder, and C.O. Justice. 2016. The Collection 6 MODIS Active Fire Detection Algorithm and Fire Products. *Remote Sensing of the Environment* 178: 31–41.

Goldman, E., A, Wu, K. Kohn, and A, Bartuska. 2022. Visualizing Federal Funding for Wildfire Management and Response. Federation of American Scientists.

Greene, W.H. 2012. *Econometric Analysis* (Seventh edition). Pearson.

Hang, D., D. McFadden, K. Train, and K. Wise. (2016). Is Vehicle Depreciation a Component of Marginal Travel Cost?: A Literature Review and Empirical Analysis. *Journal of Transport Economics and Policy* 50(2): 132-150.

Heft-Neal, S., A. Driscoll, W. Yang, G. Shaw, and M. Burke. 2022. Associations Between Wildfire Smoke Exposure During Pregnancy and Risk of Preterm Birth in California. *Environmental Research* 203: 111872.

Heft-Neal, S., C.F. Gould, M. Childs, M.V. Kiang, K. Nadeau, M. Duggan, E. Bendavid, E., and M. Burke. 2023. Behavior Mediates the Health Effects of Extreme Wildfire Smoke Events (No. w30969). National Bureau of Economic Research.

Jones, B. 2017. Are We Underestimating the Economic Costs of Wildfire Smoke? An Investigation Using the Life Satisfaction Approach. *Journal of Forest Economics* 27: 80-90.

Kochi, I., G.H. Donovan, P.A., Champ, and J.B. Loomis. 2010. The Economic Cost of Adverse Health Effects from Wildfire-Smoke Exposure: A Review. *International Journal of Wildland Fire* 19(7): 803-17.

Kolstoe, S., and T.A. Cameron. 2017. The Non-Market Value of Birding Sites and the Marginal Value of Additional Species: Biodiversity in a Random Utility Model of Site Choice by eBird Members. *Ecological Economics*, 137, 1-12.

Korrick, S.A., L.M. Neas, D.W. Dockery, D.R. Gold, G.A. Allen, L.B. Hill, K.D. Kimball, B.A. Rosner, and F.E. Speizer. 1998. Effects of Ozone and Other Pollutants on the Pulmonary Function of Adult Hikers. *Environmental Health Perspectives* 106: 93–9.

Leggett, C. G. 2002. Environmental valuation with imperfect information the case of the random utility model. *Environmental and Resource Economics* 23(3): 343-355.

Lewis, D.J., S.J. Dundas, D. Kling, D.K. Lew, and S.D. Hacker. 2019. The Non-Market Benefits of Early and Partial Gains in Managing Threatened Salmon. *PLOS ONE* 14(8): e0220260.

Loureiro, M. L., M. Alló, and P. Coello. 2022. Hot in Twitter: Assessing the Emotional Impacts of Wildfires with Sentiment Analysis. *Ecological Economics* 200: 107502.

Lupi, F., D.J. Phaneuf, and R.H. von Haefen. 2020. Best Practices for Implementing Recreation Demand Models. *Review of Environmental Economics and Policy* 14(2): 302-23.

McCoy, S.J., and X. Zhao. 2021. Wildfire and Infant Health: A Geospatial Approach to Estimating the Health Impacts of Wildfire Smoke Exposure. *Applied Economics Letters* 28(1): 32-7.

Miller, N., D. Molitor, and E. Zou. 2021. A Causal Concentration-Response Function for Air Pollution: Evidence from Wildfire Smoke. Working paper, College of Business, University of Illinois, Urbana-Champaign.

Parsons, G.R. 2017. Travel Cost Models. In *A Primer on Nonmarket Valuation*, 2nd ed., edited by P.A. Champ, K.J. Boyle, and T.C. Brown, 269-329. New York: Springer.

Parthum, B., and P. Christensen. 2022. A Market for Snow: Modeling Winter Recreation Patterns Under Current and Future Climate. *Journal of Environmental Economics and Management* 113: 102637.

Preisler, H.K., D. Schweizer, R. Cisneros, T. Procter, M. Ruminski, and L. Tarnay. 2015. A Statistical Model for Determining Impact of Wildland Fires on Particulate Matter (PM2.5) in Central California Aided by Satellite Imagery of Smoke. *Environmental Pollution*: 205: 340-9.

Reid, C.E., M. Brauer, F.H. Johnston, M. Jerrett, J.R. Balmes, and C.T. Elliott. 2016. Critical Review of Health Impacts of Wildfire Smoke Exposure. *Environmental Health Perspectives* 124: 1334–43.

Richardson, L.A., P.A. Champ, and J.B. Loomis. 2012. The Hidden Cost of Wildfires: Economic Valuation of Health Effects of Wildfire Smoke Exposure in Southern California.

Journal of Forest Economics 18(1): 14-35.

Richardson, L.A., P.A. Champ, and J.B. Loomis. 2013. Valuing Morbidity from Wildfire Smoke Exposure: A Comparison of Revealed and Stated Preference Techniques. *Land Economics* 89(1): 76-100.

Schroeder, W., M. Ruminski, I. Csiszar, L. Giglio, E. Prins, C. Schmidt, and J. Morisette. 2008. Validation Analyses of an Operational Fire Monitoring Product: The Hazard Mapping System. *International Journal of Remote Sensing* 29(20): 6059-66.

Smith, J.W., E.J. Wilkins, and Y.F. Leung. 2019. Attendance Trends Threaten Future Operations of America's State Park Systems. *Proceedings of the National Academy of Sciences* 116(26): 12775-80.

Terza, J.V. 2009. Parametric Nonlinear Regression with Endogenous Switching. *Econometric Reviews* 28(6): 555-80.

Train, K.E. 2009. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Walls, M., C. Wichman, and K. Ankney. 2018. *Nature-Based Recreation: Understanding Campsite Reservations in National Parks*. Washington, DC: Resources for the Future.

Wen, J., and Burke, M. 2022. Lower Test Scores from Wildfire Smoke Exposure. *Nature Sustainability* 5(11): 947-55.

Westerling, A.L. 2016. Increasing Western US Forest Wildfire Activity: Sensitivity to Changes in the Timing of Spring. *Philosophical Transactions of the Royal Society B: Biological Sciences* 371(1696): 20150178.

Westerling, A.L. 2018. *Wildfire Simulations for California's Fourth Climate Change Assessment: Projecting Changes in Extreme Wildfire Events with a Warming Climate: A Report for California's Fourth Climate Change Assessment* (p. 57). Sacramento, CA: California Energy Commission.

Whitehead, J. C., T. Haab, S. L. Larkin, J. B. Loomis, S. Alvarez, and A. Ropicki. 2018. Estimating Lost Recreational Use Values of Visitors to Northwest Florida Due to the Deepwater Horizon Oil Spill Using Cancelled Trip Data. *Marine Resource Economics*, 33(2): 119-132.

Williams, A.P., J.T. Abatzoglou, A. Gershunov, J. Guzman-Morales, D.A. Bishop, J.K., Balch, and D.P. Lettenmaier. 2019. Observed Impacts of Anthropogenic Climate Change on Wildfire in California. *Earth's Future* 7(8): 892-910.

Wooldridge, J.M. 2015. Control Function Methods in Applied Econometrics. *Journal of Human Resources* 50(2): 420-45.

