



Assessing the Vulnerability of California Water Utilities to Wildfires

Juhee Lee¹ · Mehdi Nemati² · Jose J. Sanchez³

Received: 24 March 2022 / Accepted: 4 July 2022
© The Author(s) 2022

Abstract

Wildfires are becoming more frequent and destructive in California, and it is essential to quantify their potential impacts on drinking water utilities. This study aims to measure the severity of wildfires in each California water utility based on the exposure frequency and the extent of area burned by wildfires in each service area. Our quantitative models show an association between water utility characteristics and their vulnerability to wildfires. Findings indicate that wildfire vulnerability is higher in government-owned utilities than private ones, utilities primarily relying on surface water than groundwater, and utilities using local-sourced water than purchased water. Also, we find a stronger association between wildfire vulnerability and large utilities in terms of population served than small or medium ones. Regarding geography, we find wildfire vulnerability is higher in southern and coastal California utilities than in Northern and inland California. These results help water utilities and land managers identify vulnerable locations and develop wildfire management and disaster preparedness strategies.

Keywords California · Urban water utilities · Wildfires severity

1 Introduction

Over the past several decades, climate change, extensive droughts, population growth, urbanization, and land-use changes have raised wildfire risks and severity throughout the western US. In recent years, California has experienced a substantial increase in the

✉ Mehdi Nemati
mehdin@ucr.edu

Juhee Lee
juheel@ucr.edu

Jose J. Sanchez
jose.sanchez2@usda.gov

¹ Postdoctoral Scholar, School of Public Policy, University of California, 900 University Ave. Riverside, Interdisciplinary South 4120, Riverside, CA 92521, USA

² Assistant Professor of Cooperative Extension in Water Resource Economics and Policy, School of Public Policy, University of California, 900 University Ave. Riverside, Interdisciplinary South 4109, Riverside, CA 92521, USA

³ Research Economist, USDA Forest Service Pacific Southwest Research Station, 4955 Canyon Crest Drive, CA 92507 Riverside, USA

frequency and intensity of wildfires from ongoing drought conditions. According to the California Department of Forestry and Fire Protection (2021), in 2018, California experienced the deadliest and most destructive wildfires—the Camp Fire¹ and the Mendocino Complex fire. However, this trend continued in 2021; California has a new record for the largest fire in the state's history. At that time, the Dixie Fire, named for the road where it started, had burned more than 963,309 acres in Northern California, taking hundreds of buildings down with it and threatening nearly 14,000 structures.

Correspondingly, the risk of wildfire damage to humans and environmental and ecosystem health is increasing. The impacts of wildfire on soil, wildlife, watersheds, air quality, vegetation, and human health are well-documented (DeBano 2000; DeBano et al. 2003; Ice et al. 2004; Neary et al. 2005; Kim et al. 2008; Seibert et al. 2010; Prior and Eriksen 2013; Sham et al. 2013; Batelis and Nalbantis 2014; Liu et al. 2016; Bladon 2018; Hallema et al. 2018). However, studies on the risk of wildfires on water utilities remain limited. Wildfires can impact water utilities² through several channels, including changes in water availability, source water quality from ash build-up (both surface and groundwater sources), soil erosion, and fire debris. Additionally, costs can be substantial for water utilities to treat and disinfect water for drinking effectively, damages to infrastructure and pipelines, and changes in taste, color, and smell of treated drinking water (Hohner et al. 2016; Robinne et al. 2019, 2016; Mansilha et al. 2019).

Despite all the potential damages listed above and the frequency and intensity of the wildfires in California, we have found a handful of recent studies documenting California water utilities' vulnerability to wildfires (Schulze and Fischer 2020; Robinne et al. 2019). In this paper, we measure the severity of wildfires and derive the wildfire vulnerability for each water utility service area. To accomplish this, we combine and use 30-year historical records on California wildfires and water utility boundaries and characteristics (e.g., ownership, primary source, trade, geographic locations, and population size). Specifically, in this paper, we: (i) calculate wildfire severity for each period and water utility by dividing the total acres burned by wildfires by the frequency of the wildfires (i.e., the number of events that occurred); (ii) assess the vulnerability of California water utilities to wildfires by estimating the association between water utility characteristics and the degree of wildfire severity; and (iii) examine how this vulnerability changes over time and across utilities.

Our findings indicate that government-owned utilities are at a higher level of wildfire risk than those owned privately. Similarly, we find higher wildfire risk in utilities that primarily use surface water than groundwater, those relying on local sources than purchased water, those located in Southern or coastal California than those in Northern or inland California, and the large utilities than the small- or medium-sized utilities. Our findings identify water utilities at a higher risk of wildfires and provide information to help develop wildfire management and disaster preparedness strategies. These findings highlight the need for higher caution and coordinated management by relevant agencies to utilities with the following characteristics, which are more susceptible to wildfire risk.

¹ The CalFire reports the Camp Fire burned a total of 153,336 acres, destroyed nearly 18,804 structures and occurred at least 85 civilian fatalities.

² Water utilities means water-related entities that own and operate public drinking water systems and provide water to their customers. They also called public water agencies, public water systems, water systems, and water agencies.

