Non-market damages of wildfire smoke: evidence from administrative recreation data*

Jacob Gellman†, Margaret Walls‡, Matthew Wibbenmeyer§

10 November 2022

Job market paper: Click `here` for the latest version.

Abstract

Wildfire smoke pollution is growing in the western United States. Estimates of the health impacts of wildfire smoke are numerous, but there are few revealed preference estimates for the damages of smoke. Smoke is challenging to value with revealed preference methods because it is a transient environmental bad: it may blanket an area for several days before winds change or a fire is extinguished. In this paper we study a setting where individuals are directly exposed to wildfire 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. These data are rich among most studies of recreation, with nearly 1,000 campgrounds, a large geographic reach, detailed individual-level behavior, and high frequency daily data. The data allow us to model sequential recreation decisions under varying information, where preferences are correlated across decisions using a novel control function approach. We estimate that wildfire smoke reduces welfare by $107 per person per trip. The damage function is increasing and convex when campgrounds are affected by consecutive days of smoke, and is attenuated when smoke events are sufficiently far from active fires. In total, 21.5 million outdoor recreation visits in the western United States are affected by wildfire 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

*We thank Andrew Plantinga, Olivier Deschênes, Kelsey Jack, Max Moritz, Dave Lewis, Steve Dundas, Jeff Englin, and Molly Robertson for comments and suggestions on this paper. We are also grateful to seminar participants in the UCSB Environmental Economics group, the Western Economics Association 2022 Conference, the Association of Environmental and Resource Economists 2021 Summer Conference, 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 (USDA), National Institute of Food and Agriculture (NIFA) Agriculture and Food Research Initiative (AFRI) grant, award number 2020-67023-33258.

†University of California, Santa Barbara. Email: `gellman@ucsb.edu`.

‡Resources for the Future, Washington, DC. Email: `walls@rff.org`.

§Resources for the Future, Washington, DC, `wibbenmeyer@rff.org`.

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, Westerling 2018, Williams et al. 2019). Increased wildfire activity has also brought an increase in wildfire smoke, which can transport pollution hundreds of miles from the point of origin. The smoke produced by wildfires has large costs for society. Wildfire smoke now accounts for up to half of particulate matter pollution in some areas of the western United States (Burke et al. 2021a). Health damages from wildfire smoke are distinct from other air quality damages, as smoke harms health more severely than fine particles from other sources (Aguilera et al. 2021, Kochi et al. 2010). Negative health effects include increased morbidity, higher mortality, and reduced mental health (Cullen 2020, Heft-Neal et al. 2022, McCoy and Zhao 2020, Miller et al. 2021, Reid et al. 2016, Wen and Burke 2021).

While measures of health impacts are numerous, there are few revealed preference estimates for the welfare damages of smoke. In contrast to stated preference approaches such as surveys, revealed preference methods directly measure individuals' behavioral responses to an amenity or disamenity. Smoke is particularly challenging to study in a revealed preference setting because it is a transient environmental bad: it may blanket an area for several days before winds change or the fire is extinguished. Estimation of revealed preference values for wildfire smoke requires a context where individuals are both exposed to the environmental bad and where the researcher can observe their behavior at a high temporal resolution.

One setting where exposure is likely to be high is outdoor recreation. Researchers frequently use changes in outdoor recreation activity, such as camping and hiking, for revealed preference estimates of environmental amenities. Natural areas hold implicit non-market value due to the time and travel cost that people expend to visit them; the difference in a site's value across levels of an amenity identifies the amenity's value. Empirical measures of recreation value often form a large portion of natural resource appraisals or damage assessments (Phaneuf and Requate 2016). These estimates inform conservation decisions, natural resource management, and legal settlements for environmental accidents (English et al. 2018, Phaneuf and Requate 2016).

Wildfire smoke has numerous consequences for recreation. 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). This smoke puts visitors at an increased risk of respiratory health problems (Reid et al. 2016). Visitors to natural areas spend the majority of their trip outdoors, and vigorous activity such as hiking or rock climbing may exacerbate the effects of this direct exposure (Korrick et al. 1998, Richardson et al. 2012). Apart from health impacts, smoke may also reduce the visibility and amenity value from visitation. Visitors wary of air pollution and reduced site quality may avoid their trip altogether, with documented reductions in participation due to smoke (Cai 2021, Gellman et al. 2022). However, despite the numerous effects of wildfire smoke on outdoor visitation, most recreation data has not allowed for welfare estimation because it typically lacks the necessary temporal resolution to study avoidance behavior.

In this paper we provide the first revealed preference welfare estimates of the damage of wildfire smoke for outdoor recreation. We combine millions of administrative campground reservation records with satellite data on wildfire, smoke, and air pollution. These data are especially rich among most studies of recreation. The combined dataset features more than 16 million transactions from 2 million unique users, high frequency daily data over eight years, nearly 1,000 federally managed campgrounds across the west, and detailed records of individual-level behavior. The detail and variation afforded by this data are particularly necessary in our setting. Because wildfire smoke events affect large areas, campgrounds in a single region are often affected by smoke at exactly the same time, limiting the effectiveness of a region- or year-specific study. Our approach exploits daily variation across many years and many regions, which is necessary for proper identification of smoke effects.

We also account for the transient nature of wildfire smoke by measuring decisions both before and after visitors have knowledge of smoke conditions. Most visitors to campgrounds reserve their site several weeks or months in advance, before smoke conditions can be known. But, by the time smoke conditions are likely known, many campgrounds are either completely full or completely empty, which would limit the identifying variation from measuring new visitors. Our setting allows us to study one decision where visitors both have knowledge of site conditions and where they are unconstrained by congestion: cancellations of existing reservations due to smoke.

We consider visitors' cancellation decisions with a unique two stage discrete choice. In a first stage, a visitor chooses to reserve ahead of time based on expected site conditions; in a second stage, they decide whether to cancel or follow through with the reservation based on realized site conditions. A key feature of this setting is the correlation of preferences between the two decisions. A visitor can only cancel a trip if they previously made a reservation, meaning they have already demonstrated a taste for the site. Using a numerical example we show how failure to account for this form of sample selection would bias welfare estimates of wildfire smoke damages. The bias operates by attenuating estimates of visitors' marginal disutility in expenditure, a key input for welfare calculation. To correct for this bias we develop a novel control function approach to link preferences across choices (Wooldridge 2015), and demonstrate its effectiveness through numerical simulations.

There are several main findings of this paper. First, we estimate that wildfire smoke reduces welfare by $107 per person per trip. This estimate uses the aforementioned control function to account for sample selection. Without accounting for sample selection, the analysis would have implied damages of $154 per person per trip, which would overstate welfare damages by 44%. Inclusion of the control function increases the estimated magnitude of the marginal disutility in expenditure, as the numerical example predicts.

In further analyses we explore the shape of the damage function for wildfire smoke. Welfare losses vary by the duration of smoke events. When a campground was affected by smoke on only one day in the week of arrival, damages are as low as $32 per person per trip; when affected by smoke on all seven days in the week of arrival, losses are as high as $432 per person per trip. The damage function increases at an increasing rate, implying convexity of losses in the duration of smoke events. In addition, damages vary by proximity to active wildfires. Previous research has found that visitors to natural areas are less avoidant of smoke that originates from distant fires (Cai 2021). We find results that are in line with this research, with 20% lower welfare damages for smoke-affected campgrounds that are far from active fires. In general, the estimates are robust to multiple specifications, including a placebo for wildfire smoke. The placebo reassigns smoke events to either one or two weeks after the scheduled arrival date, testing whether visitors actually respond to wildfire smoke. We indeed find null results for this placebo, building confidence that the main estimates measure smoke responses.

The scale of wildfire smoke impacts for outdoor recreation is large. We combine estimates of the proportion of smoke-affected campers in our dataset, at park-specific and forest-specific levels, with total visitation data from state and federal agencies to determine the total number of outdoor visits affected by smoke each year. As a back of the envelope calculation we multiply this total smoke-affected visitation by the empirical per trip welfare estimate to approximate the total annual welfare loss due to smoke in the west. This analysis carries the limitation that welfare estimates are derived from camping activity, which may not be representative of other forms of outdoor recreation such as swimming, fishing, or day hiking. In addition, it accounts for lost welfare to inframarginal visitors and does not include the value of lost trips due to smoke. However, it provides an approximation of the magnitude of total annual smoke damages for recreation in the west. We find that an average of 21.5 million outdoor recreation visits per year 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 at state parks. A high proportion of outdoor trips are affected by smoke, at roughly 4.2% of the more than 511 million annual visits. Applying the empirical welfare estimate of $107 per person per trip, this figure implies welfare losses of roughly $2.3 billion per year due to wildfire smoke.

This paper makes several contributions. First, it adds to the literature on both market and non-market damages of wildfire smoke. To the best of our knowledge, this study is the first to directly value smoke damages using revealed preference methods and observational data. Existing non-market estimates have used survey methods, healthcare costs, or have applied the value of a statistical life (VSL) to changes in mortality. This paper's results complement these estimates. Richardson et al. (2012) surveyed individuals about self-protective expenditures following one fire in Los Angeles County. They found a willingness to pay (WTP) to reduce one wildfire smoke induced symptom day of $102.[1] We measure the value of an exposure day, rather than a symptom day; by comparison, this paper's estimate of $107 per trip roughly translates to $38 per day. The total cost of smoke for recreation, at $2.3 billion per year, is also informative for the literature. Miller et al. (2021) estimated the increase in mortality among elderly Medicare recipients in the United States, finding annual damages of $6 billion to $170 billion, depending on VSL assumptions. Borgschulte et al. (2022) found annual lost labor earnings of $125 billion per year due to wildfire smoke. Other studies have found costs of wildfire smoke for test scores, crime, and hospital visits (Burkhardt et al. 2019, Cullen 2020, Wen and Burke 2021). Aside from smoke, this paper complements the literature on the costs of wildfires more generally (Baylis and Boomhower 2021, Graff Zivin et al. 2020, Plantinga et al. 2022).

Measuring the cost of wildfire smoke is crucial to inform public policy. The federal government has spent an average of $2.8 billion per year on fire suppression over the period 2017 to 2021, while California has spent an average of $900 million per year from 2018 to 2022.[2,3] Proactively, California

[1] This figure is inflation-adjusted to 2020 dollars from the published estimate of $84 in 2009 dollars.

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

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

has proposed spending $1.2 billion over Fiscal Years 2022-23 and 2023-24 on fire mitigation measures including vegetation management, prescribed burns, home hardening, and related activities.[4] These activities are consistent with the state's recently declared goal to treat 1 million acres of hazardous fuels per year.[5] Understanding the averted costs of wildfire and smoke is critical to assess the benefit of these public expenditures.

This paper also contributes by using novel methods and data sources. We value a transient environmental good, wildfire smoke, in a setting where users make decisions under evolving sets of information. The two stage choice structure which links preferences across decisions is informed by literature on sample selection correction in non-linear models (Greene 2012, Terza 2009), as well as in 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 transient environmental amenities such as temperature, rainfall, 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 areas (Cameron and Kolstoe 2022, Dundas and von Haefen 2020, English et al. 2018).

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 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 west. Section 6 concludes.

2 Data

We combine data on recreation, wildfire smoke, air pollution, wildfire activity, and weather. We build 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 daily smoke, wildfire activity, pollution, weather, and climate normals at each campground. The second dataset is

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

[5] 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.

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 into travel cost zones around each campground to show daily reservation activity at various distance radii.

2.1 Recreation

We obtained data on campground use from Recreation.gov.[6] Recreation.gov is the web portal used to make reservations for 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 Recreation.gov 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.

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 of the transaction is known. For the 999 campgrounds in our analysis, 84% of transactions were made online, 9% over the phone, and 7% on-site (such as walk-ins or early checkouts).

2.2 Travel costs

We calculate travel costs using the distance and travel time between an origin zip code and a destination campground. We use GraphHopper, an open source routing engine which calculates routes using Djikstra's algorithm and OpenStreetMap data.[7,8] In total, 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 coordinates of each user's zip

[6]Recreation.gov. `https://www.recreation.gov`.
[7]GraphHopper. `https://www.graphhopper.com`.
[8]GraphHopper GitHub. `https://github.com/crazycapivara/graphhopper-r`.

code, we matched zip codes to Census Zip Code Tabulation Areas (ZCTAs) and found the centroid of each ZCTA.[9,10] Figure A2 in Appendix A displays an example automobile route.

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

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

for travel distance D_{zj} and travel time T_{zj}. The per-kilometer cost of traveling between ZCTA z and campground j is given by p_{zt}^D and includes costs of gasoline, per-kilometer vehicle maintenance costs, and per-kilometer average 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 nationwide average fleet fuel economy.[11,12] We use per-kilometer average depreciation and vehicle maintenance costs from AAA data, as in English et al. (2018).[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

For each day we record whether a campground was covered by wildfire smoke. We use daily observations of wildfire smoke from the National Oceanic and Atmospheric Administration (NOAA) Hazard Mapping System (HMS) by Schroeder et al. (2008).[14] Each day NOAA analysts manually trace the perimeters of wildfire 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. 2021a, Burke et al. 2021b, Burkhardt et al. 2019, Cullen 2020, Gellman et al. 2022, Heft-Neal et al. 2022, Miller et al. 2021, Preisler et al. 2015).

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

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

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

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

[13]For example: AAA. Your Driving Costs 2016. `https://jacobgellman.github.io/files/aaa/aaa_your_driving_costs_2016.pdf`.

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

One challenge presented by this dataset is that satellite photography does not reveal where in the air column a smoke plume is: the smoke could be at the ground level or it could be high in the atmosphere. If the plume is located high in the atmosphere it might not reflect on-the-ground conditions. To address this challenge we code an area as smoke-affected only if it is both covered by a smoke plume and if 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 example of that restriction using kriged $PM_{2.5}$ data from Burkhardt et al. (2019). The map shows that while 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. To measure wildfire activity 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 wildfires occurring in the United States. 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 final perimeter of the fire, as well as the start and containment dates of the fire, 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), maximum temperature (°C), 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 which reflect average conditions over the period 1980 to 2010.

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

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

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

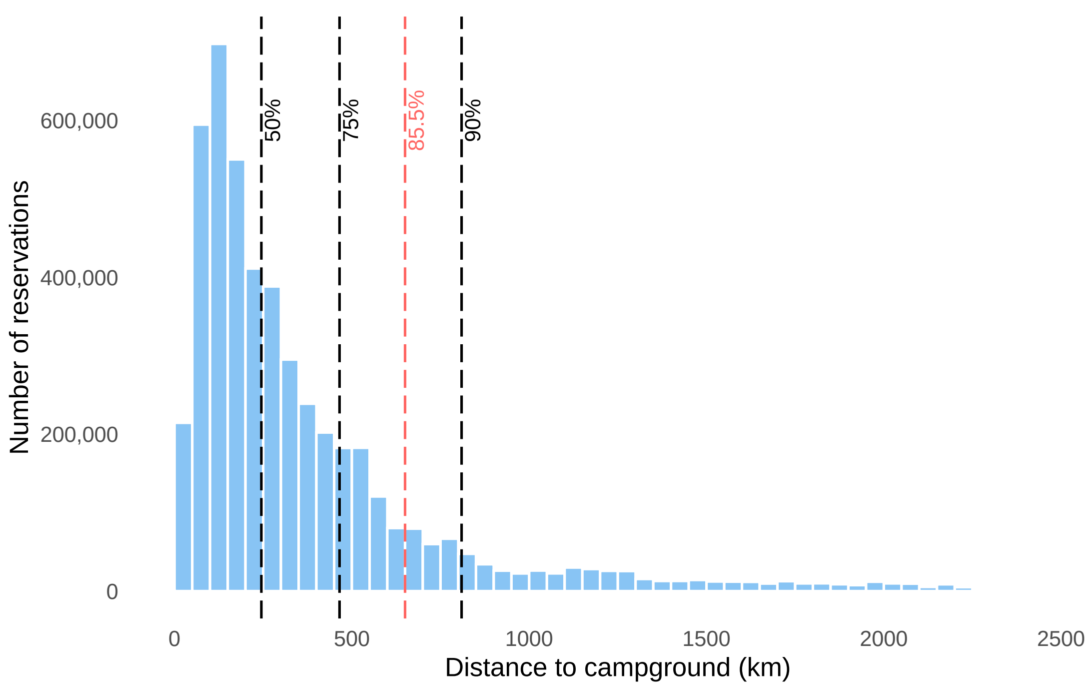
[18] Northwest Alliance for Computational Science and Engineering, Oregon State University. PRISM Climate Data. https://www.prism.oregonstate.edu.

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 analysis. While most of the campgrounds are managed by the Forest Service, the most-visited campgrounds tend to be National Parks. Table A1 in Appendix A reports the most-visited campgrounds in the dataset.