Appendix to "Welfare Losses from Wildfire Smoke: Evidence from Daily Outdoor Recreation Data"

Jacob Gellman, Matthew Wibbenmeyer, Margaret Walls

15 April 2024

Appendix A: Additional figures

Figure A1: Recreation.gov Web Interface

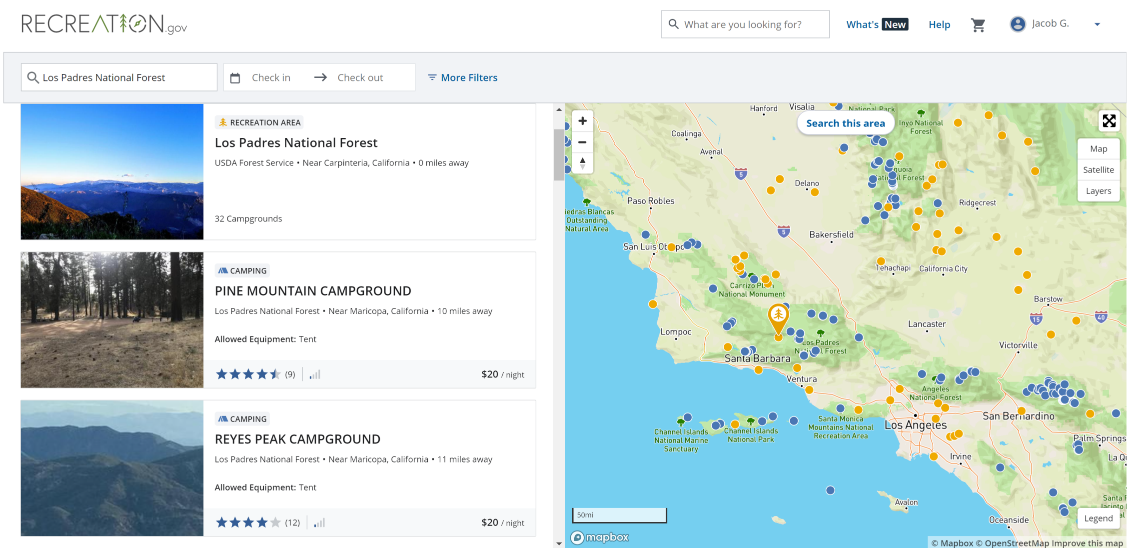


Figure A2: Automobile Route from Santa Barbara, California to Yosemite National Park

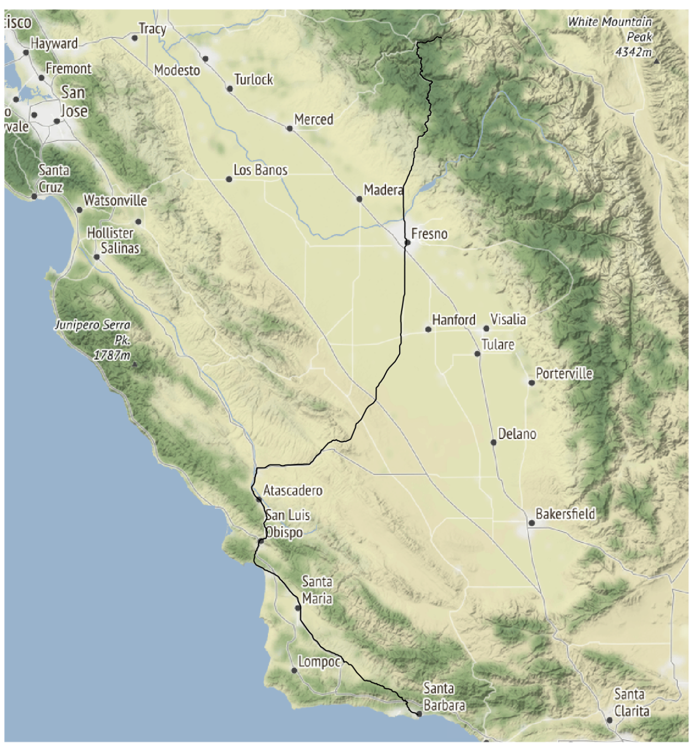
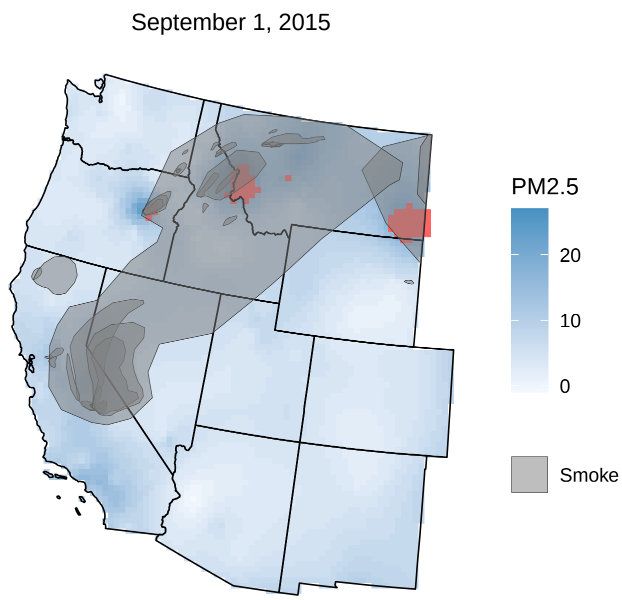


Figure A3: NOAA Smoke Plumes and $PM_{2.5}$



Notes: Red areas are affected by smoke and poor air quality.