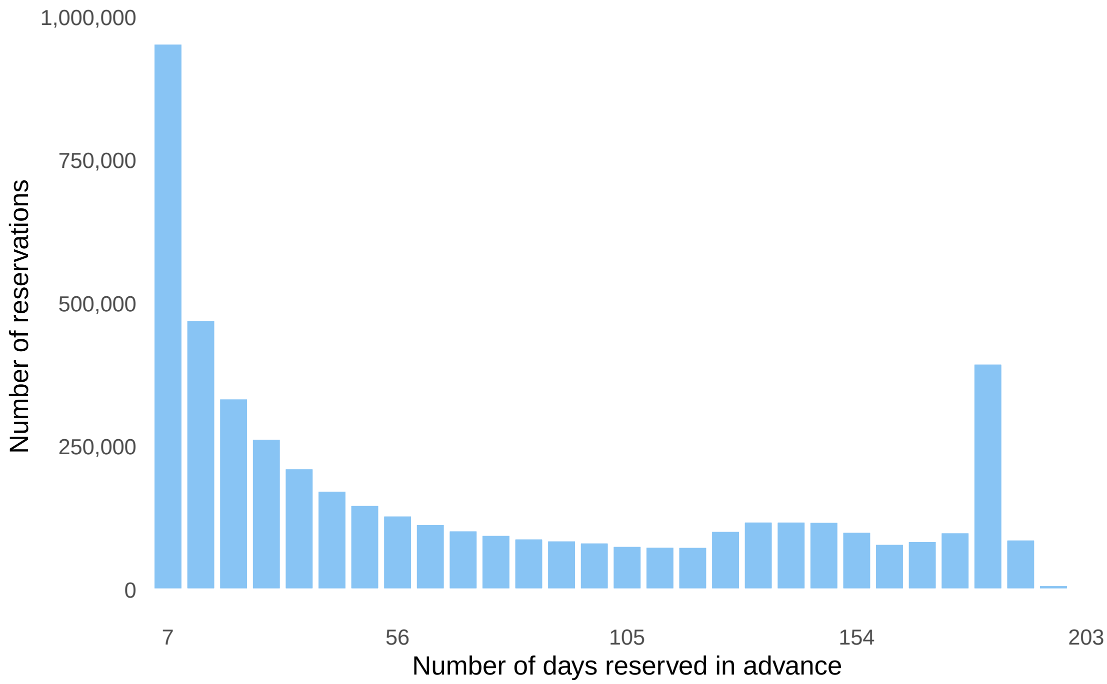
For the main analysis we restrict the set of potential reservers to residents living within one day's driving distance of a given campground. We set this restriction at 650 km (400 miles) of one-way driving distance. English et al. (2018) report survey results showing that, beyond 500 miles of driving distance, a substantial portion of recreation visitors are likely to have flown to their destination, which adds additional complications in the calculation of travel cost. Figure 1 shows that our 650 km restriction results in inclusion of more than 85% 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).

Figure 1: One-way driving distance of reservations from campground. Red line indicates 650 km cutoff.



The timing of a reservation is also key for our setting. Wildfire smoke is a random event, meaning that visitors who reserved far in advance could not have known that their chosen campground would be smoke-affected. Figure 2 shows that most visitors reserve far in advance of their arrival date, 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, there is significant mass 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.

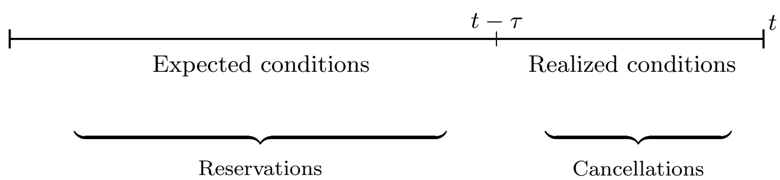
Figure 2: Days reserved in advance of arrival date.



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 for this setting is that smoke is ephemeral; it is a random event that may affect a site for several days before eventually disappearing. However, most campground reservations are made well in advance, before site conditions are known. But, by the time smoke conditions are likely known, many campgrounds are either completely full or completely empty, which would limit the identifying variation from measuring new visitors.[19] We therefore consider the cancellation decisions of visitors who reserved ahead of time.[20] We model a two part sequential process. In a first stage, visitors choose whether to reserve at a campground based on expected site conditions. In a second stage, close to the arrival date, they decide whether to cancel or follow through with the reservation based on realized site conditions. We allow for correlation of preferences across these decisions using a control function. For the reservation decision we use a pooled zonal travel cost model, which provides parameters for the control function in a trip-level model of cancellations. Figure 3 illustrates the timing of these decisions, where t gives the arrival date and τ denotes a bandwidth sufficiently close to the arrival date.

Figure 3: Timing of decisions.



3.1 Reservations

Define utility for the initial reservation decision as follows:

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

where V_{ijt} gives the indirect utility of person i for campground j on arrival date t. The variable ε_{ijt} represents preferences known to the individual but unobserved by the researcher. Define the observable portion of utility as:

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

The variable c_{ijt} gives the travel cost for person i to site j at time t, while s_{jt} is equal to 1 if there are smoke conditions at campground j on date t. The vector X_{jt} contains campground-level conditions including precipitation, temperature, and proximity to an active wildfire.

[19]See Figure A6 in Appendix A.

[20]We could have also considered new reservations close to arrival. For a discussion of late reservers, see Appendix B.

Additional variables include the alternative-specific constant ψ_j, which denotes campground-specific, time-invariant traits such as the quality of a campground. The variable λ_t similarly captures time-specific factors including seasonality and yearly trends. The parameter of interest is the willingness to pay (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 $= \phi/\delta$.

To operationalize our model, we estimate the reservation decision using a pooled zonal travel cost model. For each campground j and each day t, we sum the number of reservers and non-reservers in concentric zones z around a campground. Consider a representative agent in zone z. Because the agent reserves in advance, before site conditions are known, they compare utility over expected site conditions, $\mathbb{E}[U_{ijt}]$, to the expected utility of the outside option, $\mathbb{E}[U_{i0t}]$. Denote $R_{ijt} = 1$ if the individual chooses to reserve at campground j for arrival date t. When ε_{ijt} is distributed iid type I extreme value, the probability of observing $R_{ijt} = 1$ is given by:

$$\begin{aligned}
\mathbb{P}(R_{ijt} = 1) &= \mathbb{P}(\mathbb{E}[U_{ijt}] \geq \mathbb{E}[U_{i0t}]) \\
&= \mathbb{P}(\mathbb{E}[V_{ijt} + \varepsilon_{ijt}] \geq \mathbb{E}[V_{i0t} + \varepsilon_{i0t}]) \\
&= \mathbb{P}(\mathbb{E}[V_{ijt}] + \varepsilon_{ijt} \geq 0 + \varepsilon_{i0t}) \\
&= \mathbb{P}(\varepsilon_{i0t} - \varepsilon_{ijt} \leq \mathbb{E}[V_{ijt}]) \\
&= \frac{\exp(\mathbb{E}[V_{ijt}])}{1 + \exp(\mathbb{E}[V_{ijt}])}, \qquad (4)
\end{aligned}$$

noting that $V_{i0t} = 0$, and that ε_{ijt} is non-random from the perspective of the individual such that $\mathbb{E}[\varepsilon_{ijt}] = \varepsilon_{ijt}$.

The reservers are 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 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 $R_{ijt} \in \{0, 1\}$. For example, on August 1, 2015, Diamond Lake in Oregon saw 49 reservers ($R_{ijt} = 1$) from distance bin $(350, 400]$ and approximately 2.7 million non-reservers ($R_{ijt} = 0$) from distance bin $(350, 400]$, which includes non-reserving residents from Portland, Oregon and Redding, California. This implies a reservation rate of 1.8 per 100,000 individuals. All estimations use frequency weights for the number of individuals choosing either $R_{ijt} = 0$ or $R_{ijt} = 1$.

We use a zonal model for the reservation decision for several reasons. The primary purpose of the reservation estimation is to construct a control function that accounts for preferences in the cancellation estimation. As we will see, a person can only cancel a trip if they previously held a reservation. Therefore, preferences from the reservation decision likely play a role in the cancellation. The zonal reservation model accounts for these preferences while providing substantial flexibility over a multinomial logit approach. In this setting we have more than 1,200 arrival dates to define choice occasions, nearly 5 million reservations for the reserving individuals, 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 of the data in a multinomial logit model. One 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 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, a matter which we discuss further in Appendix C. 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 comptuationally less expensive than the contraction mapping method used in many multinomial logit studies (Berry 1994). This computational speed makes a difference when bootstrapping standard errors in the two stage model.

Denoting the set of parameters $\{\delta, \phi, \gamma, \psi_j, \lambda_t\}$ as ω, the likelihood and log likelihood function of a representative individual's reservation decision are written as:

$$\mathcal{L}(\omega|R_{ijt}) = \prod_{i=1}^{N}\prod_{j=0}^{J}\prod_{t=1}^{T} \mathbb{P}(R_{ijt} = 1|\omega)^{R_{ijt}} \left(1 - \mathbb{P}(R_{ijt} = 1|\omega)\right)^{1-R_{ijt}}; \tag{5}$$

$$\ell(\omega|R_{ijt}) = \sum_{i=1}^{N}\sum_{j=0}^{J}\sum_{t=1}^{T} R_{ijt} \log\left(\mathbb{P}(R_{ijt} = 1|\omega)\right) + (1 - R_{ijt}) \log\left(1 - \mathbb{P}(R_{ijt} = 1)\right), \tag{6}$$

for N agents, T choice occasions, and J sites. At the zonal level, we group each visitor i into travel cost zone $z \in \{1, 2, ..., Z\}$ to estimate the zonal travel cost model. Let N^0_{zjt} and N^1_{zjt} denote the number of non-reservers and reservers in each zone, respectively. For each zone the average travel costs for non-reservers and reservers are $\bar{c}^0_{zjt} = \frac{1}{N^0_{zjt}} \sum_{i \in z} (1 - R_{ijt}) c_{ijt}$ and $\bar{c}^1_{zjt} = \frac{1}{N^1_{zjt}} \sum_{i \in z} R_{ijt} c_{ijt}$. The likelihood function and log likelihood function are written as:

$$\mathcal{L}(\omega | R_{ijt}) = \prod_{z=1}^{Z} \prod_{j=0}^{J} \prod_{t=1}^{T} \mathbb{P}(R_{ijt} = 1 | \omega)^{N^1_{zjt}} \big(1 - \mathbb{P}(R_{ijt} = 1 | \omega)\big)^{N^0_{zjt}}; \tag{7}$$

$$\ell(\omega | R_{ijt}) = \sum_{z=1}^{Z} \sum_{j=0}^{J} \sum_{t=1}^{T} N^1_{zjt} \log \big(\mathbb{P}(R_{ijt} = 1 | \omega)\big) + N^0_{zjt} \log \big(1 - \mathbb{P}(R_{ijt} = 1 | \omega)\big). \tag{8}$$

Maximization of equation 8 yields utility parameters given a representative agent i from zone z.

3.2 Cancellations

For the second stage cancellation decision we model a binary choice at the level of the individual trip.[21] Assume that an agent chose to reserve at campground j. Close to the arrival date, within τ days, new preferences υ_{ijt} are realized. The agent chooses whether or not to cancel based on realized conditions. Let the utility from cancellation be:

$$U_{ijt} = V_{ijt} + \upsilon_{ijt}. \tag{9}$$

Because agents only face a cancellation decision if they previously made a reservation, we allow for correlation between their preferences close to the time of arrival υ_{ijt} and their preferences at the time of reservation ε_{ijt}. We assume a linear correlation structure:

$$\upsilon_{ijt} = \rho \varepsilon_{ijt} + \eta_{ijt}, \tag{10}$$

where η_{ijt} is distributed iid type I extreme value. The variable η_{ijt} reflects additional shocks to the agent's preferences close to the trip. For example, an unforeseen work obligation might raise the opportunity cost of the visit, or the agent could learn new information that increases their anticipation of the trip. A value of $\rho \neq 0$ implies that preferences at the time of reservation influence the cancellation decision, which we will see is an empirically testable hypothesis.

Let $C_{ijt} = 1$ if the agent cancels their reservation and 0 if they follow through. Within τ

[21] In Appendix C we show that very few users cancel their trip and rebook at another site for the same choice occasion. Close to the arrival date, many campgrounds are fully booked, which can prevent substitution. In addition, because smoke conditions are spatially and temporally correlated, substitution is unlikely an important factor in the identification of the smoke parameter. Therefore, a binary cancellation decision is a reasonable representation of the choice that visitors face.

days of arrival, site conditions such as smoke s_{jt} are approximately known to the individual, so they maximize utility over realized conditions by comparing U_{ijt} to U_{i0t}. The probability that an individual does not cancel, i.e. that they follow through with their reservation, is:

$$
\begin{aligned}
\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1) &= \mathbb{P}(U_{ijt} \geq U_{i0t}) \\
&= \mathbb{P}(V_{ijt} + \upsilon_{ijt} \geq V_{i0t} + \upsilon_{i0t}) \\
&= \mathbb{P}(V_{ijt} + \rho\varepsilon_{ijt} + \eta_{ijt} \geq 0 + \rho\varepsilon_{i0t} + \eta_{i0t}) \\
&= \mathbb{P}(\eta_{i0t} - \eta_{ijt} \leq V_{ijt} - \rho(\varepsilon_{i0t} - \varepsilon_{ijt})), \qquad (11)
\end{aligned}
$$

again substituting for $V_{i0t} = 0$.

Equation 11 presents challenges for the econometrician. The variables ε_{i0t} and ε_{ijt} are unobserved. However, omission of these variables will bias parameter estimates because they are correlated with travel cost among the selected sample. The selected sample is such that we only observe the cancellation decision for visitors that have already made a reservation. For those that did reserve, the correlation is $\mathbb{E}[c_{ijt}\varepsilon_{ijt}|R_{ijt} = 1] > 0$; among the selected sample of reservers, those with a high travel cost tend to have a high taste for the site. This relationship downward biases estimates of the travel cost parameter δ in the cancellation decision and thus inflates estimates of WTP $= \phi/\delta$. Appendix D explores this relationship using a numerical example. We show that the bias arises only when preferences are correlated ($\rho \neq 0$ in equation 10) 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 novel 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:

$$\begin{aligned} f(\ \tilde{\varepsilon}_{ijt} \mid R_{ijt} = 1) &= f(\ \tilde{\varepsilon}_{ijt} \mid \tilde{\varepsilon}_{ijt} \leq \mathbb{E}[V_{ijt}] - \mathbb{E}[V_{i0t}]) \\ &= \frac{f(\tilde{\varepsilon}_{ijt}) \cdot \mathbb{1}\{\tilde{\varepsilon}_{ijt} \leq \mathbb{E}[V_{ijt}] - \mathbb{E}[V_{i0t}]\}}{F(\mathbb{E}[V_{ijt}] - \mathbb{E}[V_{i0t}])} \\ &= \frac{f(\tilde{\varepsilon}_{ijt}) \cdot \mathbb{1}\{\tilde{\varepsilon}_{ijt} \leq \mathbb{E}[V_{ijt}]\}}{\mathbb{P}(R_{ijt} = 1)}, \end{aligned} \tag{12}$$

where the first line follows from the reservation condition in equation 4, the second line from the definition of a truncated density, and the third line by noting that $V_{i0t} = 0$ and that $F(\mathbb{E}[V_{ijt}]) = \mathbb{P}(R_{ijt} = 1)$. An estimand for $\tilde{\varepsilon}_{ijt}$ is given by:

$$\begin{aligned} \mathbb{E}[\ \tilde{\varepsilon}_{ijt} \mid \tilde{\varepsilon}_{ijt} \leq \mathbb{E}[V_{ijt}]] &= \int_{-\infty}^{\infty} \tilde{\varepsilon}_{ijt}\ f(\tilde{\varepsilon}_{ijt} \mid \tilde{\varepsilon}_{ijt} \leq \mathbb{E}[V_{ijt}])\ d\tilde{\varepsilon}_{ijt} \\ &= \frac{\int_{-\infty}^{\mathbb{E}[V_{ijt}]} \tilde{\varepsilon}_{ijt}\ f(\tilde{\varepsilon}_{ijt})\ d\tilde{\varepsilon}_{ijt}}{\mathbb{P}(R_{ijt} = 1)} \\ &= \frac{\mathbb{E}[V_{ijt}] \cdot \frac{\exp(\mathbb{E}[V_{ijt}])}{1+\exp(\mathbb{E}[V_{ijt}])} - \log(1 + \exp(\mathbb{E}[V_{ijt}]))}{\mathbb{P}(R_{ijt} = 1)} \\ &= \frac{\mathbb{E}[V_{ijt}] \cdot \mathbb{P}(R_{ijt} = 1) - I_{ijt}}{\mathbb{P}(R_{ijt} = 1)} \\ &= \mathbb{E}[V_{ijt}] - \frac{I_{ijt}}{\mathbb{P}(R_{ijt} = 1)}. \end{aligned} \tag{13}$$

The first line follows from the definition of a conditional expectation, the second line by substituting in equation 12, the third line by evaluating the definite integral, the fourth line by substituting equation 4 and by defining $I_{ijt} \equiv \log(1 + \exp(\mathbb{E}[V_{ijt}])$, and the final line through simplification. Equation 13 contains familiar terms. The $\mathbb{E}[V_{ijt}]$ term gives the expected utility of the site choice from the reservation decision. The second term contains the value I_{ijt}, which is equivalent to the inclusive value in the nested logit literature (Train 2009). This term approximates the expected maximal utility a visitor could expect from holding the reservation, which includes either the trip or the cancellation. The I_{ijt} term is scaled by the inverse of the probability that they would reserve at the site.

The estimand $\tilde{\varepsilon}_{ijt}$ captures the preferences of individual i from their reservation decision, allowing for unbiased estimation of the travel cost parameter in the cancellation problem. Since 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 11. In Appendix D we illustrate the bias correction of this estimand 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 8, for reservations made earlier than $t - \tau$ and using expected site conditions. Then, we use the parameters to create a fitted value $\hat{\tilde{\varepsilon}}_{ijt}$ 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 the trip. In this second stage the independent variables in V_{ijt} use realized rather than expected site conditions since users approximately know the site conditions close to the arrival date. For individual i, the likelihood and log likelihood function for the cancellation decision are:

$$\mathcal{L}(\omega | C_{ijt}, R_{ijt} = 1) = \prod_{i=1}^{N} \prod_{j=0}^{J} \prod_{t=1}^{T} \mathbb{P}(C_{ijt} = 0 | \omega, R_{ijt} = 1)^{1 - C_{ijt}} \left(1 - \mathbb{P}(C_{ijt} = 0 | \omega, R_{ijt} = 1)\right)^{C_{ijt}}; \tag{14}$$

$$\ell(\omega | C_{ijt}, R_{ijt} = 1) = \sum_{i=1}^{N} \sum_{j=0}^{J} \sum_{t=1}^{T} (1 - C_{ijt}) \log\left(\mathbb{P}(C_{ijt} = 0 | \omega, R_{ijt} = 1)\right) + C_{ijt} \log\left(1 - \mathbb{P}(C_{ijt} = 0 | \omega, R_{ijt} = 1)\right). \tag{15}$$

Because of the two stage estimation the researcher can use a bootstrapping process to obtain appropriate standard errors (Cameron and Miller 2015, Wooldridge 2015).

3.4 Numerical example

In Appendix D we provide a numerical example of the bias correction of our control function. We simulate 10,000 draws with $N = 100,000$ users who reserve and cancel. We assign each user random travel costs, smoke conditions, and preferences ε_{ijt} and $\upsilon_{ijt} = \rho\varepsilon_{ijt} + \eta_{ijt}$. Arbitrarily, we assert a true WTP to avoid smoke of $\phi/\delta = 2$. We vary two dimensions in the simulation. First, we test the role of dependent preferences in the two stages by turning ρ on ($\rho \neq 0$) and off ($\rho = 0$). Second, we test the role of sample selection. Unlike with the real recreation data, in the simulation we observe the counterfactual cancellation decisions of users who never held a reservation. In the simulation we test the cancellation estimation on both the selected sample of reservers and among the full sample, which includes non-reservers.

Table 1 summarizes results from the simulation. Columns 2 and 3 show that either correlated preferences or sample selection alone do not bias WTP estimates. It is only in column 4, when both conditions are present, that WTP is biased. In Appendix D 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 5, we maintain both sample selection and correlated preferences, but introduce our control function $\tilde{\varepsilon}_{ijt}$. 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 D.

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)	(5)
WTP	2.00**	2.00**	2.00**	3.77**	1.98**
	(0.08)	(0.10)	(0.11)	(0.46)	(0.17)
Users	All users	All users	Reservers	Reservers	Reservers
ρ		Yes		Yes	Yes
2-step estim.					$\tilde{\varepsilon}_{ijt}$

Notes: True WTP = 2. N = 100,000 users. * $p < 0.05$, ** $p < 0.01$.

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 ahead of time to cancellations close to arrival. Figure 3 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 3, and who subsequently decided whether to cancel within a week of the arrival date. We therefore exclude reservations which 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. Lastly, 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.[22]

Our analysis explores the sample selection issues characterized in the previous section, namely that a visitor can only cancel a trip if they previously demonstrated a taste for the site by reserving. Without accounting for this sample selection, the results would imply that wildfire smoke causes \$154 in welfare damages per person per trip. However, when accounting for sample selection using a control function, we find damages of \$107 per person per trip. We show that the control function operates by correcting for the correlation between a user's preferences and their travel cost in the selected sample.

We also discuss the shape of the damage function for smoke. Welfare damages monotonically increase in the duration of smoke events. The main results set the variable of interest s_{jt} equal to one if there was at least one hazardous smoke day in the week of arrival, $t - \tau$. However, when a campground was affected by smoke on two, three, or up to seven days in the week of arrival, we find damages of up to \$432 per person per trip. The damage function increases at an increasing rate, implying convexity in the duration of smoke events.

Damages are attenuated, however, when wildfire smoke is far from an active fire. When we remove observations for which there is an active wildfire within 20 km, we find reduced damage estimates of \$85 per person per trip. The estimates are robust to a placebo which reassigns smoke events to the weeks following an arrival date.

4.1 Cancellations close to arrival

Figure 4 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 than during non-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 D, 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 for someone with a similar travel cost that 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

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

in the estimation of cancellations $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$, which translates to a shallow slope in Figure 4.

Figure 4: Cancellation rate close to arrival.

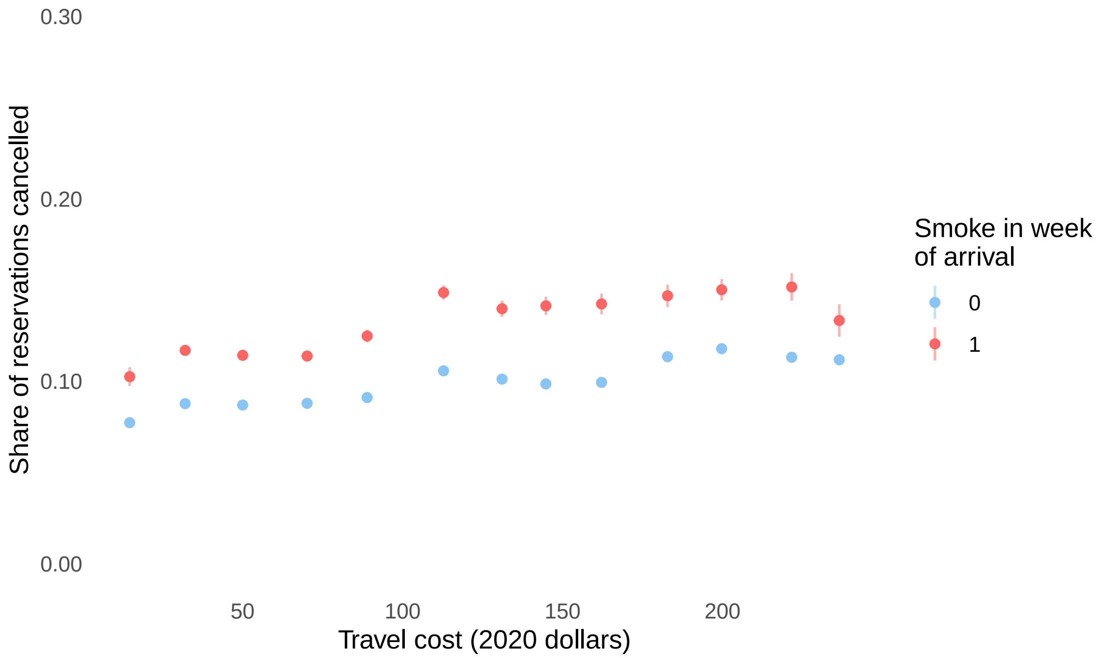


Table 2 reports results for biased estimation of cancellations $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ using the trip-level maximum likelihood function of equation 15. These estimates ignore the correlation between ε_{ijt} and travel cost among the set of users that chose to reserve. WTP is computed by taking the ratio of marginal disutility in smoke to marginal disutility in expenditure, i.e. the smoke coefficient divided by the travel cost coefficient. Standard errors for WTP are computed using the delta method.[23] In all estimations the observations are weighted using frequency weights since a single reservation might represent, for example, two visitors 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 fixed effect to account for location-specific, time-invariant unobservables related to site quality. We also account for differences in reservation rates based on the day of the week, since weekends see higher reservation activity than weekdays. A campground by week fixed effect controls for unobserved location-specific seasonality, such as seasonal campground-specific natural phenomena. Lastly, we include various year fixed

[23]For an example of the delta method for the ratio of two coefficients, such as the ratio in our WTP estimate, an interested reader may refer to Casella and Berger (2002), example 5.5.27.

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 since WTP $= \phi/\delta$ and we expect the travel cost parameter δ to be attenuated.

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 FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

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

To correct for the biased WTP in Table 2 we use the control function described in equation 13, $\tilde{\varepsilon}_{ijt} = \mathbb{E}[V_{ijt}] - \frac{I_{ijt}}{\mathbb{P}(R_{ijt}=1)}$. The first step is to estimate the probability of reservation earlier than one week based on expected site conditions and using a 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 our PRISM data source, which represent average weather conditions from the period 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 8. 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 zero. 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 expectations parameter.

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 FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

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

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. Figures A7 and A8 in Appendix A show the variation in fitted probability $\mathbb{P}(R_{ijt} = 1)$ and in the control function $\tilde{\varepsilon}_{ijt}$. Since we expect that preferences and travel costs are correlated in the selected sample, $\mathbb{E}[c_{ijt}\varepsilon_{ijt}|R_{ijt} = 1] > 0$, then it should be true that our control function is inversely correlated with travel cost, $\mathbb{E}[c_{ijt}(\varepsilon_{i0t} - \varepsilon_{ijt})|R_{ijt} = 1] < 0$. Figure 5 illustrates this relationship using the fitted values of $\tilde{\varepsilon}_{ijt}$ for the sample of reservers. The slope of the fitted line follows the expected direction. Using 400 bootstrapped estimations and the fixed effects from model (4) we find that $c_{ijt} = -(239.738 + 24.70\ \tilde{\varepsilon}_{ijt})$, where the intercept and slope coefficients are both significant at the 0.01 level. This empirical result is consistent with the prediction of our theory and numerical exercise in Section 3 and Appendix D.

Figure 5: Relationship between control function $\tilde{\varepsilon}_{ijt}$ and travel cost using model (4) demonstrates correlation between preferences and travel cost in the selected sample of reservers.

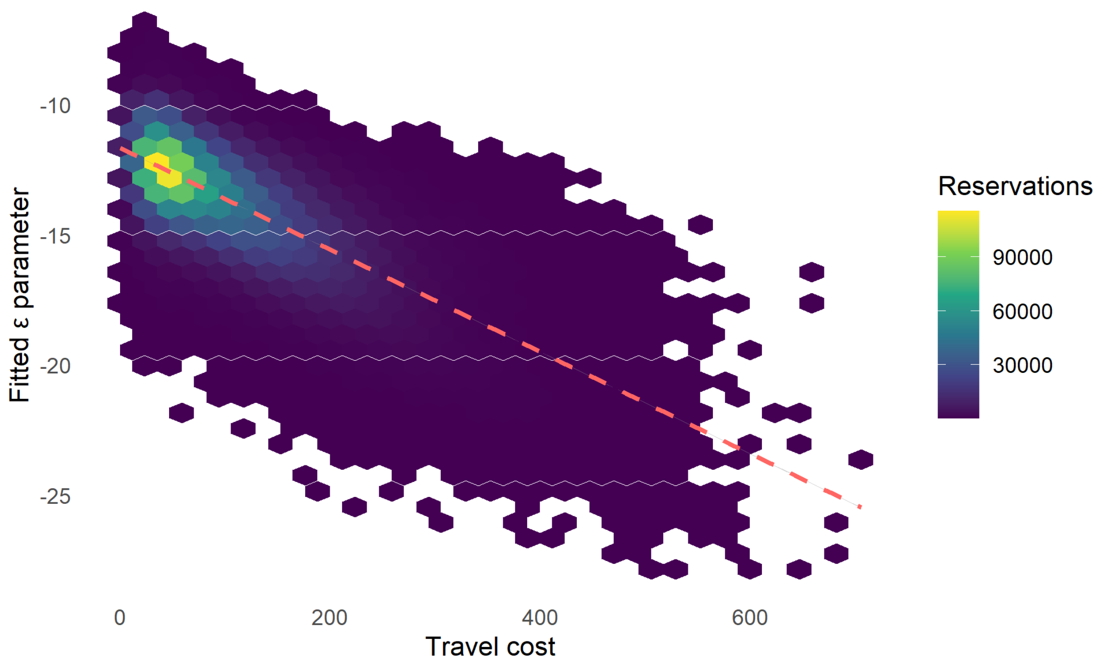


Table 4 reports the main trip-level results for the cancellation estimation $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ using the bias correcting control function $\tilde{\varepsilon}_{ijt}$. The coefficient for $\tilde{\varepsilon}_{ijt}$ is significant, suggesting that preferences at the time of reservation 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 the notion that sample selection bias operates through correlation with travel cost. Overall, the WTP estimates are reduced to \$107 per person per trip of lost utility due to cancellations. By comparison, the biased results in Table 2 were \$154 per person per trip, which is 44% higher.

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 FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year 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. * $p < 0.05$, ** $p < 0.01$.

Wooldridge (2015) recommends bootstrapping standard errors for control functions because of the two stage estimation process. We follow the clustered bootstrapping process of Cameron and Miller (2015), drawing with replacement at the campground level for 400 bootstraps. In Appendix E 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.

One concern when studying cancellations is the question of whether an individual will formally cancel their reservation, or whether they will simply not show up. For most campgrounds we do not observe whether an individual checks in to their campground or not. 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 of the reservation less a \$10 cancellation fee; and, when cancelling within 24 hours of arrival, they are still reimbursed for the full trip less the \$10 fee and the price of the first night's stay. Still, we further explore this question in Appendix F. For a small subset of campgrounds we are able to observe no shows. Among the sample of campgrounds reporting no shows we demonstrate that the inclusion or exclusion of no show observations in the estimation of cancellations $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ does not change the estimates for the smoke or travel cost coefficients. For a discussion of this issue the interested reader may refer to Appendix F.

4.2 Shape of the smoke damage function

In this section we investigate the shape of the damage function for smoke. Welfare damages may be higher for more severe smoke events, where severity could refer to the intensity or duration of a smoke event. We explore the relationship of damages to the duration of an event. In our main specification the variable of interest is an indicator equal to one if the campground was affected by at least one day of wildfire smoke in the week of arrival. We respecify the equation of interest, $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$, and allow for differential effects based on the number of smoke-affected days in the week of arrival.

Figure 6 plots the welfare damages visually as a function of the number of smoke days in the week of arrival. For full results an interested reader may refer to Table A2 in Appendix A. Damages monotonically increase in the number of smoke days. 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. Further, the damage function appears to increase at an increasing rate. That is, welfare damages are approximately convex in the number of smoke days in the arrival week. Figure 7 demonstrates this relationship by showing that the marginal willingness to pay (MWTP) generally increases in the number of smoke-affected days in the week of arrival. On average, WTP rises about \$62 with each additional day of smoke. This result confirms intuition, as we should expect greater welfare losses for more severe wildfire smoke events.

Figure 6: The smoke damage function increases in the number of smoke days in week of arrival.

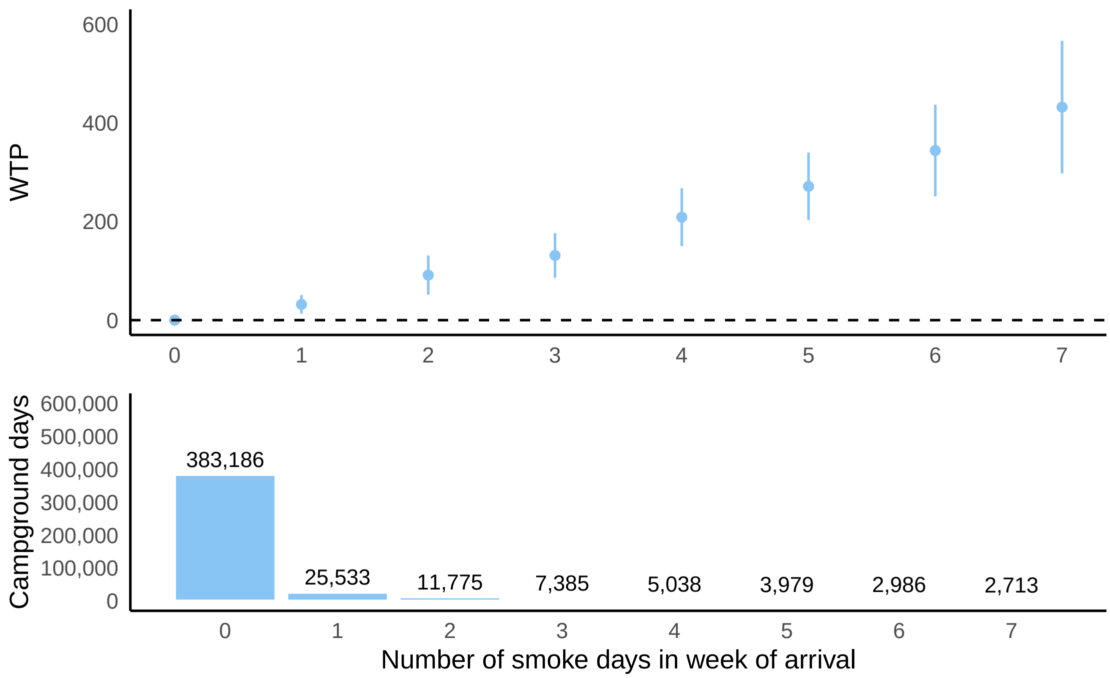
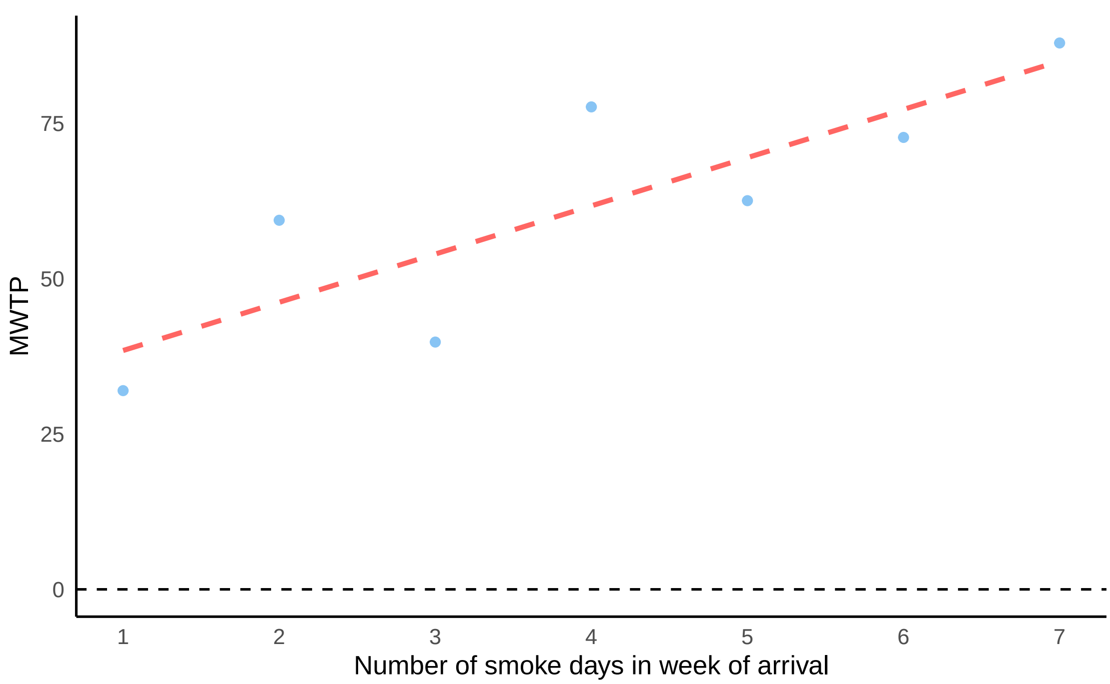


Figure 7: Marginal willingness to pay (MWTP) increases with additional smoke days, implying convexity of the damage function.