Figure A4: Fire Detection Points and Fire Perimeters

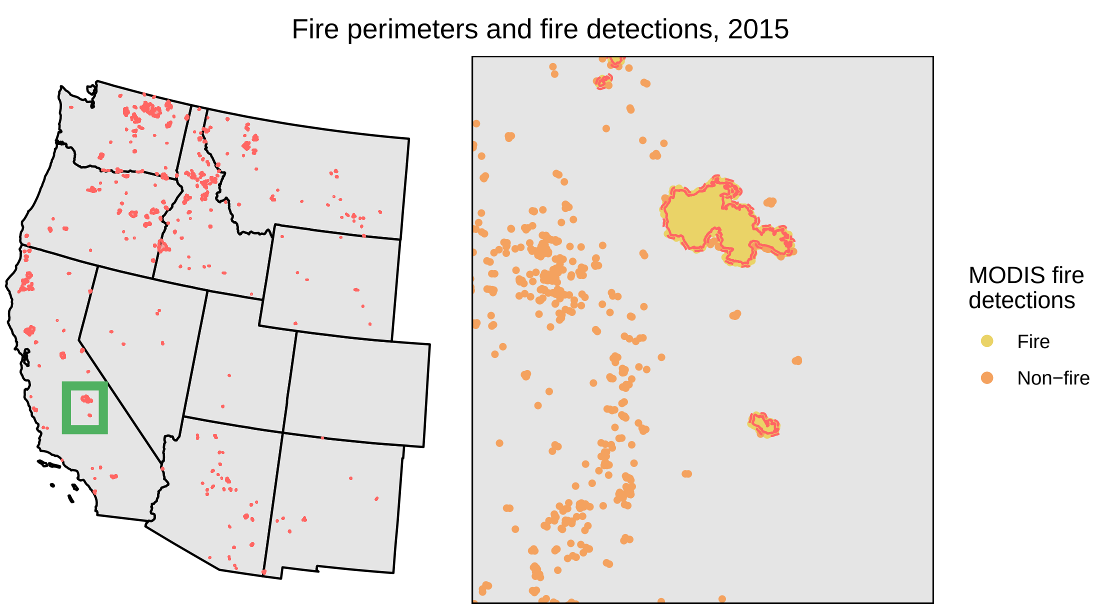


Figure A5: Map of Campgrounds in Dataset

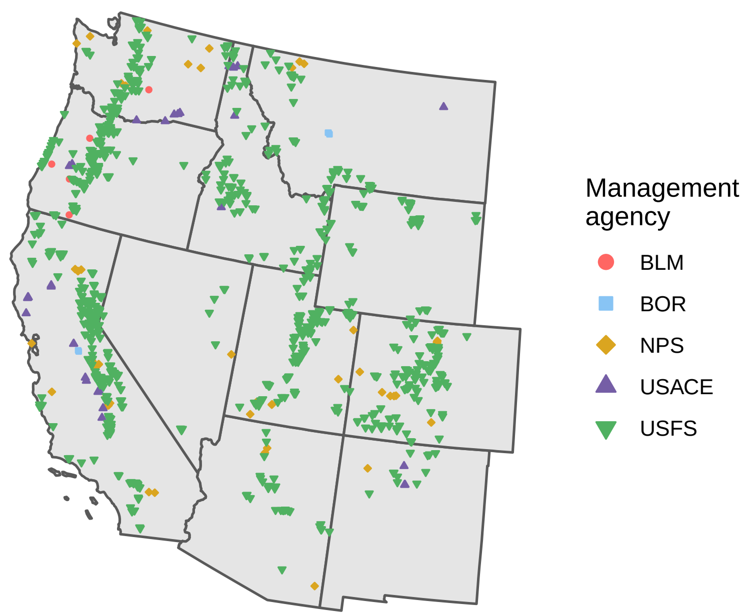


Figure A6: Relationship Between Control Function $\tilde{\varepsilon}_{ijt}$ and Travel Cost Using Model (4) of Table 4 Showing Correlation Between Preferences and Travel Cost in the Selected Sample of Reservers

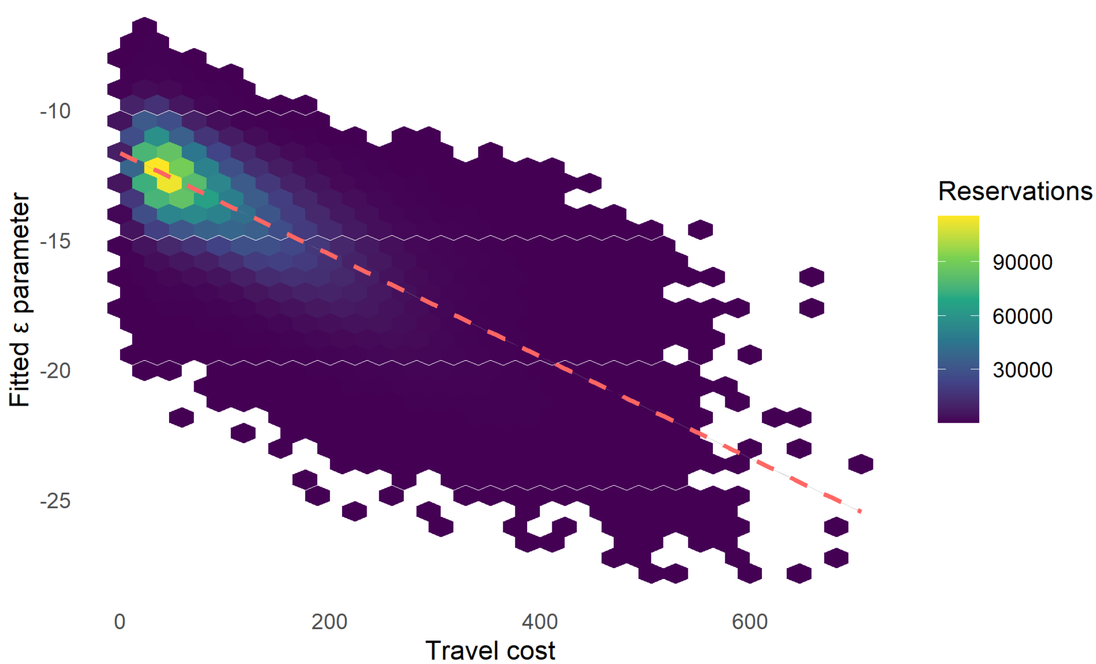


Figure A7: Total Estimated Welfare Losses and Proportion of Visits Affected by Smoke per Year

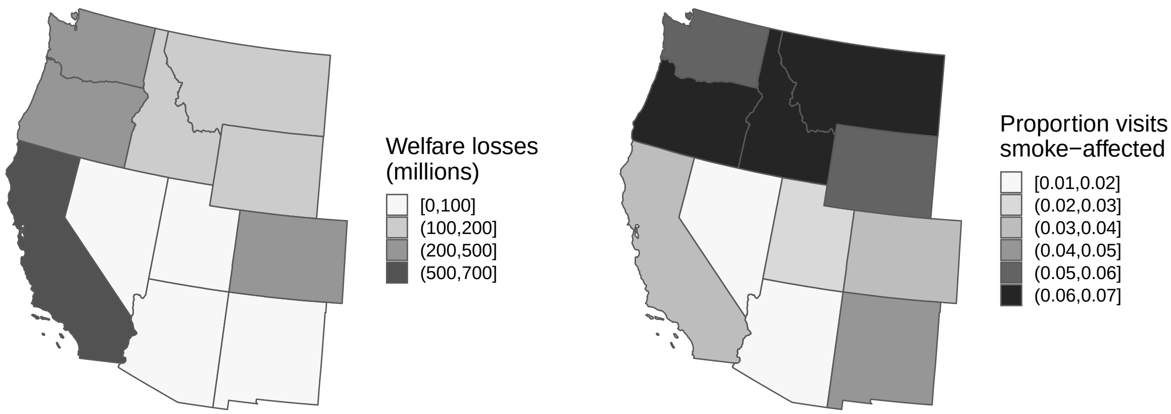


Table A1: Predicted Probability Campground Will Be Smoke-Affected, by Number of Smoke Days in the Preceding Week

	1{Campground is smoke-affected}
Intercept	0.0000
	(0.0001)
1{Smoke days in week of arrival = 1}	0.2013**
	(0.0005)
1{Smoke days in week of arrival = 2}	0.3009**
	(0.0008)
1{Smoke days in week of arrival = 3}	0.3798**
	(0.0010)
1{Smoke days in week of arrival = 4}	0.4850**
	(0.0012)
1{Smoke days in week of arrival = 5}	0.6108**
	(0.0013)
1{Smoke days in week of arrival = 6}	0.7390**
	(0.0016)
1{Smoke days in week of arrival = 7}	1.0000**
	(0.0018)
N	1,528,470

Notes: * $p < 0.05$, ** $p < 0.01$.