4.3 The role of active wildfires

The existing literature has found that visitors to National Parks are less avoidant of wildfire smoke that originates from distant sources (Cai 2021). In this section we investigate how nearby active wildfires affect the main estimates. Although the main estimation controls for proximity to active wildfire, one might still be concerned that individuals avoid recreation primarily due to fire rather than smoke. If smoke days are highly correlated with nearness to fire it could increase the estimated smoke coefficient and inflate WTP. To address this possibility, we reestimate the main specifications but remove observations for which there was a nearby active fire.

We consider a campground as near to fire on day t if there is an active burning wildfire within 20 km (12 miles), a threshold we have used in previous work (Gellman et al. 2022). Table 5 reports the number of reservations affected by either smoke or fire conditions. When there is fire nearby, there is nearly an equal number of smoke- and non-smoke-affected reservations. However, due to the large distances that smoke travels, most smoke-affected reservations are not for campgrounds near an actively burning wildfire.

Table 5: Reservations with smoke or fire conditions in the estimating dataset.

Smoke in week of arrival	Fire within 20 km	Number of reservations	Percent of sample (%)
0	0	2,356,407	86.5
1	0	322,114	11.8
0	1	24,199	0.9
1	1	21,220	0.8

Table 6 reports results for the cancellation estimation $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$ when removing observations with nearby active wildfire. We find welfare damages of \$85 per person per trip due to smoke. By comparison, in the main specification we estimated lost welfare of \$107 per person per trip. The omission of fire days thus reduced estimated welfare damages by approximately 20%. These results are consistent with the findings of Cai (2021), who found that outdoor recreationists are less responsive to smoke originating from distant sources of fire.

Table 6: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, removing days with wildfire within 20 km.

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.1678**	-0.2323**	-0.2119**	-0.2116**
	(0.0224)	(0.0195)	(0.0173)	(0.0177)
Travel cost (dollars)	-0.0028**	-0.0024**	-0.0025**	-0.0025**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Inv. distance to wildfire (km^{-1})	-9.4566**	-7.7163**	-7.2826**	-6.5691**
	(1.1591)	(1.0196)	(1.0794)	(0.8965)
High temp. (degrees C)	0.0195**	0.0275**	0.0277**	0.0298**
	(0.0043)	(0.0021)	(0.0021)	(0.0021)
Low temp. (degrees C)	-0.0029	-0.0202**	-0.0208**	-0.0249**
	(0.0056)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0043**	-0.0060**	-0.0062**	-0.0059**
	(0.0011)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0129	-0.0348**	-0.0357**	-0.0377**
	(0.0261)	(0.0121)	(0.0121)	(0.0123)
N	2,677,628	2,645,592	2,645,592	2,642,695
WTP	60.97**	95.08**	85.96**	84.66**
	(12.4)	(13.13)	(12.03)	(12)
Campground FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

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

4.4 Placebo test for smoke

As a robustness check, we devise a placebo test to check whether the smoke coefficient actually measures responses to wildfire 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, there are more than 375,000 placebo reservations for campgrounds that saw smoke in the week or second week after arrival.

Table 7 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.

Table 7: 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 FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

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

5 Total welfare losses

In the preceding sections we have 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 in the west. One limitation of this analysis is that the welfare 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 due to smoke. This back of the envelope estimate represents the lost welfare to inframarginal visitors and does not include the value of lost trips.

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 the National Association of State Park Directors (Smith et al. 2019) for the years 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, an interested reader may refer to Appendix G.

Table 8 displays estimates of total visitation, smoke-affected visitation, and total welfare losses. 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 affected by wildfire smoke, or 4.2%. When multiplied by the per trip estimate of $107, we find total annual welfare losses of approximately $2.3 billion due to smoke. 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 due to the fact that much of the agency's visitation (nearly 40%) occurs at lakes and reservoirs in the Pacific Northwest, a region which has seen particularly high wildfire smoke impacts relative to other regions (Burke et al. 2021a, Gellman et al. 2022, Miller et al. 2021). As a whole, these findings show the high cost of wildfire smoke for outdoor recreation in the western United States.

[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.

Table 8: Smoke-affected recreation visits and welfare losses for the western US, 2008 to 2017.

	Total visits/year (millions)	Smoke-affected visits/year (millions)	Welfare loss/year (millions)
National Park Service	102.6	2.3	$248.1
US Forest Service	108.0	4.8	$511.4
Bureau of Land Management	59.8	2.5	$267.3
US Army Corps of Engineers	46.4	2.4	$251.7
State Parks	194.5	9.6	$1,022.1
Total	511.4	21.5	$2,300.6

Notes: Welfare losses computed by multiplying per trip damages of $107 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, and are attenuated when smoke-affected campgrounds are far from active wildfires. 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.

The paper provides several contributions to the literature. First, we contibute by using novel methods and data. We value a temporary environmental bad, wildfire smoke, in a context where visitors face changing sets of information. To do so, we develop a two stage decision structure that links preferences with a control function. This model draws on work from economists concerned with sample selection in non-linear models (Greene 2012, Terza 2009), as well as 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. Further, our use of administrative data complements recent literature using large or innovative datasets to study recreation across multi-state regions (Cameron and Kolstoe 2022, Dundas and von Haefen 2020, English et al. 2018).

We also add to the existing 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, healthcare costs, or have valued changes in mortality using the value of a statistical life (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, namely expenditures on air purifiers, as well as health outcomes and risk perceptions. They derive a WTP to avert one wildfire induced symptom day of $84 in 2009 dollars, or $102 when adjusted to 2020 dollars. We estimate the 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 length of 2.84 days.

We can also compare 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 billion and $170 billion in annual damages, in 2021 dollars. These results mainly vary due to assumptions on remaining years of life, since 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, Cullen 2020, Wen and Burke 2021).

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 on suppression.[28,29] Wildfires destroy thousands of structures per year, which has cost tens of billions of dollars in recent years (Baylis and Boomhower 2021, Buechi et al. 2021). Both states and the federal government have pledged to increase fuels 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] Consistent with this goal, California has proposed to spend $1.2 billion across Fiscal Years 2022-23 and 2023-24 for fire mitigation activities

[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.

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.

7 References

Abatzoglou, J. T., & Williams, A. P. (2016). Impact of anthropogenic climate change on wildfire across western US forests. Proceedings of the National Academy of Sciences, 113(42), 11770-11775.

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

Baylis, P., & Boomhower, J. (2021). The Economic Incidence of Wildfire Suppression in the United States. American Economic Journal: Applied Economics, forthcoming.

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

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. RAND Journal of Economics 25(2):, 242-262.

Borgschulte, M., Molitor, D., & Zou, E. (2022). Air pollution and the labor market: Evidence from wildfire smoke (No. w29952). National Bureau of Economic Research.

Buechi, H., Weber, P., Heard, S., Cameron, D., & Plantinga, A. J. (2021). Long-term trends in wildfire damages in California. International Journal of Wildland Fire, 30(10), 757-762.

Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., & Wara, M. (2021a). The changing risk and burden of wildfire in the United States. Proceedings of the National Academy of Sciences, 118(2).

Burke, M., Heft-Neal, S., Li, J., Driscoll, A., Baylis, P. W., Stigler, M., Weill, J., & Gould, C. (2021b). Exposures and behavioral responses to wildfire smoke (No. w29380). National Bureau of Economic Research.

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

Cai, C. (2021). Wildfire and visitation in US National Parks. Working paper, Department of Agricultural and Consumer Economics, University of Illinois, Urbana-Champaign.

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

Cameron, T. A., & DeShazo, J. R. (2013). Demand for health risk reductions. Journal of Environmental Economics and Management, 65(1), 87-109.

Cameron, T. A., & Kolstoe, S. H. (2022). Using auxiliary population samples for sample-selection correction in models based on crowd-sourced volunteered geographic information. Land Economics, 98(1), 1-21.

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. Journal of Human Resources, 50(2), 317-372.

Casella, G., & Berger, R. L. (2002). Statistical Inference (Second Edition). Duxbury.

Cullen, D. (2020). The health cost of wildfire smoke. Working paper, Department of Economics, University of California, Santa Barbara.

Dundas, S. J., & von Haefen, R. H. (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), 209-242.

English, E., von Haefen, R. H., Herriges, J., Leggett, C., Lupi, F., McConnell, K., Welsh, M., Domanski, A., & Meade, N. (2018). Estimating the value of lost recreation days from the Deepwater Horizon oil spill. Journal of Environmental Economics and Management, 91, 26-45.

Fowler, M., Modaresi Rad, A., Utych, S., Adams, A., Alamian, S., Pierce, J., Dennison, P., Abatzoglou, J.T., AghaKouchak, A., Montrose, L., & Sadegh, M. (2019). A dataset on human perception of and response to wildfire smoke. Scientific Data, 6(1), 1-10.

Gellman, J., Walls, M., & Wibbenmeyer, M. (2022). Wildfire, smoke, and outdoor recreation in the western United States. Forest Policy and Economics, 134, 102619.

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

Graff Zivin, J., Liu, T., Song, Y., Tang, Q., & Zhang, P. (2020). The unintended impacts of agricultural fires: Human capital in China. Journal of Development Economics, 147, 102560.

Greene, W. H. (2012). Econometric analysis (Seventh). Pearson.

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

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

Kolstoe, S., & Cameron, T. A. (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, 93–99.

Lewis, D. J., Dundas, S. J., Kling, D. M., Lew, D. K., & Hacker, S. D. (2019). The non-market benefits of early and partial gains in managing threatened salmon. PloS one, 14(8), e0220260.

Liang, Y., Sengupta, D., Campmier, M. J., Lunderberg, D. M., Apte, J. S., & Goldstein, A. H. (2021). Wildfire smoke impacts on indoor air quality assessed using crowdsourced data in California. Proceedings of the National Academy of Sciences, 118(36).

McCoy, S. J., & Zhao, X. (2020). Wildfire and infant health: a geospatial approach to estimating the health impacts of wildfire smoke exposure. Applied Economics Letters, 28(1), 32-37.

Miller, N., Molitor, D., & Zou, E. (2021). A causal concentration-response function for air pollution: Evidence from wildfire smoke. Working paper, College of Business, University of Illinois, Urbana-Champaign.

Phaneuf, D. J., & Requate, T. (2016). A course in environmental economics: theory, policy, and practice. Cambridge University Press.

Plantinga, A. J., Walsh, R., & Wibbenmeyer, M. (2022). Priorities and effectiveness in wildfire management: evidence from fire spread in the western United States. Journal of the Association of Environmental and Resource Economists, 9(4), 603-639.

Preisler, H. K., Schweizer, D., Cisneros, R., Procter, T., Ruminski, M., & Tarnay, L. (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, 340-349.

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, 1334–1343.

Richardson, L. A., Champ, P. A., & Loomis, J. B. (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.

Schroeder, W., Ruminski, M., Csiszar, I., Giglio, L., Prins, E., Schmidt, C., & Morisette, J. (2008). Validation analyses of an operational fire monitoring product: the hazard mapping system. International Journal of Remote Sensing 29(20), 6059-6066.

Smith, J. W., Wilkins, E. J., & Leung, Y. F. (2019). Attendance trends threaten future operations of America's state park systems. Proceedings of the National Academy of Sciences, 116(26), 12775-12780.

Train, K. E. (2009). Discrete choice methods with simulation. Cambridge University Press.

Terza, J. V. (2009). Parametric nonlinear regression with endogenous switching. Econometric Reviews, 28(6), 555-580.

Walls, M., Wichman, C., & Ankney, K. (2018). Nature-based recreation: Understanding campsite reservations in National Parks. Washington, DC: Resources for the Future.

Wen, J., & Burke, M. (2021). Wildfire smoke exposure worsens learning outcomes. Working paper, Department of Earth System Science, Stanford University.

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.

Williams, A. P., Abatzoglou, J. T., Gershunov, A., Guzman-Morales, J., Bishop, D. A., Balch, J. K., & Lettenmaier, D. P. (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), 420-445.

Appendix A: Additional figures

Figure A1: Recreation.gov web interface.

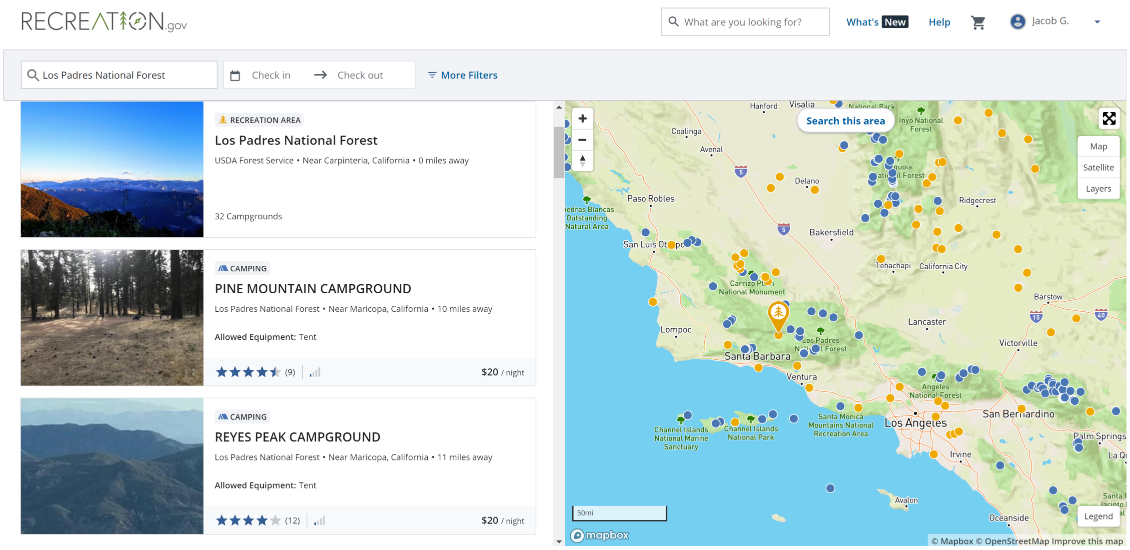


Figure A2: Automobile route from Santa Barbara, California to Yosemite National Park.

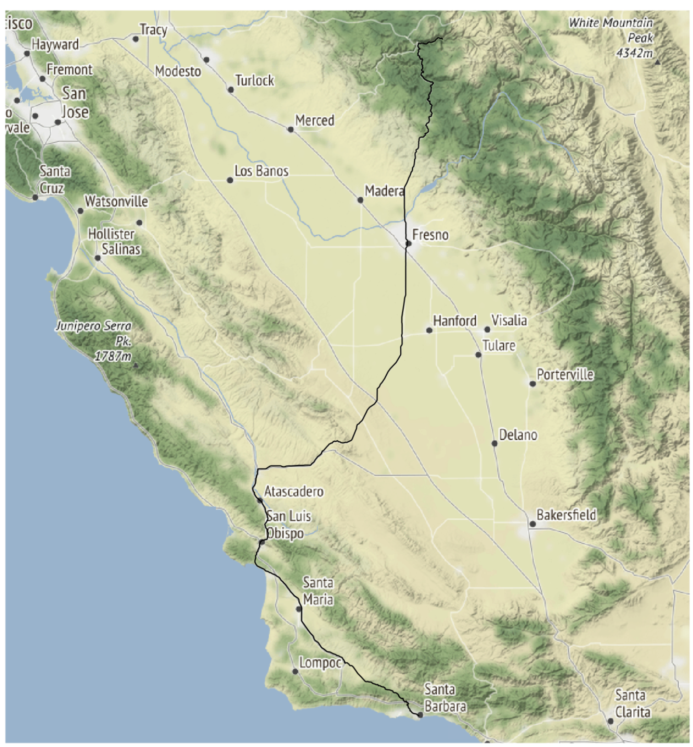


Figure A3: NOAA smoke plumes and $PM_{2.5}$. Red areas are affected by smoke and poor air quality.

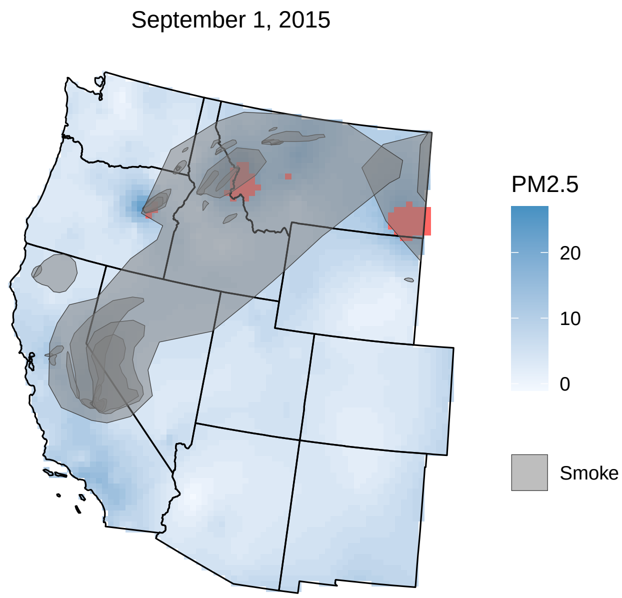


Figure A4: Fire detection points and fire perimeters.

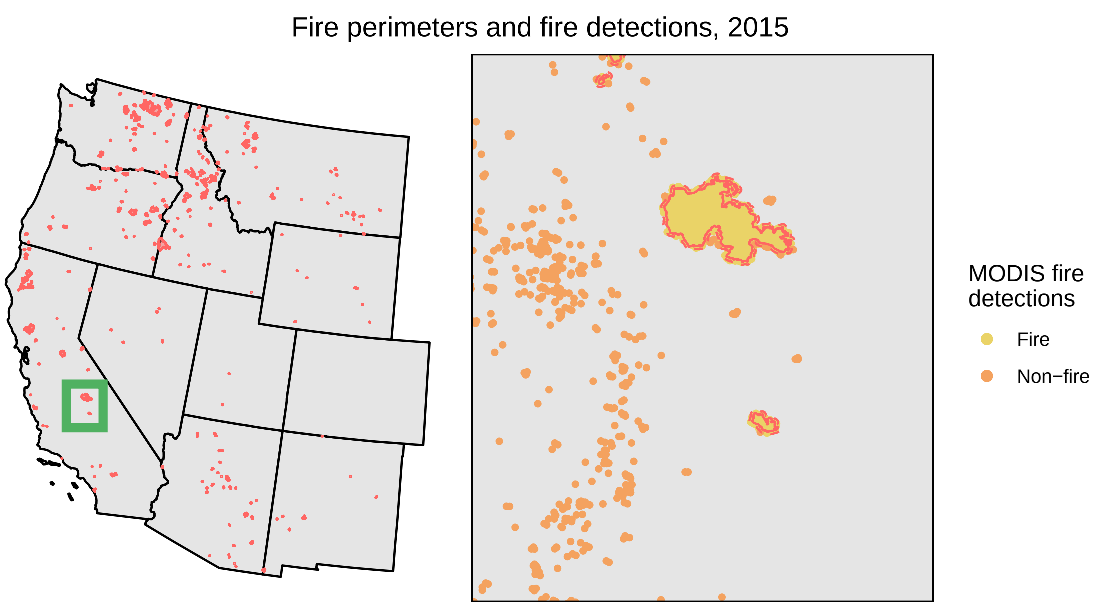


Figure A5: Map of campgrounds in dataset.

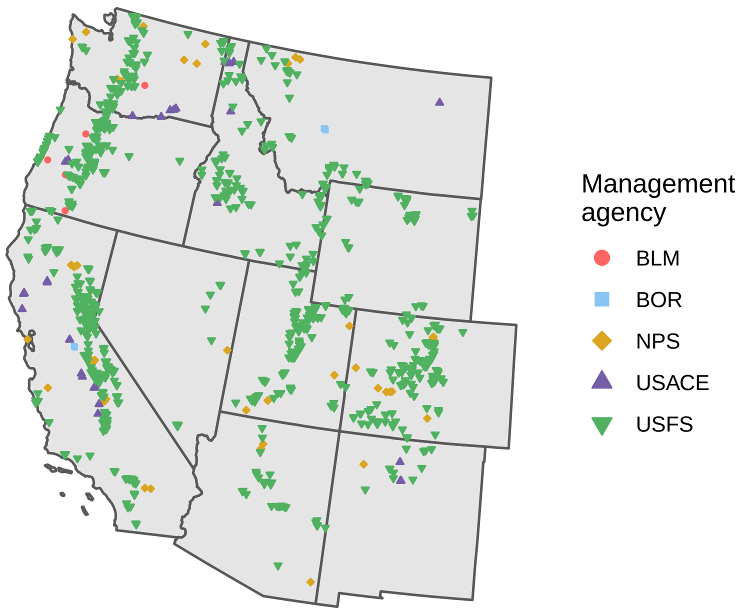


Table A1: Most visited federally-managed campgrounds.

Campground	Recreation area	State	Agency	Annual average campers
Upper Pines	Yosemite NP	CA	NPS	99,820
Mather	Grand Canyon NP	AZ	NPS	59,196
Watchman	Zion NP	UT	NPS	49,389
Serrano	Big Bear, San Bernardino NF	CA	USFS	46,610
Pinecrest	Summit RD, Stanislaus NF	CA	USFS	36,576
Fallen Leaf	Lake Tahoe Basin	CA	USFS	32,966
Lodgepole	Sequoia And Kings Canyon NP	CA	NPS	30,634
North Pines	Yosemite NP	CA	NPS	26,883
Moraine Park	Rocky Mountain NP	CO	NPS	25,884
Lower Pines	Yosemite NP	CA	NPS	25,644
Wawona	Yosemite NP	CA	NPS	25,407
Hodgdon Meadow	Yosemite NP	CA	NPS	24,746
Pinnacles	Pinnacles NP	CA	NPS	24,210
Crane Flat	Yosemite NP	CA	NPS	23,844
Indian Cove	Joshua Tree NP	CA	NPS	23,376
Dogwood	Arrow Head, San Bernardino NF	CA	USFS	21,540
Acorn	New Hogan Lake	CA	USACE	21,164
Black Rock	Joshua Tree NP	CA	NPS	19,888
Kalaloch	Olympic NP	WA	NPS	18,105
Dinkey Creek	High Sierra RD, Sierra NF	CA	USFS	16,294
Logger	Truckee RD, Tahoe NF	CA	USFS	16,253
Diamond Lake	Diamond Lake RD, Umpqua NF	OR	USFS	15,683
Kyen	Lake Mendocino	CA	USACE	15,015
Dorst Creek	Sequoia And Kings Canyon NP	CA	NPS	14,435
North Rim	Grand Canyon NP	AZ	NPS	13,898
Ohanapecosh	Mount Rainier NP	WA	NPS	13,889
Devils Garden	Arches NP	UT	NPS	13,138
Oh Ridge	Mono Lake RD, Inyo NF	CA	USFS	13,063
Fish Creek	Glacier NP	MT	NPS	12,434
Manzanita Lake	Lassen Volcanic NP	CA	NPS	12,379

Figure A6: Campground occupancy rates follow a bimodal distribution both on the date of arrival and one week in advance.

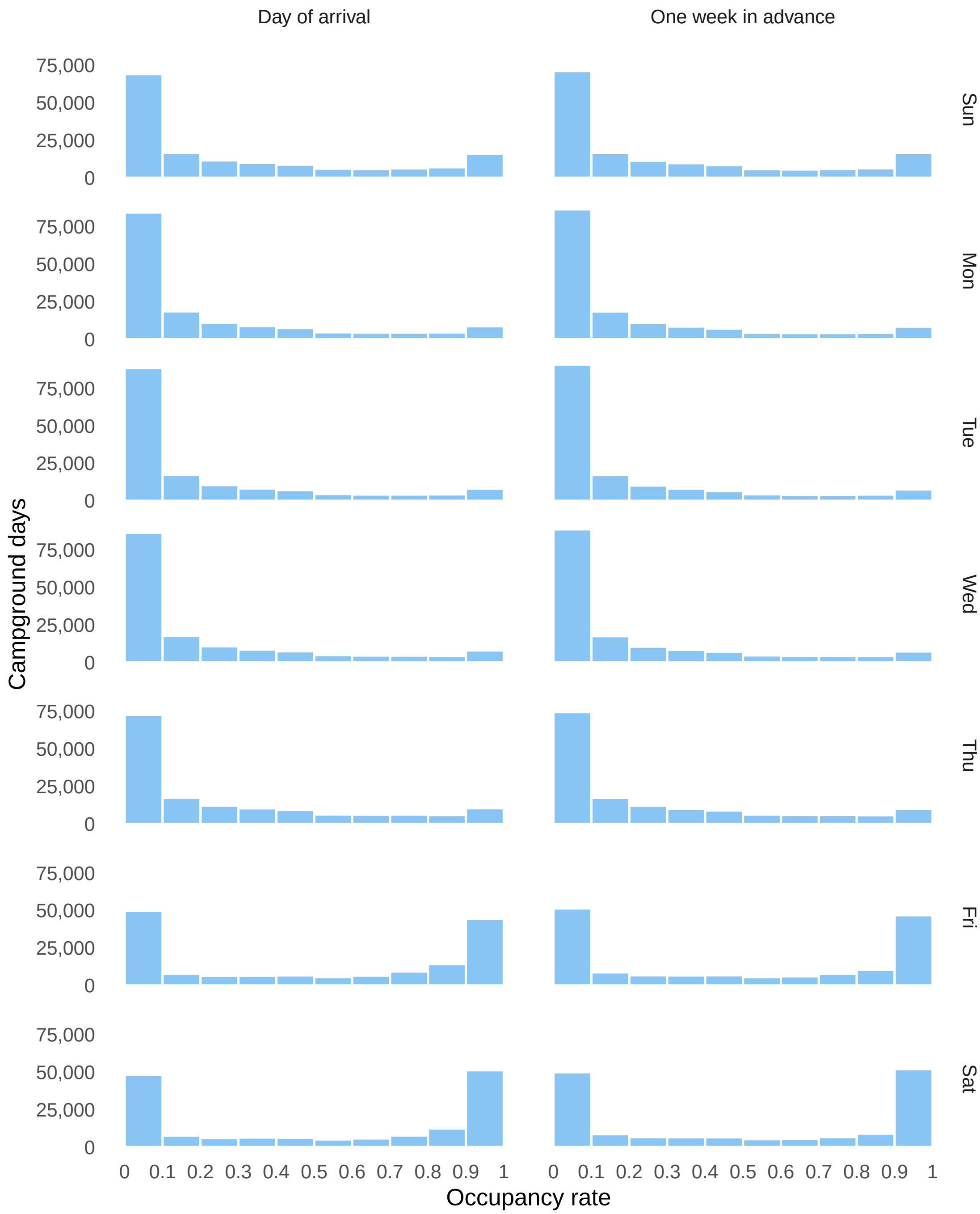


Figure A7: Fitted $\mathbb{P}(R_{ijt} = 1)$ for reservations made earlier than one week from model (4).

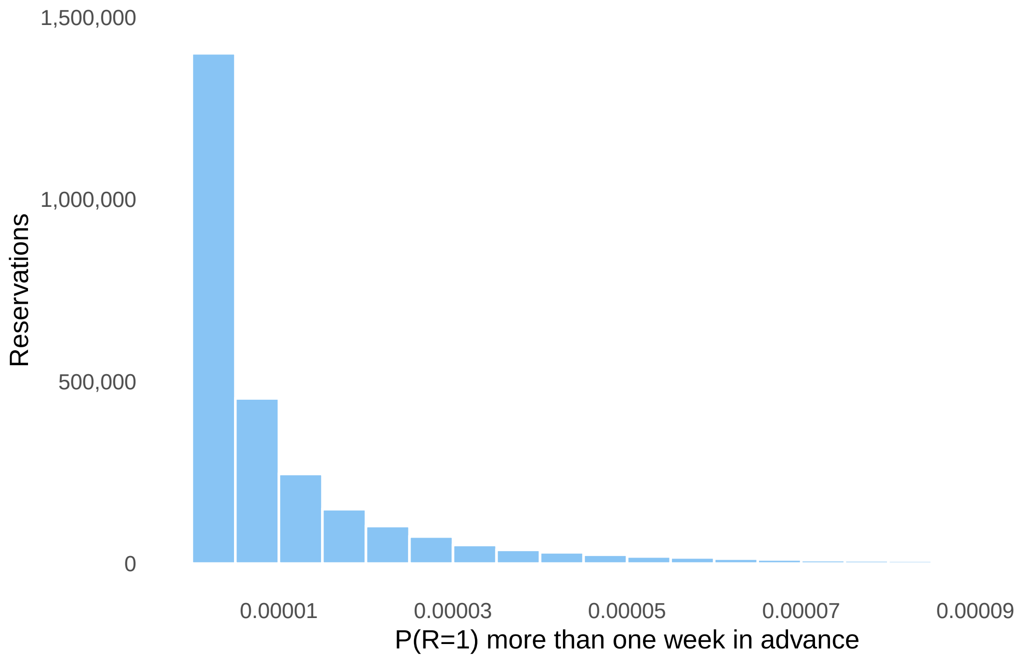


Figure A8: Fitted $\tilde{\varepsilon}_{ijt}$ for reservations made earlier than one week from model (4).

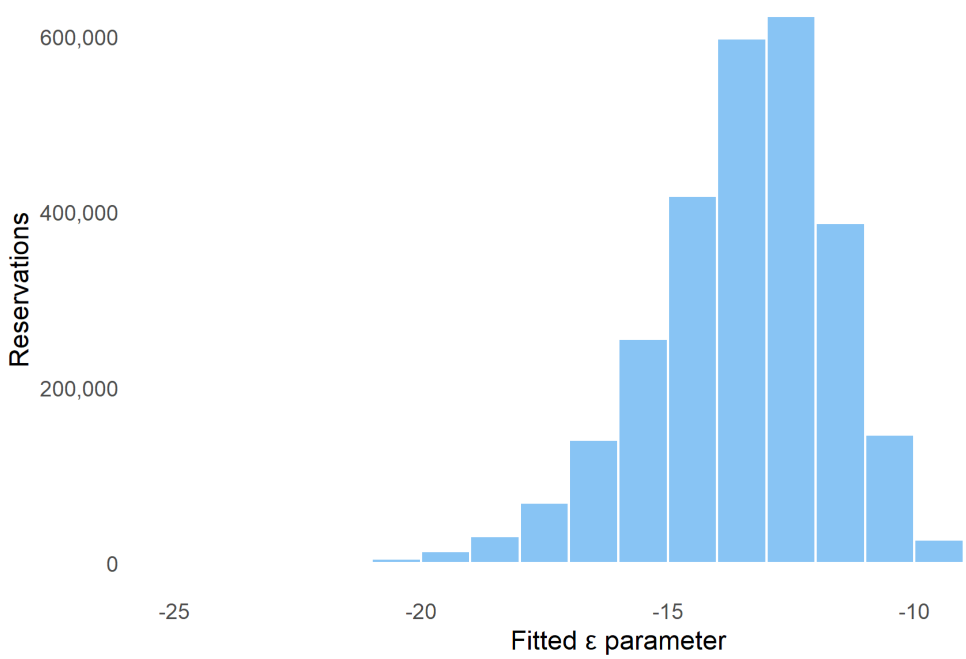


Table A2: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, heterogeneity by smoke days in week before arrival.

	(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 FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year 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: Reservations close to arrival

This paper focuses on the cancellation decisions of visitors who reserve ahead of time, before smoke conditions are known, and who subsequently decide whether to cancel close to the arrival date, after site conditions are realized. Figure 3 illustrates the timing of decisions in the main analysis. We focus on this structure for several reasons. First, most reservations are made ahead of time. Figure 2 shows that, although a plurality of reservations are made within a week of arrival, the majority are made in advance. Second, by the time smoke conditions are known, many campgrounds are either fully booked or completely empty, which limits the variation needed to identify changes in campground activity due to wildfire smoke. Figure A6 depicts this bimodal distribution. Congested campgrounds and empty campgrounds both prevent proper measurement of changes in recreation activity due to smoke. When campgrounds are completely booked, logistic regression would underestimate the latent demand for recreation on non-smoke days because campground occupancy meets a binding constraint; this analysis would lead to an underestimate of the coefficient on smoke. When campgrounds are empty on non-smoke days, there is similarly not identifying variation. We focus on cancellations because, once a visitor holds a reservation, they may always cancel it and do not face constraints.[32]

Still, we could have measured decisions for visitors who make new reservations close to the arrival date, when they are likely aware of smoke conditions. In this section we report results for a zonal travel cost model of new reservations close to the arrival date. We restrict the data to reservations made within a week of arrival during the months of May to September and over the years 2010 to 2017. We also limit attention to trips coming from within 650 km (400 miles), as described in Section 2.6. Lastly, we exclude new reservations which were also cancelled in the same week. These restrictions result in 693,501 same-week reservations. We aggregate these reservations for a zonal estimation as described in Section 3.1, but for same-week reservations rather than early reservations.