Table A2: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, Heterogeneity by Smoke Days in Week Before Arrival

	(1)	(2)	(3)	(4)
Travel cost (dollars)	-0.0025**	-0.0024**	-0.0024**	-0.0024**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-10.5524**	-11.6117**	-11.5449**	-7.4215**
	(0.9020)	(2.3964)	(2.4152)	(0.7843)
High temp. (degrees C)	0.0202**	0.0289**	0.0289**	0.0302**
	(0.0044)	(0.0023)	(0.0023)	(0.0021)
Low temp. (degrees C)	-0.0031	-0.0183**	-0.0189**	-0.0226**
	(0.0057)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0043**	-0.0060**	-0.0061**	-0.0057**
	(0.0011)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0027	-0.0342**	-0.0352**	-0.0368**
	(0.0255)	(0.0124)	(0.0124)	(0.0126)
Smoke days = 1	0.0158	-0.0718**	-0.0575*	-0.0776**
	(0.0268)	(0.0247)	(0.0246)	(0.0201)
Smoke days = 2	-0.1521**	-0.2164**	-0.1975**	-0.2217**
	(0.0436)	(0.0427)	(0.0416)	(0.0339)
Smoke days = 3	-0.2257**	-0.3050**	-0.2862**	-0.3182**
	(0.0410)	(0.0441)	(0.0437)	(0.0357)
Smoke days = 4	-0.4418**	-0.4792**	-0.4506**	-0.5066**
	(0.0472)	(0.0511)	(0.0502)	(0.0447)
Smoke days = 5	-0.5737**	-0.6032**	-0.5779**	-0.6583**
	(0.0448)	(0.0560)	(0.0551)	(0.0488)
Smoke days = 6	-0.7121**	-0.7612**	-0.7444**	-0.8348**
	(0.0603)	(0.0669)	(0.0669)	(0.0637)
Smoke days = 7	-1.0022**	-1.0065**	-0.9868**	-1.0481**
	(0.0660)	(0.0939)	(0.0922)	(0.0908)
WTP: 1 smoke day	-6.31	30.11*	23.87*	31.96**
	(10.45)	(11.92)	(11.38)	(9.66)
WTP: 2 smoke days	60.79**	90.8**	82.03**	91.32**
	(21.15)	(24.32)	(22.91)	(20.36)
WTP: 3 smoke days	90.26**	127.98**	118.9**	131.09**
	(22.07)	(27.66)	(26.04)	(23.07)
WTP: 4 smoke days	176.63**	201.07**	187.15**	208.68**
	(31.73)	(33.26)	(31.01)	(29.87)
WTP: 5 smoke days	229.38**	253.09**	240.06**	271.19**
	(38.74)	(39.86)	(36.75)	(34.92)
WTP: 6 smoke days	284.7**	319.4**	309.19**	343.87**
	(46.09)	(50.3)	(47.28)	(47.41)
WTP: 7 smoke days	400.7**	422.33**	409.86**	431.74**
	(59.86)	(73.41)	(68.21)	(68.64)
N	2,723,034	2,691,655	2,691,655	2,688,739
Campground x week-of-year FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Appendix B: Numerical example of sample selection correction

In Section 3.3, we propose a control function approach to account for unobserved preferences $\tilde{\varepsilon}_{ijt}$ that could bias estimation of $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ if omitted. In this appendix, we provide a numerical example to illustrate the source of this bias, its effect on estimation of willingness to pay (WTP), and our correction using a control function. We show that WTP is only biased when preferences for the reservation decisions influence the cancellation decision (i.e., given selection) and when the counterfactual cancellation decision of non-reservers is unobserved. Furthermore, the bias operates through correlation between preferences and travel cost: among the selected sample of reservers, those with a high travel cost tend to have had a high taste for the site. This relationship downward biases estimates of the travel cost parameter in the cancellation decision, inflating WTP estimates. Finally, we demonstrate bias correction using the control function for $\tilde{\varepsilon}_{ijt}$ given in Equation 10.

In this numerical example, we simulate the two-stage reservation and cancellation decision using a Monte Carlo of 10,000 random draws. For every iteration, we generate $N = 100,000$ users i, each with a spatial coordinate $(x, y) \in [0, 1] \times [0, 1]$, where x and y are distributed uniformly. In addition, we generate a single site j at a random coordinate $(x, y) \in [0, 1] \times [0, 1]$, where x and y are again distributed uniformly. User i's travel cost c_{ij} is given by the Euclidean distance from i to j.

Users who reserve far in advance maximize utility based on expected smoke conditions. Define the utility from the reservation as $U^R_{ij} = \alpha_j + \delta c_{ij} + \phi\mathbb{E}[s_j] + \varepsilon_{ij}$, where α_j is an intercept, c_{ij} is the travel cost, s_j denotes smoke conditions at the site, and ε_{ij} is the individual's preferences from reservation. We will assert arbitrarily that $\alpha_j = 1$, $\delta = -0.8$, and $\phi = -1.6$. Therefore, the true WTP is $\phi/\delta = 2$. Each user's site-specific preference values of ε_{i0} and ε_{ij} are drawn from a type I extreme value distribution. Based on the "time of visitation," expected smoke conditions $\mathbb{E}[s_j]$ are drawn for each user from $\{0.1, 0.2, 0.4\}$ with equal probability. Users will choose to reserve $R_{ij} = 1 \iff U^R_{ij} \geq U^R_{i0}$.

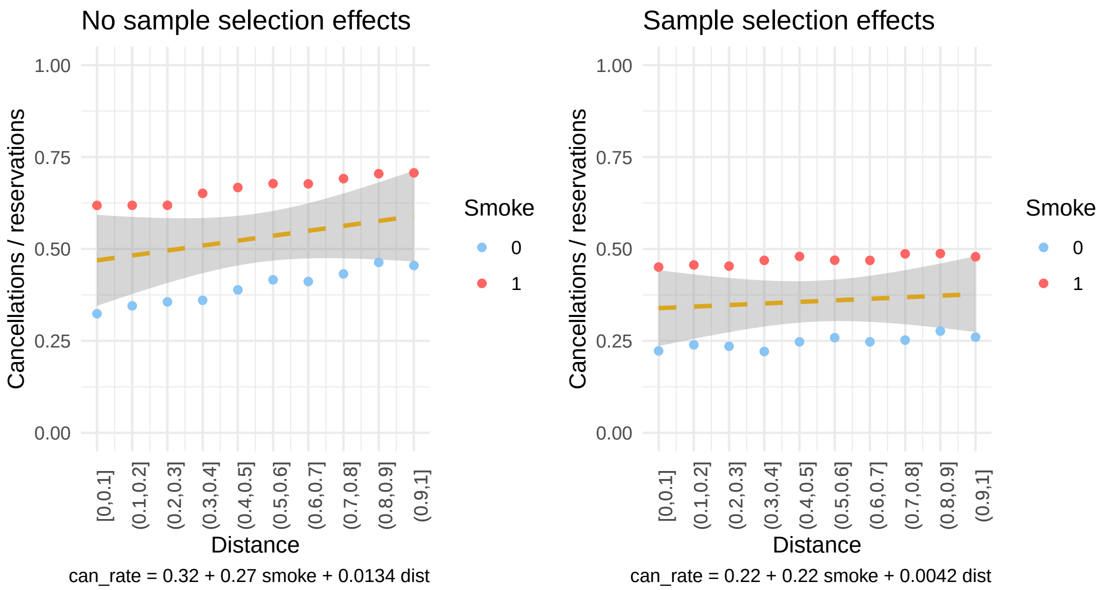
For the cancellation decision, the user decides based on realized smoke conditions. Let the utility from cancellation be $U^C_{ij} = \alpha_j + \delta c_{ij} + \phi s_j + \upsilon_{ij}$. Realized smoke s_j is drawn from $\{0, 1\}$ with $\mathbb{P}(s_j = 1) = 0.25$ for each user to create variation based on the "time of visitation."