Figure B1 shows how reservation rates vary by travel cost and wildfire smoke conditions. Reservation rates are much higher at lower levels of travel cost. Before controlling for other observable and unobservable factors, Figure B1 shows that raw reservation rates are actually higher on days with smoke than on days without smoke. This difference is likely due to the fact that wildfire

[32]For a discussion of site substitution, refer to Appendix C. Users tend not to cancel and rebook for the same choice occasion. In addition, smoke conditions are spatially and temporally correlated among choice sets, meaning there is low variation of differences in smoke-related disutility among choice alternatives.

season overlaps with popular camping times, such as the American holidays of Independence Day and Labor Day. Therefore, fixed effects for location and seasonality are likely to be important.

Figure B1: Reservation rate within one week of arrival.

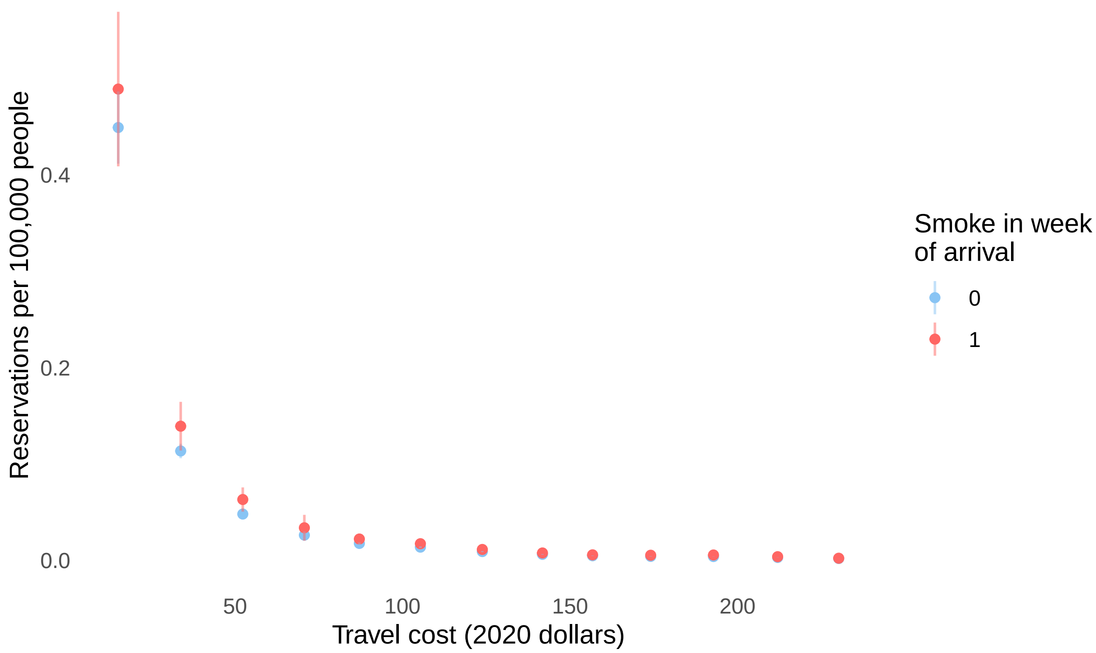


Table B1 reports results for estimation of the reservation likelihood within one week, $\mathbb{P}(R_{ijt} = 1)$, using the zonal maximum likelihood function of equation 8. In all estimations the observations are weighted using frequency weights since a single row of data might represent, for example, 20 reservers or 2.3 million non-reservers. In column 1 we display results without controlling for campground or seasonal fixed effects. As suggested by Figure B1, users are unconditionally more likely to reserve for dates that happened to be smoke-affected, yielding an unexpectedly positive coefficient on smoke. Columns 2 through 4 add fixed effects, which yield the expected sign for the smoke coefficient. The results in columns 2 through 4 indicate a WTP to avoid smoke of between \$1.45 and \$1.65 per person per trip. Given previous discussion of congested and empty campgrounds, we believe these estimates are less plausible than the paper's main set of results.

Table B1: $\mathbb{P}(R_{ijt} = 1)$ for reservations within one week.

	(1)	(2)	(3)	(4)
Smoke in week of arrival	0.1382**	-0.0434**	-0.0406**	-0.0460**
	(0.0021)	(0.0125)	(0.0104)	(0.0091)
Travel cost (dollars)	-0.0241**	-0.0279**	-0.0279**	-0.0279**
	(0.0000)	(0.0017)	(0.0017)	(0.0017)
Inv. distance to wildfire (km^{-1})	-0.6732**	-2.0485**	-2.0798**	-1.9822**
	(0.0310)	(0.3046)	(0.3029)	(0.2809)
High temp. (degrees C)	0.0602**	0.0074**	0.0075**	0.0079**
	(0.0002)	(0.0012)	(0.0011)	(0.0010)
Low temp. (degrees C)	-0.0205**	-0.0044**	-0.0039**	-0.0047**
	(0.0002)	(0.0016)	(0.0015)	(0.0014)
Precip. in week of arrival (mm)	-0.0035**	-0.0028**	-0.0027**	-0.0028**
	(0.0001)	(0.0004)	(0.0004)	(0.0003)
N	13,792,677	10,913,738	10,913,738	10,542,160
WTP	-5.73**	1.55**	1.45**	1.65**
	(0.09)	(0.45)	(0.37)	(0.32)
Campground FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year 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 C: Site substitution

In the main analysis, both the trip-level cancellation decision and the zonal travel cost model imply a binary choice for the representative visitor. While one could have used a multinomial logit to model site substitution, we discuss in Section 3.1 the practical limitations of that approach and the advantages afforded by a zonal travel cost model. Moreover, the zonal reservation estimation and the binary cancellation decision should properly identify the parameters of interest, namely the marginal disutility of smoke and the marginal disutility of expenditure. In this section we first show that a binary cancellation decision is a realistic representation of the choice that users face. We also discuss the choice to model early reservations in a zonal setting.

We begin by discussing the binary cancellation decision. Substitution following a cancellation is uncommon. Of the 2,723,940 trips in the estimating dataset, there are 268,750 cancellations, implying a 9.87% raw cancellation rate. Among the cancelled reservations, approximately 10.3% of users "rebooked," meaning they made a new reservation for a date within a year of their original scheduled arrival date. However, rebookers rarely substitute for the same choice occasion. Only 11% of rebookings substituted to a different campground on the same week of arrival; multiplying by 10.3%, this implies that only 1.1% of all cancellations substituted to a different site for the same week. Intertemporal substitution is more common: 57% of all rebookings were for either the same campground or a different campground but at a later arrival week. Multiplying by 10.3%, this means that 5.8% of all cancellations intertemporally substituted.

Because this analysis is concerned with wildfire smoke, we note the smoke status of rebooked visits. Of all rebooked visits, 0.9% were smoke-affected and rebooked for a different week; 1% were smoke-affected and rebooked for the same week. Multiplying by 10.3%, this means that 0.09% and 0.1% of all cancellations could have ostensibly substituted due to wildfire smoke. We view these substitutions as uncommon. Therefore, modeling cancellations as a binary decision is a reasonable representation of the choice that visitors face.

One additional reason not to model site substitution is that smoke conditions are spatially and temporally correlated. This correlation could wash out differences in smoke-related utility between choice alternatives, variation which is needed to properly identify the smoke parameter. Figures C1, C2, and C3 plot a visualization of this spatial and temporal correlation for Colorado, Oregon, and California. These figures sort campgrounds north to south on the vertical axis, while on the horizontal axis they plot days of the year during the summer months. Each tile represents a campground day and is colored according to the smoke conditions on those days. These figures reveal that, when one campground is smoke-affected, it tends to be the case that nearby campgrounds are also smoke-affected. Figure C4 also shows this relationship as a histogram for all campground days in the estimating dataset.

Figure C1: Spatial and temporal correlation of smoke in Colorado.

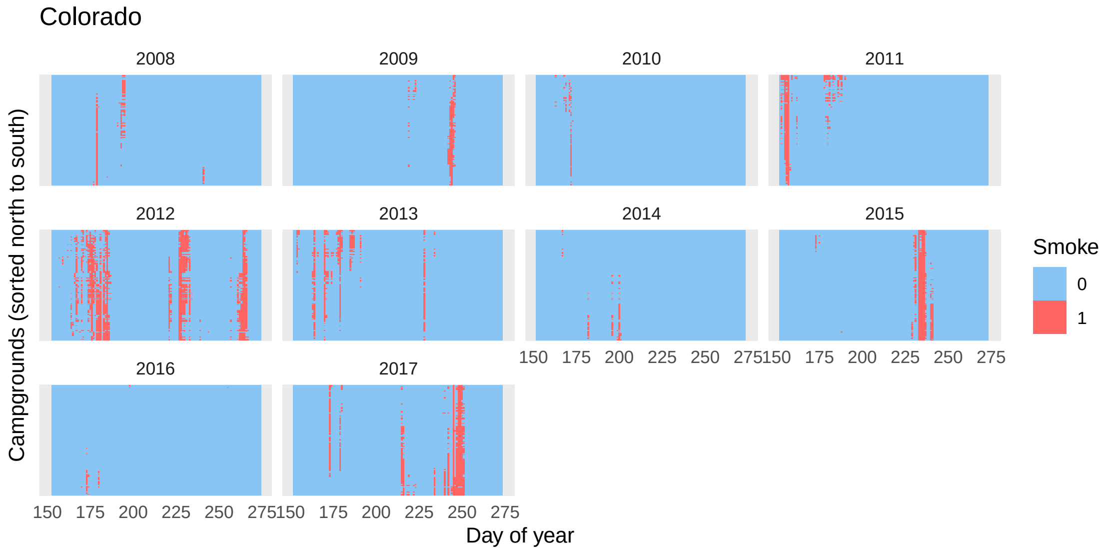


Figure C2: Spatial and temporal correlation of smoke in Oregon.

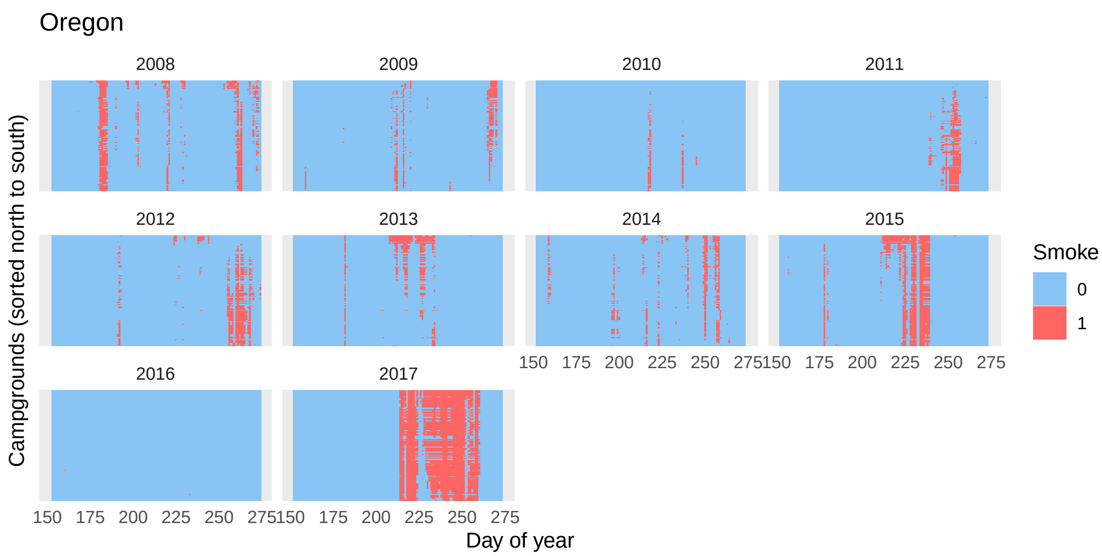


Figure C3: Spatial and temporal correlation of smoke in California.

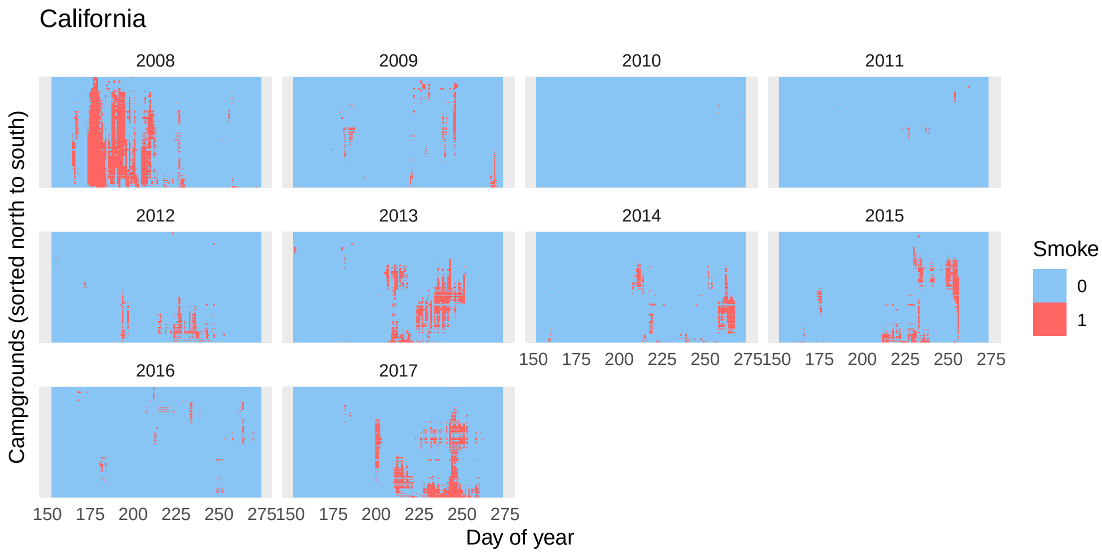
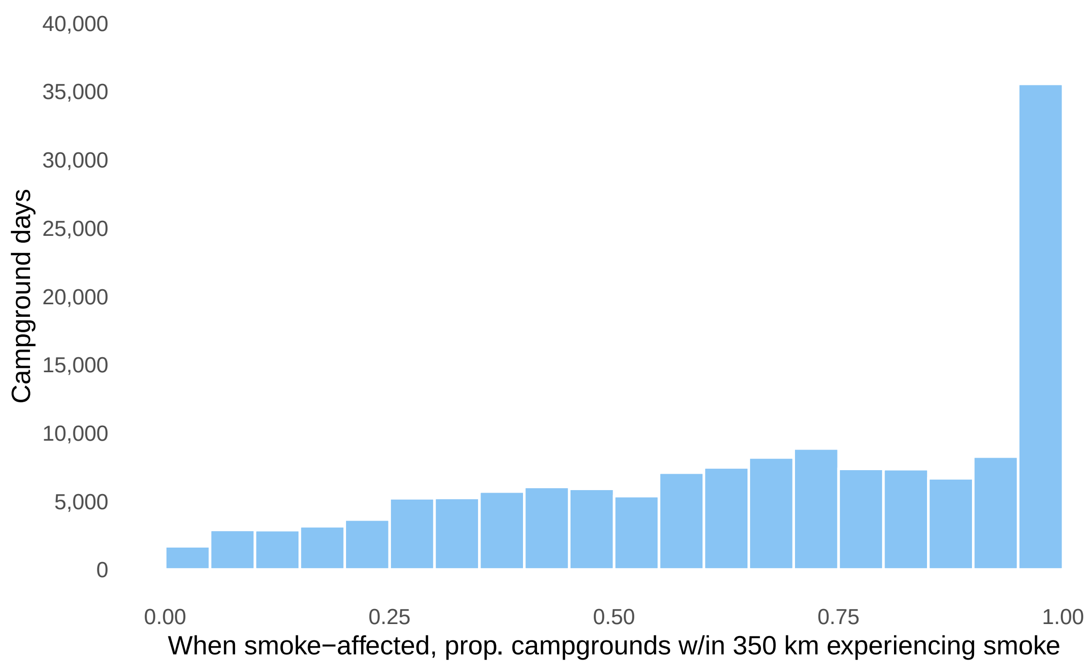


Figure C4: When a campground is smoke-affected, proportion of campgrounds within 350 km experiencing smoke in the same week.



The goal of this study is to value the non-market damages of wildfire smoke. The parameters of interest to estimate welfare damages are the marginal disutility of wildfire smoke and the marginal disutility in expenditure. Both of these parameters arise from the cancellation model, when site conditions become known to individuals. As we discuss in Section 3.1, the main purpose for the reservation estimation is to build the control function that accounts for preferences in the cancellation estimation. These preferences are likely correlated with travel cost in the selected sample, which we explore theoretically in Appendix D and show empirically in Figure 5. We are less concerned with the estimation of smoke in the reservation decision since the reservation occurs ahead of time, before smoke conditions are known. In addition, the binned travel cost zones provide variation to estimate how travel cost affects the likelihood of reservation. Overall, because we are less interested in site substitution for the reservation decision, we argue that the flexible computational advantages afforded by the zonal estimation justify this tradeoff. For more discussion, refer to Section 3.1.

Appendix D: Numerical example of sample selection correction

In Section 3.3 we proposed a control function approach to account for unobserved preferences $\tilde{\varepsilon}_{ijt}$ which 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 WTP, and correction using a control function. We show that WTP is only biased when preferences for the reservation and cancellation decisions are correlated, and when the counterfactual cancellation decision of non-reservers is unobserved. Further, 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. Finally, we demonstrate bias correction using the control function for $\tilde{\varepsilon}_{ijt}$ given in equation 13.

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 uniform. 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 uniform. 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 $U^R_{ij} = \alpha_j + \delta c_{ij} + \phi\mathbb{E}[s_j] + \varepsilon_{ij}$. 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 value 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 $R_{ij} = 1 \iff U_{ij}^R \geq U_{i0}^R$.