We consider two types of error structures υ_{ij} in the cancellation decision. The first is an independent error, $\upsilon_{ij}^{ind} = \frac{1}{\rho}\eta_{ij}$, where $\eta_{ij} \sim$ type I extreme value, which assumes that the user's preferences in the decision are completely unrelated to their choice to have reserved. The second is a dependent error, $\upsilon_{ij}^{dep} = \varepsilon_{ij} + \frac{1}{\rho}\eta_{ij}$, which allows preferences at the time of reservation to affect the decision. We assume $\eta_{ij} \sim$ type I extreme value and arbitrarily set $\rho = 0.7$. Users will cancel $C_{ij} = 1 \iff U_{ij}^C \leq U_{i0}^C$. Because of the differing error structures, we consider two decisions under both υ_{ij}^{ind} and υ_{ij}^{dep}, which we denote as C_{ij}^{ind} and C_{ij}^{dep}, respectively.

The selection issue in the real recreation data arises because we can only observe the cancellation decision for reservers. However, under the Monte Carlo simulation, we can also examine the counterfactual decision of the non-reservers to see if they "would have" cancelled. We show that, even with a dependent error υ_{ij}^{dep}, estimation of $\mathbb{P}(C_{ij} = 0)$ on the full sample (reservers and non-reservers) without observing ε_{ij} will still recover the true WTP because no selection effect exists. That is, the biased estimation of $\mathbb{P}(C_{ij} = 0|R_{ij} = 1)$ is because ε_{ij} and c_{ij} are correlated in the selected sample, not the full sample.

Figure B1 illustrates the effects of selection by contrasting cancellation rates with and without selection effects, illustrating several key points. First, the overall cancellation rate is lower in the presence of selection, as indicated by the intercept of the fitted gold line. Users who made a reservation had a high initial preference for the site, so they are less likely overall to cancel. Second, the average effect of smoke, which is the distance between the red and blue points, is similar both with and without sample selection effects. Third, the effect of travel cost, which is the slope of the gold fitted line, is attenuated when preferences at the time of reservation affect the cancellation decision. This attenuation illustrates that the selection effect likely operates through positive correlation between ε_{ij} and travel cost.

Figure B1: Example Cancellation Rate to Illustrate the Effects of Selection



Next, we show that WTP estimates are only biased under a selected sample and when preferences at the time of reservation affect the cancellation decision. We estimate a logit regression for the reservation and cancellation decisions, varying whether we use the full sample or selected sample of reservers and the dependent error υ_{ij}^{dep} or the independent error υ_{ij}^{ind} for the cancellation decision.

Table B1 shows results from 10,000 draws. In Column 1, we use the full sample for the reservation decision. In Columns 2 and 3, we estimate the cancellation decision with both errors υ_{ij}^{ind} and υ_{ij}^{dep} but with the full sample. These regressions show that the selection effects would not cause biased estimation if the counterfactual cancellation decision of the non-reservers were known. In Column 4, we estimate the cancellation decision among only the selected sample but with an independent error υ_{ij}^{ind} (i.e., assuming no selection effects). Regression 4 demonstrates that sample selection is not an issue if the user's preferences at the time of reservation are unrelated to their cancellation decision. Biased estimation arises in Column 5 when there is both sample selection and when the reservation preferences affect the cancellation decision. Finally, Column 6 shows that inclusion of a control function corrects for the bias of Column 5. This result lends support to the use of this bias corrector in the empirical dataset.

Table B1: Monte Carlo 10,000 Simulated Regressions of Reservation and Cancellation Decisions

	(1)	(2)	(3)	(4)	(5)	(6)
WTP	2.00**	2.01**	2.01**	2.01**	6.70	2.00**
	(0.10)	(0.11)	(0.14)	(0.15)	(8.52)	(0.25)
Dep. var.	R_{ij}	C_{ij}	C_{ij}	C_{ij}	C_{ij}	C_{ij}
Users	All users	All users	All users	Reservers	Reservers	Reservers
Error	ε_{ij}	υ_{ij}^{ind}	υ_{ij}^{dep}	υ_{ij}^{ind}	υ_{ij}^{dep}	υ_{ij}^{dep}
Control function		No	No	No	No	Yes

Notes: True WTP = 2. * $p < 0.05$, ** $p < 0.01$.

Appendix C: Bootstrapped standard errors for $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$

In Section 4, we used a two-stage sample selection correction to estimate $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$. Wooldridge (2015) recommends that researchers bootstrap standard errors when estimating two-stage control functions. Because we cluster standard errors at the campground level, our bootstrap follows the process outlined by Cameron and Miller (2015) in a methods guide for clustered standard errors: for B bootstraps and G clusters, (1) sample with replacement G times from the original sample of clusters, and (2) compute parameter estimates. The estimating dataset contains $G = 999$ clusters. The resampling occurs over entire clusters; in some bootstraps, some clusters will not be represented, whereas some clusters will have all of their observations appear multiple times in the estimating dataset. Cameron and Miller (2015) note that $B = 400$ should be "more than adequate" in most settings.

In this section, we test that the bootstrapped coefficients follow a normal distribution, assessing whether $B = 400$ is adequate. Table C1 reports W statistics from Shapiro-Wilk tests of normality for the smoke and travel cost coefficients from the main estimation of Table 4. We fail to reject the null hypothesis that the bootstrapped smoke and travel cost coefficients follow a normal distribution. These tests imply that 400 bootstraps are adequate for the analysis. Figures C1 and C2 plot the bootstrapped coefficients visually.[32]

Table C1: W Statistics from Shapiro-Wilk Test of Normality for Bootstrapped Coefficients of $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ with Sample Selection Correction (parentheses indicate p values; the null hypothesis is that the coefficients are normally distributed)

	(1)	(2)	(3)	(4)
Smoke in week of arrival	0.996	0.998	0.998	0.994
	(0.450)	(0.979)	(0.852)	(0.084)
Travel cost (dollars)	0.990	0.996	0.995	0.995
	(0.006)	(0.343)	(0.291)	(0.255)
Campground x week-of-year FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

[32] One single iteration of the bootstrap produced abnormally large coefficient estimates, where, for example, the travel cost coefficient was estimated as -23 trillion. Still, that iteration produced a WTP in line with the other iterations; the WTP was equal to $74, suggesting that despite the abnormal magnitude of the coefficients, the ratios of coefficients to one another were proper. Nevertheless, we remove this single outlier from Figures C1 and C2, and from the analysis in the main text.