For the cancellation decision the user decides based on realized smoke conditions. Let the utility from cancellation be $U_{ij}^C = \alpha_j + \delta c_{ij} + \phi s_j + \upsilon_{ij}$. 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 errors υ_{ij} in the cancellation decision. The first is an independent error, $\upsilon_{ij}^{ind} \sim$ type I extreme value, which assumes the user's preferences in the cancellation decision are completely uncorrelated with their choice to have reserved. The second is a dependent error, $\upsilon_{ij}^{dep} = \rho\varepsilon_{ij} + \eta_{ij}$, which allows correlation of preferences between the reservation and the cancellation 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 cancellation decisions under both υ_{ij}^{ind} and υ_{ij}^{dep}, which we will denote 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 users that chose to make a reservation. However, under the Monte Carlo simulation, we can also examine the counterfactual cancellation decision of the non-reservers to see if they "would have" cancelled. We will show that, even with a dependent error υ_{ij}^{dep}, estimation of $\mathbb{P}(C_{ij} = 1)$ on the full sample (reservers and non-reservers) without observing ε_{ij} will still recover the true WTP since there is no selection effect. That is, the biased estimation of $\mathbb{P}(C_{ij} = 1 | R_{ij} = 1)$ comes from the fact that ε_{ij} and c_{ij} are correlated in the selected sample, not the full sample.

Table D1: Example of users' reservation and cancellation decisions.

R_{ij}	C_{ij}^{ind}	C_{ij}^{dep}	N
0	0	0	9,004
0	0	1	15,135
0	1	0	5,453
0	1	1	12,849
1	0	0	24,759
1	0	1	8,215
1	1	0	16,141
1	1	1	8,444

Table D1 shows an example of users' reservation and cancellation decisions from one iteration of the Monte Carlo. In this case non-reservers were more likely to cancel with correlated errors and reservers were less likely to cancel with correlated errors. This result is driven by their initial preferences about the site since reservers have a higher ε_{ij}. Figure D1 illustrates this point by comparing the ε_{ij} of reservers to the total population.

Figure D1: Example distribution of ε_{ij} for reservers and for all users.

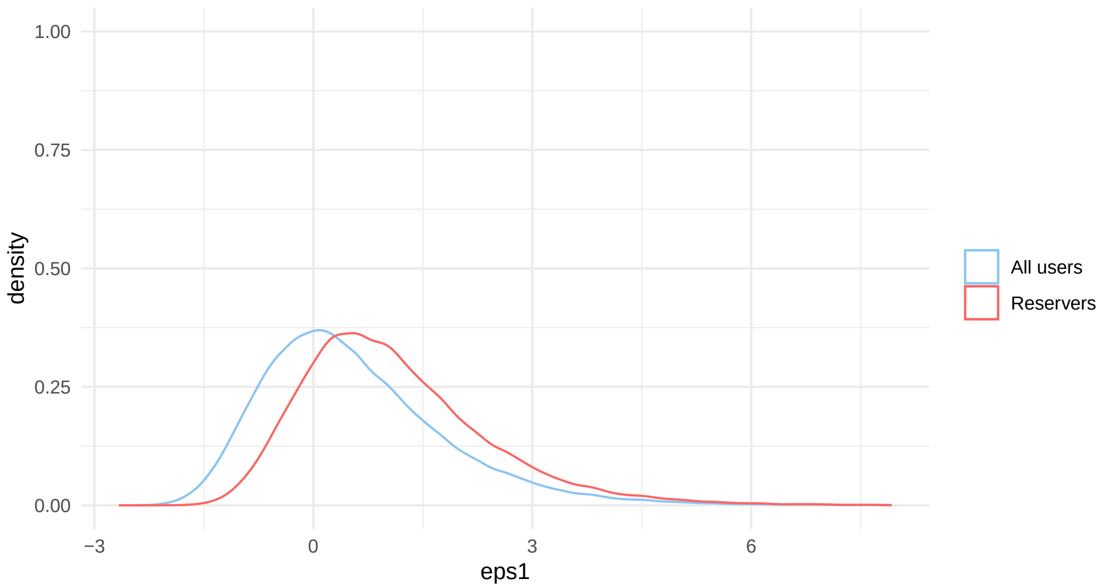


Figure D2 plots the cancellation rate for reservers at various distances for smoke and non-smoke conditions. Results are shown under both independent and dependent errors. The figure illustrates several key points. First, the overall cancellation rate is lower with dependent errors, as indicated by the intercept of the fitted golden line. Users that 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 with independent and dependent errors. Third, the effect of travel cost, which is the slope of the golden fitted line, is attenuated when errors are dependent. This attenuation illustrates that the selection effect likely operates through positive correlation between ε_{ij} and travel cost.

We can further demonstrate this relationship by regressing distance on ε_{ij} in the full sample and the selected sample. Table D2 shows an example of such a regression using one draw from the Monte Carlo simulation. Travel cost and distance are correlated among the selected sample, but not among all users.

Figure D2: Example cancellation rate for reservers by distance and smoke.

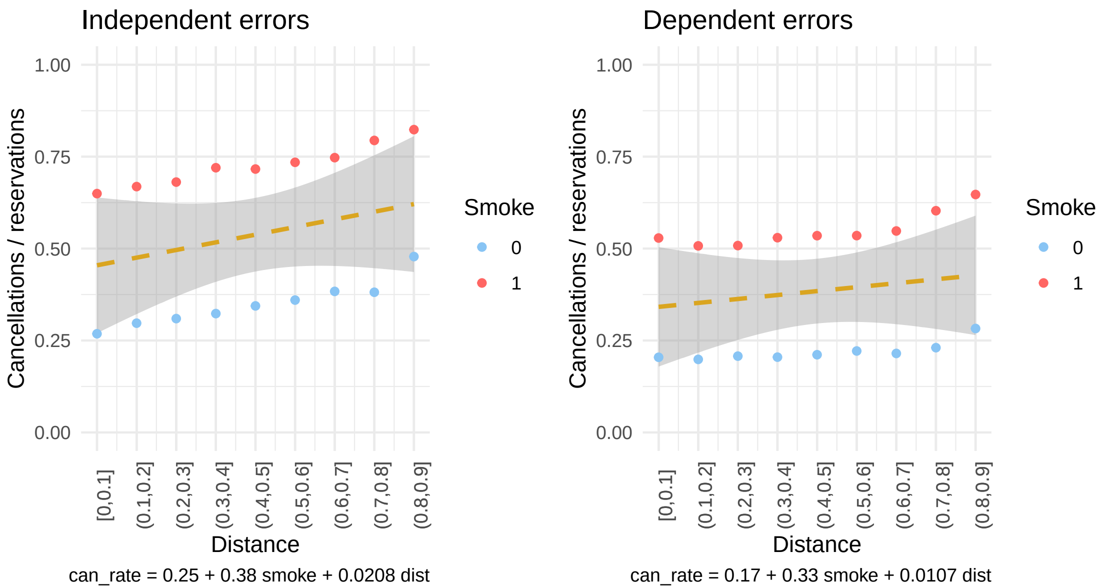


Table D2: Example regression of distance on ε_{ij} in the full and selected sample.

	(1)	(2)
ε_{ij}	-0.0003	0.005^{**}
	(0.0004)	(0.001)
Intercept	0.398^{**}	0.384^{**}
	(0.001)	(0.001)
Observations	100,000	57,559
R^2	0.00001	0.001
Users	All users	Reservers
Note:		*p<0.05; **p<0.01.

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

Table D3 shows an example from one iteration of the Monte Carlo simulation. 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 dependency of the error 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}. Regression 4 demonstrates that sample selection is not an issue if the user's preferences at cancellation are uncorrelated with their preferences at the time of reservation. Finally, column 5 shows that WTP estimates are biased when preferences are correlated and the sample is selected.

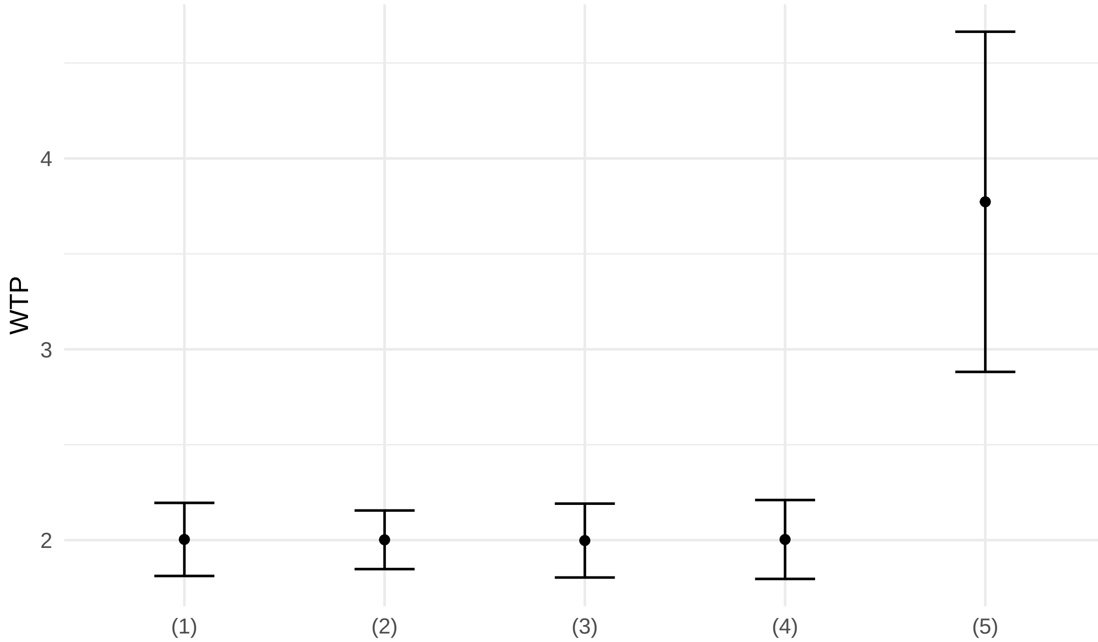
Table D3 uses only one draw from the full set of 10,000 random draws. In Figure D3 we show the same results over the full set of 10,000 draws. The logic holds: estimation of WTP is biased only under a selected sample with correlated preferences.

Table D3: Example regressions of reservation and cancellation decisions under various samples and error structures.

	(1)	(2)	(3)	(4)	(5)
Distance	-0.7998**	-0.8131**	-0.5826**	-0.8261**	-0.2727**
	(0.0405)	(0.0426)	(0.0414)	(0.0563)	(0.0606)
$\mathbb{E}$[Smoke]	-1.5951**				
	(0.0515)				
Smoke		-1.5814**	-1.2402**	-1.6025**	-1.4407**
		(0.0160)	(0.0155)	(0.0211)	(0.0205)
Intercept	0.9988**	1.0091**	0.7576**	1.0190**	1.4310**
	(0.0214)	(0.0189)	(0.0183)	(0.0245)	(0.0266)
N	100,000	100,000	100,000	57,559	57,559
Dep. var.	R_{ij}	C_{ij}	C_{ij}	C_{ij}	C_{ij}
Users	All users	All users	All users	Reservers	Reservers
Error	ε_{ij}	υ_{ij}^{ind}	υ_{ij}^{dep}	υ_{ij}^{ind}	υ_{ij}^{dep}
WTP	1.99	1.95	2.13	1.94	5.28

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

Figure D3: Monte Carlo 10,000 simulated regressions of reservation and cancellation decisions.



	(1)	(2)	(3)	(4)	(5)
WTP	2.00**	2.00**	2.00**	2.00**	3.77**
	(0.10)	(0.08)	(0.10)	(0.11)	(0.46)
Dep. var.	R_{ij}	C_{ij}	C_{ij}	C_{ij}	C_{ij}
Users	All users	All users	All users	Reservers	Reservers
Error	ε_{ij}	υ_{ij}^{ind}	υ_{ij}^{dep}	υ_{ij}^{ind}	υ_{ij}^{dep}

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

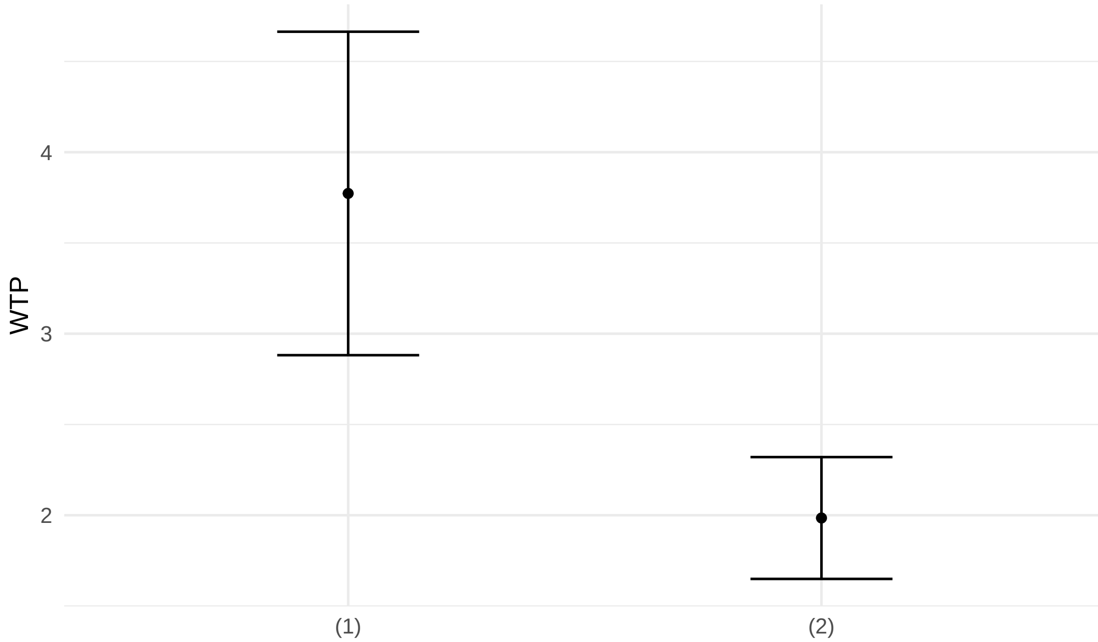
We next demonstrate the bias correction of the estimand $\tilde{\varepsilon}_{ij}$ derived in equation 13. We first estimate the reservation decision, then use the fitted values of $\mathbb{E}[V_{ij}]$ to form $\tilde{\varepsilon}_{ij}$. Table D4 shows an example from one draw of the 10,000 simulations. In this example we see that the smoke coefficient is unaffected by the bias corrector. Instead, the value of the intercept is reduced and the value of the distance coefficient is inflated. In this single random draw the true WTP was not exactly recovered. However, over the full set of 10,000 simulations we see that the inclusion of $\tilde{\varepsilon}_{ij}$ results in unbiased estimation. Figure D4 shows WTP results for the full set of 10,000 simulations. The inclusion of the $\tilde{\varepsilon}_{ij}$ estimator resulted in recovery of the true WTP. This result lends support to the use of this bias corrector in the empirical dataset.

Table D4: Example regression of cancellation decision for reservers using bias correction.

	(1)	(2)
Intercept	1.4310**	0.7590**
	(0.0266)	(0.0813)
Smoke	-1.4407**	-1.4417**
	(0.0205)	(0.0205)
Distance	-0.2727**	-0.6088**
	(0.0606)	(0.0718)
$\tilde{\varepsilon}_{ijt}$		-0.6862**
		(0.0786)
N	57,559	57,559
WTP	5.28	2.37
2-step estimator	None	$\tilde{\varepsilon}_{ijt}$

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

Figure D4: Monte Carlo 10,000 simulated regressions showing bias correction.



	(1)	(2)
WTP	3.77**	1.98**
	(0.45)	(0.17)
2-step estim.	None	$\tilde{\varepsilon}_{ij}$

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

There are several key assumptions from this exercise. First, the data generating process asserts that users react the same way to expected smoke as to realized smoke. That is, the coefficient for expected smoke and realized smoke is the same. This assumption may not hold for real users; it is reasonable to believe that decision makers may respond differently to expected conditions than to realized conditions. Still, the purpose of $\tilde{\varepsilon}_{ij}$ is to account for selection from the first stage and should therefore serve as an appropriate control function, regardless of whether the coefficients are identical between stages.

The second key assumption is that the decision maker selects from a single choice alternative. This setup matches our conceptual framework in Section 3, where we assumed a binary site choice. The main reason for this assumption is for computational tractability of the dataset, which features millions of users and nearly one thousand campgrounds over eight years. For an extended treatment of this matter see Appendix C.

Appendix E: 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. Their process is as follows: for B bootstraps and G clusters, (1) sample with replacement G times from the original sample of clusters, (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 an adequate number of bootstraps. Table E1 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 E1 and E2 plot the bootstrapped coefficients visually.

Table E1: 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 FE		Yes	Yes	Yes
Day-of-week FE		Yes	Yes	Yes
Campground x week-of-year FE		Yes	Yes	Yes
Year FE		Yes		
State x year FE			Yes	
Campground x year FE				Yes

Figure E1: Distribution of estimated smoke coefficient from models (1) through (4) in bootstrapped estimation of $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ with sample selection correction. Red line indicates mean.