Figure C1: Distribution of Estimated Smoke Coefficient from Models (1) to (4) in Bootstrapped Estimation of $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ with Sample Selection Correction (red line indicates mean)

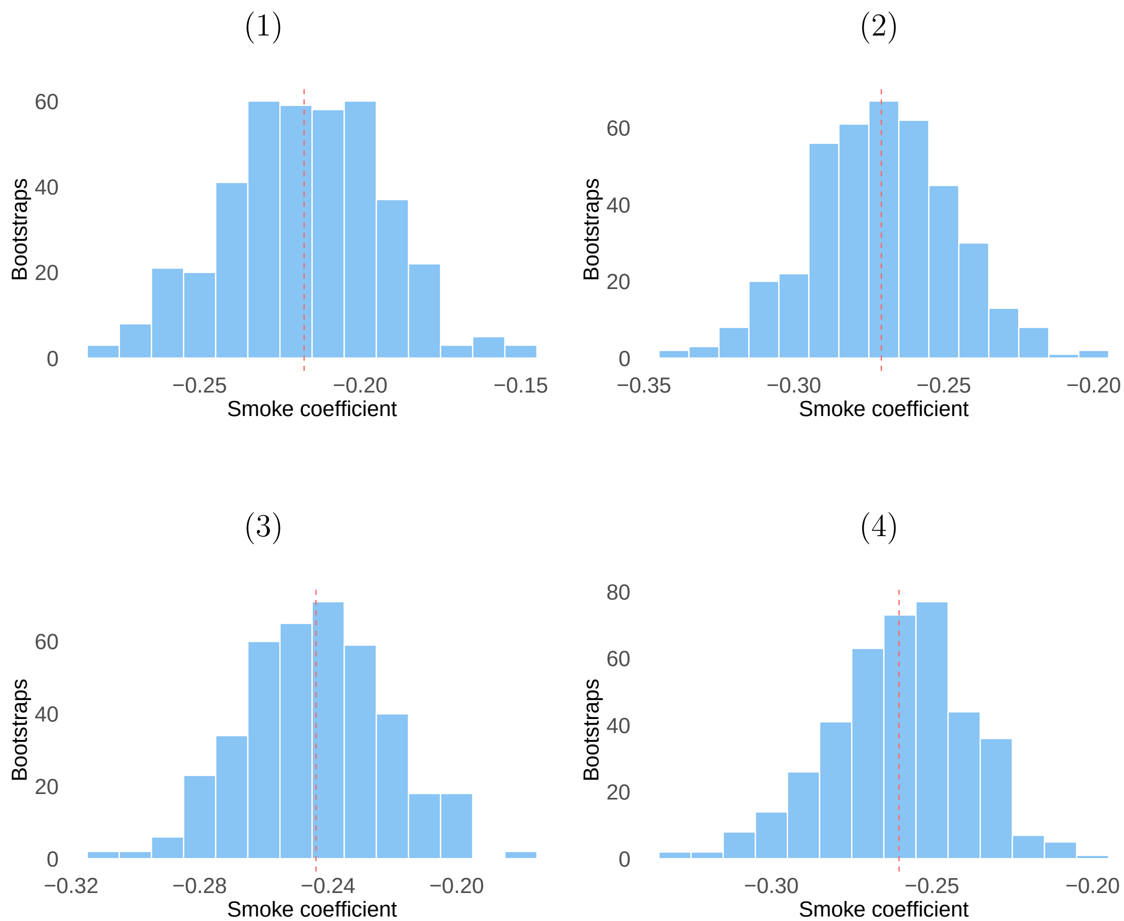
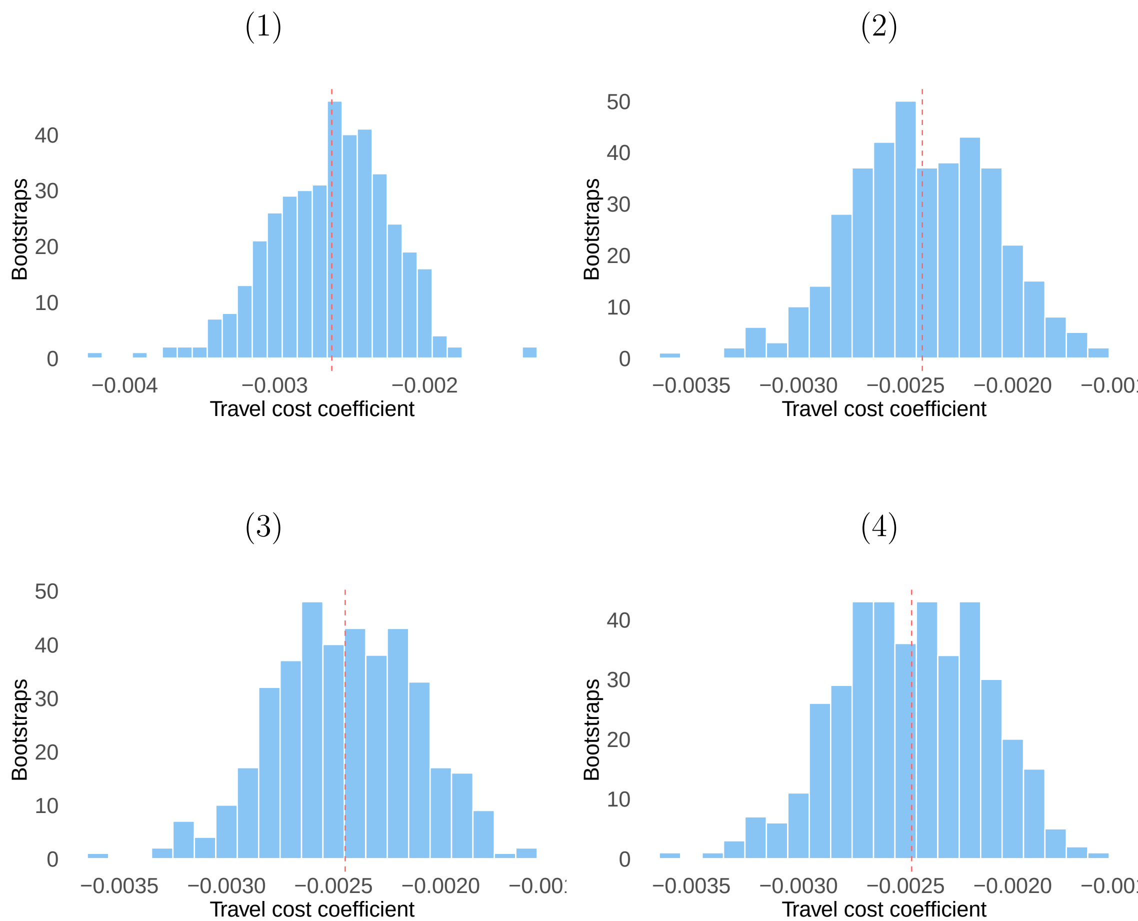


Figure C2: Distribution of Estimated Travel Cost Coefficient from Models (1) to (4) in Bootstrapped Estimation of $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ with Sample Selection Correction (red line indicates mean)



Appendix D: Testing the influence of no-shows in cancellations

One may be concerned that recreationists do not formally cancel their reservation when they decide not to complete a trip. Unreported no-shows threaten the identification of any willingness to pay (WTP) that is based on cancellations, as it could underestimate them. Although most of the campgrounds in the Recreation.gov dataset do not report check-ins or no-shows, a subset do. Just 36 out of 999 campgrounds (3.6 percent) report no-shows, but they represent 19.7 percent of the reservations used in the cancellation estimation. Overall, no-shows represent approximately 14.1 percent of all cancellations at the campgrounds which report no-shows.

As a robustness check, we recode no-shows at the reporting campgrounds to arrivals, which would reflect what the data would show at non-reporting campgrounds. If there is no difference in WTP when recoding no-shows, then it suggests that no-shows do not threaten the results of the main analysis.

Table D1 tests two models. Column 1 allows smoke and travel cost to respond differentially for no-show and non-no-show campgrounds. This model shows that no-show and non-no-show campgrounds have different overall measures of WTP. In Column 2 we recode no-shows as arrivals for the campgrounds that report no-shows. Comparing the WTP of no-show campgrounds with and without recoding (\$116.66 and \$118.44), WTP is virtually unchanged and is not statistically different. This analysis should alleviate concerns that no-shows influence the estimate of WTP in the full sample.