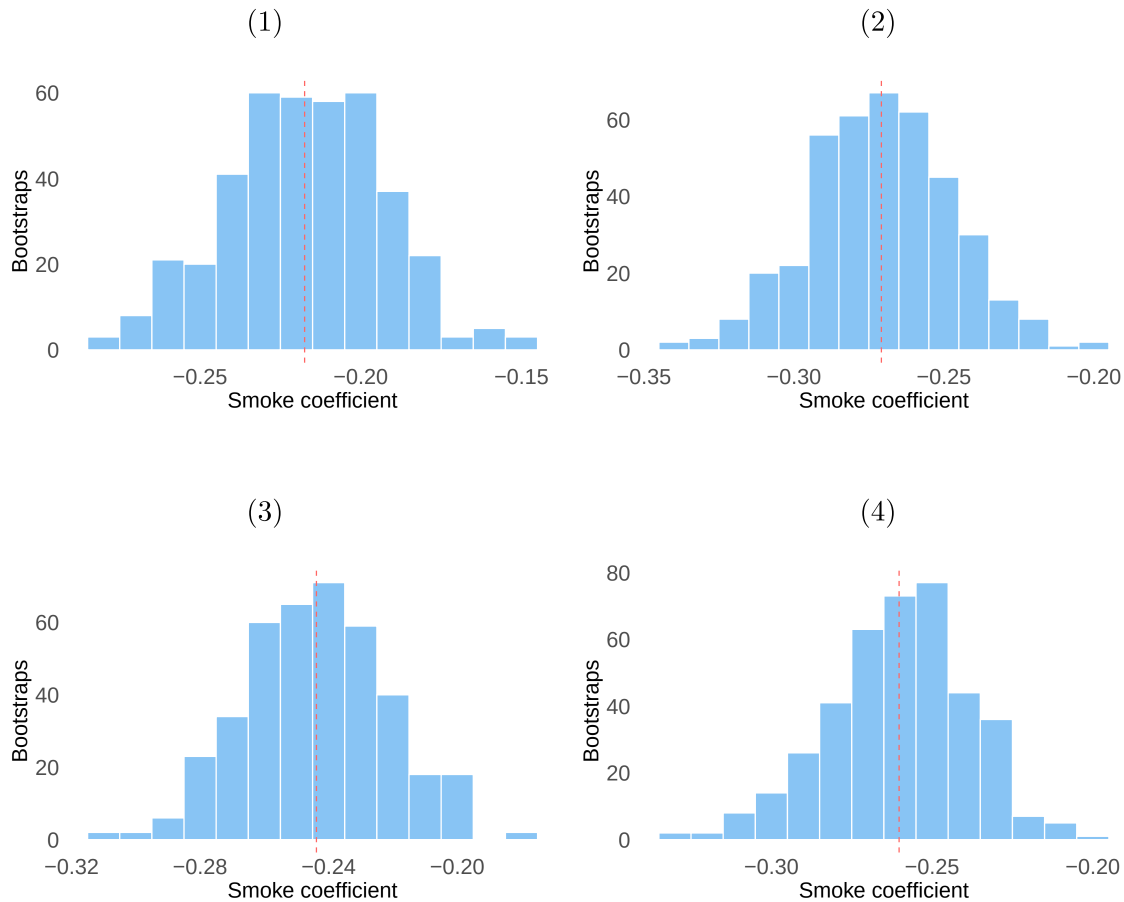
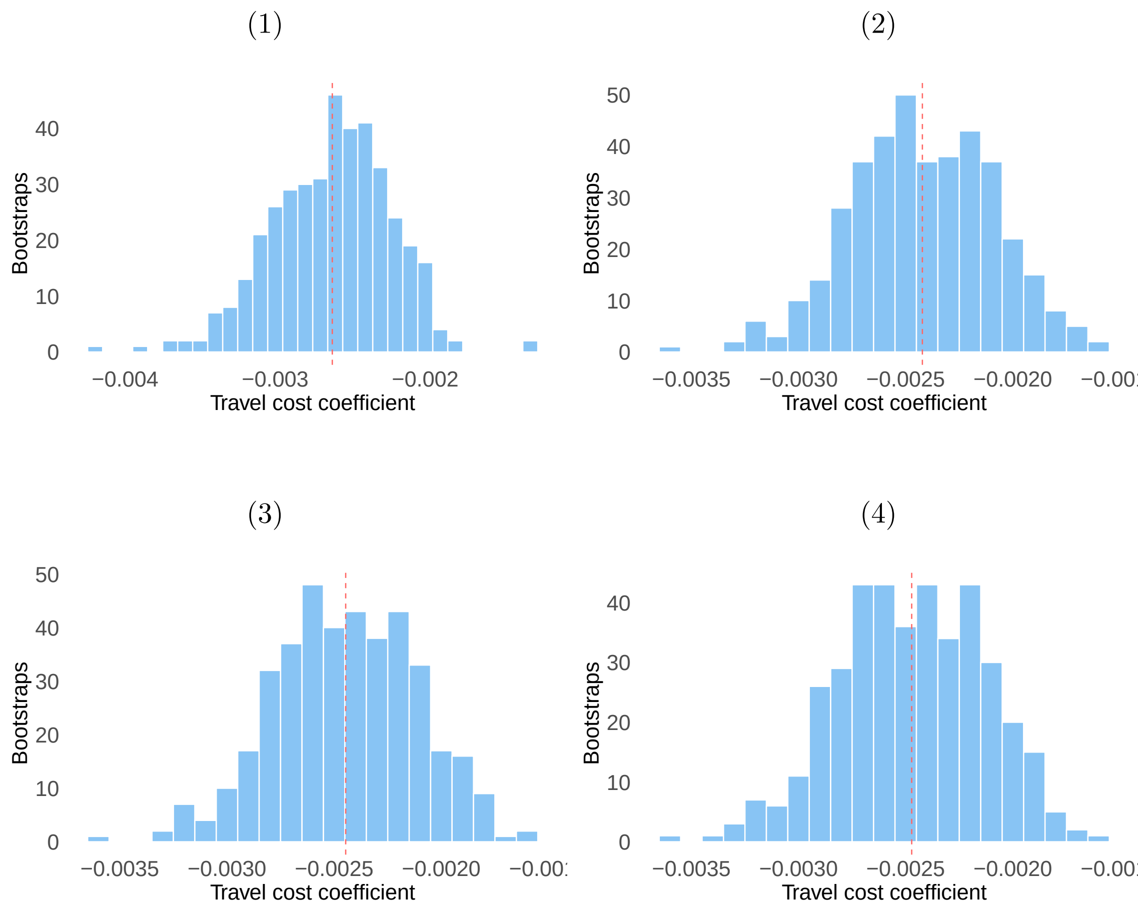


Figure E2: Distribution of estimated travel cost coefficient from models (1) through (4) in bootstrapped estimation of $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$ with sample selection correction. Red line indicates mean.



Appendix F: Testing the influence of no shows in cancellations

One may be concerned that some recreationists do not formally cancel their reservation when they decide not to complete a trip. Unreported no shows threaten the identification of any WTP that is based on cancellations, since it could potentially underestimate cancellations. While most of the campgrounds in the Recreation.gov dataset do not report check ins or no shows, a subset of campgrounds do.

In this section we compare estimates at these select campgrounds with and without the inclusion of no shows. We demonstrate that although no shows are not infrequent at these campgrounds, omitting them does not influence measures of responses to smoke and travel cost. This analysis should mitigate some concern that we understate avoidance behavior.

In total, just 36 out of 999 campgrounds (3.6%) report no shows. However, these campgrounds represent a large proportion of the reservations used in the cancellation estimation. Of the reservations made greater than a week ahead of time, 2,188,444 reservations were at non-no show facilities (80.3%), while 535,496 were reservations at facilities that report no shows (19.7%).

To gauge the importance of no shows in the cancellation estimation, Figure F1 and Table F1 report the share of all cancellations that are no shows at each campground. While most of the overall dataset is comprised of US Forest Service campgrounds, Table F1 shows that many campgrounds reporting no shows are managed by the National Park Service and US Army Corps of Engineers. For most of these campgrounds, no shows represent less than 15% of all cancellations, though at some campgrounds no shows represent nearly a third of all cancellations.

Figure F1: No shows as a proportion of all cancellations among campgrounds reporting no shows.

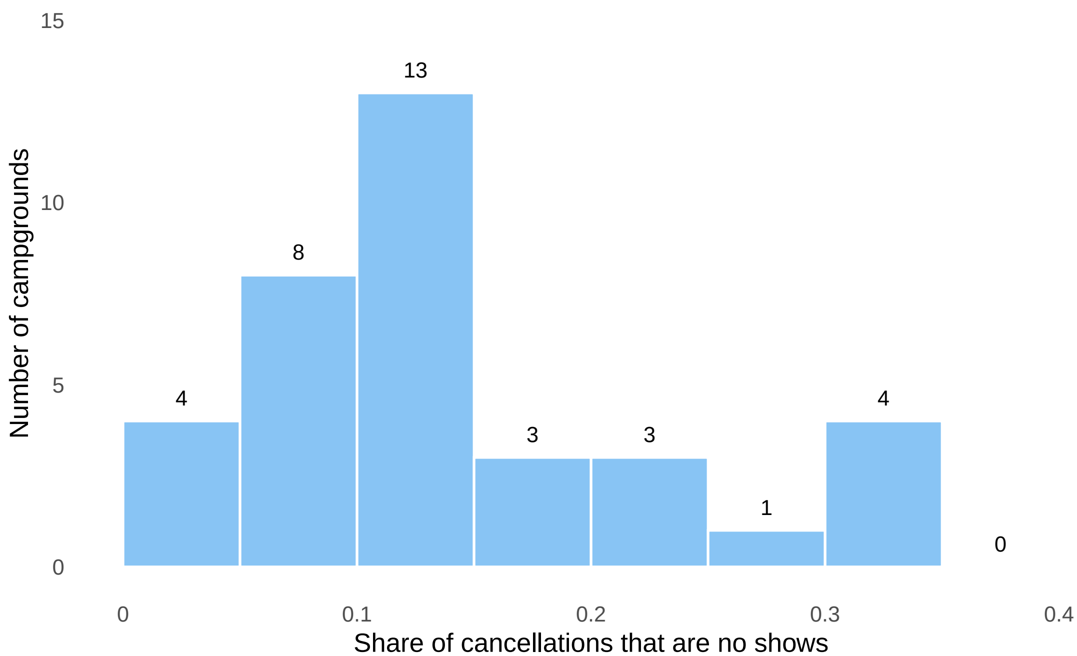


Table F1: Campgrounds reporting no shows.

Campground	Recreation area	State	Agency	No show % of cancellations
Aspenglen Campground	Rocky Mountain National Park	CO	NPS	33.69
Glacier Basin Campground	Rocky Mountain National Park	CO	NPS	32.8
Moraine Park Campground	Rocky Mountain National Park	CO	NPS	32.51
Watchman Campground	Zion National Park	UT	NPS	32.36
Mather Campground	Grand Canyon National Park	AZ	NPS	27.59
Schwarz Park	Dorena Lake	OR	USACE	24.62
Buckhorn	Black Butte Lake	CA	USACE	23.37
Springy Point	Albeni Falls Dam	ID	USACE	20.44
Hood Park	McNary Lock And Dam	WA	USACE	18.51
Hodgdon Meadow	Yosemite National Park	CA	NPS	18.49
Fishhook Park	Ice Harbor Lock	WA	USACE	15.09
Dinkey Creek	High Sierra RD	CA	USFS	14.86
Meeks Bay	Lake Tahoe Basin	CA	USFS	14.03
Serrano	Big Bear	CA	USFS	14.02
Crane Flat	Yosemite National Park	CA	NPS	13.89
Riley Creek	Albeni Falls Dam	ID	USACE	13.83
Dogwood	Arrow Head	CA	USFS	13.24
Wawona	Yosemite National Park	CA	NPS	13.2
Charbonneau Pk	Ice Harbor Lock	WA	USACE	12.74
Pine Meadows Campground	Cottage Grove Lake	OR	USACE	12.65
North Rim Campground	Grand Canyon National Park	AZ	NPS	12.47
Kyen Campground	Lake Mendocino	CA	USACE	10.83
Nevada Beach Campground	Lake Tahoe Basin	CA	USFS	10.81
William Kent Campground	Lake Tahoe Basin	CA	USFS	10.1
Lepage Park	John Day Lock	OR	USACE	9.74
Fish Creek Campground	Glacier National Park	MT	NPS	8.5
Rancheria	High Sierra RD	CA	USFS	7.61
Oh Ridge	Mono Lake RD	CA	USFS	7.19
Deer Creek	High Sierra RD	CA	USFS	7.1
Diamond Lake	Diamond Lake RD	OR	USFS	6.01
Downstream	Fort Peck Project	MT	USACE	5.68
Pinecrest	Summit RD	CA	USFS	5.47
Fallen Leaf Campground	Lake Tahoe Basin	CA	USFS	4.56
Acorn Campground	New Hogan Lake	CA	USACE	4.46
Lodgepole Campground	Sequoia And Kings Canyon National Park	CA	NPS	2.5
Dorst Creek Campground	Sequoia And Kings Canyon National Park	CA	NPS	2.33

We would like to use this subset of campgrounds to demonstrate that unreported no shows likely do not matter in the full sample. Table F2 tests whether these campgrounds are systematically different than the full sample. The table shows that these no show campgrounds tend to have higher cancellation rates under both smoke and non-smoke conditions.

However, the estimates in Table F3 should alleviate concerns that no shows are influential in cancellation estimates. We test four models with the same sets of fixed effects, but vary the estimating sample. In column 1 we include all campgrounds and estimate $\mathbb{P}(C_{ijt} = 0|R_{ijt} = 1)$. Column 2 removes no shows from the dataset, finding that WTP is unchanged. In column 3 we allow smoke and travel cost to respond differentially for no show and non-no show campgrounds, but include no shows in the dataset. This model shows that no show and non-no show campgrounds have different overall measures of WTP. Finally, column 4 removes no shows from the dataset. Comparing the WTP of no show campgrounds with and without the inclusion of no shows, WTP is virtually unchanged. This analysis should alleviate concerns that no shows influence the estimate of WTP in the full sample.

Table F2: Cancellation rate mean by campground type and by smoke.

	All campgrounds	Non-no show campgrounds	No show campgrounds	t-statistic
Baseline	0.09	0.09	0.13	(9.44)
No. of res.	2,380,606	1,898,955	481,651	
Smoke	0.13	0.12	0.19	(2.02)
No. of res.	343,334	289,489	53,845	
t-statistic	(13.41)	(13.19)	(3.06)	

Notes: The righthand column gives the t-statistic for the difference in mean cancellation rates by campground type in either smoke or non-smoke conditions. The bottom row gives the t-statistic for the difference in mean cancellation rate for smoke and non-smoke days among the different campground types.

Table F3: $\mathbb{P}(C_{ijt} = 0 | R_{ijt} = 1)$, testing effect of no shows on cancellation.

	(1)	(2)	(3)	(4)
Smoke in week of arrival	-0.2613**	-0.2659**		
	(0.0215)	(0.0215)		
Smoke x 1(Non-no show campground)			-0.2608**	-0.2605**
			(0.0188)	(0.0188)
Smoke x 1(No show campground)			-0.2628**	-0.2846**
			(0.0689)	(0.0736)
Travel cost (dollars)	-0.0024**	-0.0025**		
	(0.0003)	(0.0004)		
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})	-7.8194**	-7.9970**	-7.8180**	-7.9907**
	(0.8239)	(0.8659)	(0.8223)	(0.8631)
High temp. (degrees C)	0.0306**	0.0308**	0.0306**	0.0308**
	(0.0022)	(0.0022)	(0.0022)	(0.0022)
Low temp. (degrees C)	-0.0252**	-0.0254**	-0.0252**	-0.0254**
	(0.0025)	(0.0025)	(0.0025)	(0.0025)
Precip. in week of arrival (mm)	-0.0057**	-0.0058**	-0.0057**	-0.0058**
	(0.0009)	(0.0009)	(0.0009)	(0.0009)
$\tilde{\varepsilon}_{ijt}$	-0.0370**	-0.0370**	-0.0376**	-0.0373**
	(0.0126)	(0.0131)	(0.0134)	(0.0137)
WTP	107.95**	107.92**		
	(17.14)	(17.48)		
WTP, non-no show campgrounds			103.82**	103.9**
			(17.39)	(17.75)
WTP, no show campgrounds			116.66**	119.75**
			(32.1)	(33.5)
No shows included?	Yes	No	Yes	No
N	2,688,739	2,677,763	2,688,739	2,677,763
Campground FE	Yes	Yes	Yes	Yes
Day-of-week FE	Yes	Yes	Yes	Yes
Campground x week-of-year FE	Yes	Yes	Yes	Yes
Campground x year FE	Yes	Yes	Yes	Yes

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

Appendix G: 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. To do so 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 of Engineers, and the National Association of State Park Directors. Each of these agencies 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 data is reported 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.[33] This dataset reports total monthly visitation at all National Parks, National Monuments, National Recreation Areas, and other lands managed by the National Park Service. In the western states, 27 National Parks are included in the Recreation.gov dataset, while 82 are not. For the 27 parks in the Recreation.gov dataset, we determine each park's monthly proportion of campers that were affected by wildfire smoke. We then multiply this proportion by each park's monthly visitation from the Annual Summary Reports to infer the total number of visits affected by smoke. For the 82 parks not in the Recreation.gov dataset, 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.[34] These data report visitation at all National Forests at an annual level. In the west, 70 National Forests are included in the Recreation.gov dataset, while 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 Manage-

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

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

ment Information System (RMIS).[35] We contacted the program administrator and received data on site by year visitation for all BLM sites.[36] Most visitation to BLM lands 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 the agency's Value to the Nation (VTN) reports.[37] 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.

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

[36] Ridenhour, L. & Leitzinger, K. Bureau of Land Management. Personal correspondence.

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