Table D1: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, Testing Effect of No-Shows on Cancellation by Recoding No-Shows as Arrivals

	(1)	(2)
Smoke x 1(Non-no-show campground)	-0.2608**	-0.2605**
	(0.0188)	(0.0188)
Smoke x 1(No-show campground)	-0.2628**	-0.2786**
	(0.0689)	(0.0730)
Travel cost x 1(Non-no-show campground)	-0.0025**	-0.0025**
	(0.0004)	(0.0004)
Travel cost x 1(No-show campground)	-0.0023**	-0.0024**
	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-7.8180**	-7.9904**
	(0.8223)	(0.8606)
High temp. (degrees C)	0.0306**	0.0306**
	(0.0022)	(0.0022)
Low temp. (degrees C)	-0.0252**	-0.0253**
	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0057**	-0.0058**
	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0376**	-0.0368**
	(0.0134)	(0.0136)
WTP, non-no-show campgrounds	103.82**	104.32**
	(17.39)	(17.8)
WTP, no-show campgrounds	116.66**	118.44**
	(32.1)	(33.75)
No-shows recoded as arrivals?	No	Yes
N	2,688,739	2,688,582
Campground x week-of-year FE	Yes	Yes
Day-of-week FE	Yes	Yes
Campground x year FE	Yes	Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$

Appendix E: Alternative distance thresholds for sample restriction

The main estimates of this paper restrict the estimating sample to reservations from origins within driving distance of a site, which we define as 650 km of one-way driving distance, or approximately 400 miles. Figure 1 shows that this threshold admits approximately 85 percent of the total reservations into the estimation. In this section, we show results from the main estimation using alternative distance thresholds of 350 km (approximately 217 miles) and 950 km (approximately 590 miles).

Table E1 illustrates how willingness to pay (WTP) estimates increase as the distance threshold is relaxed. Using a restrictive threshold of 350 km, WTP is estimated to be $79 per person per trip; with a wider threshold of 950 km, WTP is estimated to be $140 per person per trip. Column 2 reports the main estimates, identical to Table 4. The main estimate of $107 is included in the confidence interval of the 950 km estimate ([90.30, 189.36]), and falls narrowly outside the confidence interval for the more restrictive 350 km estimate ([51.26, 106.22]).

As the distance threshold is relaxed, the increasing WTP estimates are driven by a decline in the magnitude of the travel cost coefficient. In other words, increasing the pool of potential reservers decreases the estimated response to travel cost. This phenomenon could result from including visitors at greater distances who chose not to cancel their reservations. An additional difference across estimations is the magnitude of the coefficient for the $\tilde{\varepsilon}_{ijt}$ preference parameter. The magnitude is likely smaller at lower thresholds due to the correlation of preferences with travel cost; removing reservations from larger distances eliminates some visitors with both high travel costs and high tastes for the chosen site.

Table E1: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ within one week, restricting sample distance to within 350 km (217 miles), 650 km (404 miles), or 950 km (590 miles) of site

	350 km	650 km	950 km
Smoke in week of arrival	-0.2651**	-0.2603**	-0.2589**
	(0.0241)	(0.0218)	(0.0206)
Travel cost (dollars)	-0.0034**	-0.0025**	-0.0019**
	(0.0005)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-8.9377**	-7.8141**	-7.5899**
	(0.8670)	(0.7920)	(0.8101)
High temp. (degrees C)	0.0331**	0.0306**	0.0300**
	(0.0025)	(0.0023)	(0.0021)
Low temp. (degrees C)	-0.0248**	-0.0252**	-0.0249**
	(0.0028)	(0.0025)	(0.0024)
Precip. in week of arrival (mm)	-0.0061**	-0.0057**	-0.0056**
	(0.0010)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0240	-0.0385**	-0.0390**
	(0.0165)	(0.0106)	(0.0128)
N	2,044,062	2,688,739	2,851,414
WTP	78.74**	107.14**	139.83**
	(14.02)	(16.33)	(25.27)
Campground x week-of-year FE	Yes	Yes	Yes
Day-of-week FE	Yes	Yes	Yes
Campground x year FE	Yes	Yes	Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

Appendix F: Heterogeneous results within τ days of arrival

The main estimation considers the cancellation decisions of users within $\tau = 7$ days of arrival, where τ is defined in Figure 2. In this section, we explore heterogeneous results for alternative temporal thresholds. We reconstruct the dataset to estimate visitors' probability of cancellation within $\tau = 3$, 5, 7, and 9 days of arrival. The variable of interest is an indicator equal to 1 if a smoke-affected day occurred within the τ day threshold. We consider only standing, uncancelled reservations as of τ days before arrival.

Table F1 reports these results.[33] For $\tau = 3$, 5, 7, and 9 days, we find welfare damages of \$137, \$129, \$107, and \$92 per person per trip, respectively. These results are consistent with an information mechanism, which was explored in Section 4.2. For smaller values of τ, the occurrence of one smoke day corresponds to a greater likelihood of smoke on the actual day of arrival. Visitors may have a greater propensity to cancel when observing smoke closer to the date of arrival. The travel cost coefficient is largely stable as τ decreases; greater willingness to pay (WTP) is driven by a growth in the magnitude of the smoke coefficient. Overall, however, the main estimate of \$107 is included in the confidence interval of the estimates from each model.

Table F1: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ Within τ Days of Arrival

τ	3 days	5 days	7 days	9 days
Smoke	-0.2961**	-0.2929**	-0.2603**	-0.2281**
	(0.0536)	(0.0241)	(0.0218)	(0.0207)
Travel cost (dollars)	-0.0022**	-0.0023**	-0.0025**	-0.0025**
	(0.0004)	(0.0004)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-}1)	-12.2894*	-7.9469**	-7.8141**	-7.7057**
	(4.9610)	(0.8285)	(0.7920)	(0.8411)
High temp. (degrees C)	0.0384**	0.0343**	0.0306**	0.0284**
	(0.0026)	(0.0023)	(0.0023)	(0.0020)
Low temp. (degrees C)	-0.0310**	-0.0279**	-0.0252**	-0.0235**
	(0.0029)	(0.0026)	(0.0025)	(0.0023)
Precip. (mm)	-0.0052**	-0.0054**	-0.0057**	-0.0055**
	(0.0009)	(0.0008)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0291*	-0.0353**	-0.0385**	-0.0380**
	(0.0147)	(0.0136)	(0.0106)	(0.0126)
N	2,917,431	2,783,520	2,688,739	2,602,897
WTP	136.73**	128.9**	107.14**	91.69**
	(37.43)	(22.22)	(16.33)	(15.03)
Campground x week-of-year FE	Yes	Yes	Yes	Yes
Day-of-week FE	Yes	Yes	Yes	Yes
Campground x year FE	Yes	Yes	Yes	Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

[33] In addition, refer to Table 4 in the main text for results when $\tau = 7$.

Appendix G: Heterogeneous results by campground popularity

This section explores heterogeneous welfare damages based on the popularity of a campground. We define popularity based on the average number of visitors per year for years in which it was open. The most popular campgrounds tend to belong to high-profile national parks, such as Yosemite, Grand Canyon, and Rocky Mountain. The least popular tend to be small, local, or regional US Forest Service campgrounds. We rerun the main estimation but allow the smoke and travel cost coefficients to vary by the quartile of popularity. Given 999 campgrounds, each quartile contains approximately 250 locations.

Table G1 displays full results, including point estimates for smoke responses, travel cost responses, and willingness to pay (WTP). Across specifications, the magnitude for both the smoke and travel cost coefficients are lower at more popular campgrounds. These results suggest visitors are more willing to incur both higher travel costs and some environmental disamenity for highly desirable locations.

The translation of these responses to welfare impacts is less clear. WTP is estimated as the ratio of marginal disutility in smoke to that in expenditure (the smoke coefficient divided by the travel cost coefficient). Because WTP is a ratio, it could be either higher or lower given reductions in both the smoke parameter (the numerator) and the travel cost parameter (the denominator). Table G1 shows that the reduction in the smoke parameter dominates, resulting in lower WTP at popular campgrounds. In general, welfare damages tend to be largest for the middle two quartiles of popularity.

Appendix H: Total welfare estimate data construction

In Section 5, we report estimates for the total annual number of recreation visits affected by smoke in the west; we combine the Recreation.gov data with overall visitation data from various federal and state agencies. In particular, we use total visitation numbers from the National Park Service, US Forest Service, Bureau of Land Management, US Army Corps

Table G1: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, Heterogeneity by Campground Popularity

	(1)	(2)	(3)	(4)
Inv. distance to wildfire (km^{-1})	-11.0284**	-12.0844**	-11.9595**	-7.8196**
	(0.9225)	(2.4306)	(2.4448)	(0.8254)
High temp. (degrees C)	0.0187**	0.0289**	0.0293**	0.0307**
	(0.0044)	(0.0023)	(0.0023)	(0.0022)
Low temp. (degrees C)	-0.0012	-0.0204**	-0.0213**	-0.0252**
	(0.0055)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0046**	-0.0059**	-0.0061**	-0.0057**
	(0.0010)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0037	-0.0377**	-0.0387**	-0.0402**
	(0.0256)	(0.0120)	(0.0120)	(0.0122)
Smoke x first quartile (most popular)	-0.2208**	-0.2297**	-0.2035**	-0.2446**
	(0.0320)	(0.0345)	(0.0338)	(0.0286)
Smoke x second quartile	-0.2563**	-0.3296**	-0.3007**	-0.2915**
	(0.0425)	(0.0417)	(0.0407)	(0.0335)
Smoke x third quartile	-0.2364**	-0.3161**	-0.2889**	-0.3301**
	(0.0462)	(0.0482)	(0.0482)	(0.0482)
Smoke x fourth quartile (least popular)	-0.2488**	-0.3577**	-0.3457**	-0.2781**
	(0.0576)	(0.0673)	(0.0681)	(0.0743)
Travel cost x first quartile (most popular)	-0.0028**	-0.0023**	-0.0024**	-0.0024**
	(0.0004)	(0.0003)	(0.0003)	(0.0004)
Travel cost x second quartile	-0.0010*	-0.0027**	-0.0027**	-0.0026**
	(0.0005)	(0.0003)	(0.0003)	(0.0003)
Travel cost x third quartile	-0.0009	-0.0030**	-0.0030**	-0.0030**
	(0.0006)	(0.0004)	(0.0004)	(0.0004)
Travel cost x fourth quartile (least popular)	-0.0009	-0.0031**	-0.0031**	-0.0032**
	(0.0006)	(0.0005)	(0.0005)	(0.0005)
WTP: first quartile (most popular)	79.36**	98.63**	86.49**	102.87**
	(17.55)	(21.64)	(20.09)	(18.25)
WTP: second quartile	249.52	123.65**	112.36**	110.23**
	(136.76)	(22.99)	(21.6)	(19.31)
WTP: third quartile	253.55	106.16**	96.52**	108.62**
	(169.66)	(22.64)	(21.48)	(20.82)
WTP: fourth quartile (least popular)	272.18	116.18**	110.39**	87.8**
	(195.78)	(29.58)	(28.58)	(27.56)
N	2,723,034	2,691,655	2,691,655	2,688,739
Campground x week-of-year FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Notes: Std. err. clustered at campground level. * $p < 0.05$, ** $p < 0.01$.

of Engineers, and National Association of State Park Directors. Each reports visitation at varying spatial and temporal levels. For example, the National Park Service reports visitation at a park by month level; the US Forest Service reports at a forest by year level; and the state parks report at a state by year level. For each data source, we aggregate the daily Recreation.gov data to the most relevant spatial and temporal scale to determine the proportion of visits affected by smoke. We then multiply this proportion by the total visitation data. In this section, we detail this process for each data source.

For the National Park Service, we use the agency's Annual Summary Reports.[34] This dataset reports total monthly visitation at all national parks, national monuments, national recreation areas, and other lands that the agency manages. In the western states, 27 national parks are included in the Recreation.gov dataset, and 82 are not. For those 27 parks, we determine each park's monthly proportion of campers that were smoke-affected. We then multiply this proportion by each park's monthly visitation from the Annual Summary Reports to infer the total number of smoke-affected visits. For those 82 parks, we calculate a statewide proportion of smoke-affected campers in the data. We multiply these state by month proportions by each park's visitation levels in the Annual Summary Reports based on its location.

To estimate smoke-affected visits at national forests, we use the US Forest Service's National Visitor Use Monitoring (NVUM) Program.[35] These data report visitation at all National Forests at an annual level. In the west, 70 forests are included in the Recreation.gov dataset, and 8 are not. For the 70 forests in the Recreation.gov data, we calculate each forest's annual proportion of campers affected by smoke and multiply it by the corresponding annual visitation totals in the NVUM data. For the 8 forests not in the Recreation.gov dataset, we use a statewide annual proportion of smoke-affected campers.

The Bureau of Land Management records visitation statistics as part of its Recreation Management Information System (RMIS).[36] We contacted the program administrator and

[34] National Park Service. Annual Summary Report. `https://irma.nps.gov/STATS`.

[35] US Forest Service. National Visitor Use Monitoring Program. `https://www.fs.usda.gov/about-agency/nvum`.

[36] Bureau of Land Management. Public Land Statistics. `https://www.blm.gov/about/data/public-land-statistics`.

received data on site by year visitation for all BLM sites.[37] Most visitation is not reservable, and a large portion is considered backcountry. Therefore, the Recreation.gov dataset contains very few BLM campgrounds. We thus combine annual state level proportions of smoke-affected campers from the Recreation.gov data with annual site visitation from the RMIS.

For sites managed by the US Army Corps of Engineers, we use data from its Value to the Nation (VTN) reports.[38] For the study period of 2008 to 2017, the agency only has one year of recreation data, which is for the year 2016. We treat this year as representative of typical annual visitation over the study period. For each site, we multiply the total number of visitors by the state level average of smoke-affected campers from the Recreation.gov data over all years.

Lastly, we estimate smoke impacts at state parks. We use visitation data from the National Association of State Park Directors, which was compiled by Smith et al. (2019). For these data, the unit of observation is a state by year. We again use annual state level proportions of smoke-affected campers from the Recreation.gov data multiplied by the NASPD data.

Having approximated total visitation, we multiply each agency's annual smoke-affected visits by the empirical estimate of per-trip losses due to wildfire smoke. We estimate that more than 21.5 million recreation visits per year are affected by smoke in the west, with annual losses of $2.3 billion. For further discussion, see Section 5.

[37] Ridenhour, L. and Leitzinger, K. Bureau of Land Management. Personal correspondence.

[38] US Army Corps of Engineers. Value to the Nation. `https://www.iwr.usace.army.mil/Missions/Value-to-the-Nation`.