Table 1 Description for the selected variables

<i>Outcome Variables</i>	Description
Exposure frequency	Total number of wildfires that occurred for the specific period (number counted)
Total area burned	Summarized total land acres burned for the specific period (acres)
Severity	Total land acres burned divided by the total number of events that occurred (1,000 acres) (i.e., wildfire extent/wildfire exposure frequency)
Log severity	Taking the logarithm of the severity
<i>Explanatory Variables</i>	Description
Ownership	Local government-owned utility (1 = government, 0 = private)
Primary source	Utility relying primarily on groundwater as a primary water source (1 = groundwater, 0 = surface water)
Trade	Utility that is mainly dependent on purchased water (1 = purchased, 0 = owned)
1st geographic location	Utility located in Southern California (1 = SoCal, 0 = NoCal)
2 nd geographic location	Utility located in coastal California (1 = coastal, 0 = inland)
Population size ^a	Population size served by an utility (i.e., 0 = small-populated level, 1 = medium-populated level, 2 = large-populated level utility)

^aA small-populated level is when utilities serve less than 3,300 people; a medium-populated level is when utilities serve less than 100,000 people; and a large-populated level denotes when utilities serve more than 100,000 people

2 Data Description

Our study focuses on California's 7,500 water utilities throughout the state. Our primary study data is from California wildfires that occurred from 1991 to 2020 within the boundaries of water utility service areas. Our final dataset for the analysis combines (1) spatial-temporal data on wildfires obtained from the California Fire and Resource Assessment Program (FRAP) (2021) and the National Interagency Fire Center (2021); data on approximately 7,500 water utility boundaries,³ and characteristics acquired from the California State Water Resources Control Board, the US Environmental Protection Agency, and California Department of Water Resources. We used ArcGIS 10.6 and programming languages to spatially merge these data to the water utility and generate outcome variables.

Our final pooled cross-sectional data contains information on wildfire-related factors used as an outcome variable and on characteristics of water utilities used as explanatory variables (Table 1). The wildfire-related factors that the FRAP provides include the date of wildfire and its name, location, causes, acres burned, and the date of fire suppression. This data is an annual digital record of fire perimeters in California generated by compiling fire perimeters and establishing an ongoing fire perimeter data capture process. Upon release publicly, the data is current as of the last calendar year and is limited to fires already suppressed by the FRAP.

Using these factors, we measure the historical fire exposure for each water utility service area based on the number and acres of historical wildfires in each utility service area and

³ The water utilities' service area boundaries in California do not change much over time.

Table 2 Descriptive statistics for the selected variables by period

	Last 30-years (1991–2020)		Last 20-years (2001–2020)		Last 10-years (2011–2020)		Last 5-years (2016–2020)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Wildfires								
Exposure frequency	14	172	10	119	4	45	3	36
Total land acres burned (1,000)	468	6,580	444	6,320	173	3,510	164	3,470
Severity (1,000)	12	54	12	56	10	55	100	56
Water Utility Characteristics								
Government (Yes = 1, No = 0)	0.35	0.48	0.35	0.48	0.36	1.21	0.32	1.16
Groundwater (Yes = 1, No = 0)	0.80	0.40	0.80	0.40	0.35	0.48	0.35	0.48
Purchased (Yes = 1, No = 0)	0.08	0.28	0.08	0.28	0.80	0.40	0.80	0.40
Southern CA (Yes = 1, No = 0)	0.23	0.42	0.23	0.42	0.08	0.28	0.08	0.28
Coastal (Yes = 1, No = 0)	0.40	0.49	0.40	0.49	0.23	0.42	0.23	0.42
Size-Medium (Yes = 1, No = 0)	0.13	0.34	0.13	0.34	0.40	0.49	0.40	0.49
Size-Large (Yes = 2, No = 0)	0.02	0.14	0.02	0.14	0.13	0.34	0.13	0.34

measure the wildfire severity. We expect that wildfire severity varies by each utility's characteristics, thereby assessing their vulnerability to wildfires.

The raw historical data of wildfires is available from 1878 to 2020. To examine different effects by different periods according to one of our research questions, we created four subsample datasets from the latest data point, which classify: the last 30-year period (1991–2020), last 20-year period (2001–2020), last 10-year period (2011–2020), and last 5-year period (2016–2020). We combine each of these fire-related datasets with data on 4,957 water utilities. Along with each subsample, we use data on a set of water utility characteristics such as ownership (local government versus private sector), the primary water source (groundwater versus surface water), and water imports (purchased water versus local sources of water), geographic location (Northern versus Southern California, and inland versus coastal California), and population size served by utilities (small, medium, and large). More specifically, we defined population size served by utilities as (1) a small-populated level when utilities serve less than 3,300 people, (2) a medium-populated level when utilities serve less than 100,000 people, and (3) a large-populated level when utilities serve more than 100,000 people. Table 2 provides the descriptive statistics for the selected variables by different periods. Note that each subsample data composed of the same outcome variable and explanatory variable show missing values on the FRAP survey or zero values because there were no wildfires.

3 Methodology

The wildfire survey has the nature of non-periodic outbreaks; accordingly, our data include missing values (i.e., no survey implemented) and zero values (i.e., no wildfire occurred at all). Since wildfire severity takes a positive censored distribution starting from zero, the estimation without considering censoring brings bias to the estimated coefficients from regression analysis using the Ordinary Least Squares (OLS) estimator. In a similar vein, when analyzing the characteristics of water utilities that determine wildfire vulnerability, the data contains two kinds of information: (1) whether wildfire occurs and (2) the severity

of wildfire. The result may lead to a biased inference if OLS regression analysis is performed on only the latter information. That is because it is not a random selection without controlling the cases with no wildfire outbreaks.

Therefore, when the range of values taken by the outcome variable is limited and complex information is embedded, an analysis model that considers these features must be applied. Considering these, we applied a Tobit model⁴ using one of the maximum likelihood estimations (MLE) and a Heckman model using a two-step procedure⁵ for the analysis. Both models have the advantage of bias adjustment, but the Tobit model is based on the assumption that the effects of explanatory variables on the occurrence of wildfire and the magnitude of wildfire are identical. The Heckman model analyzes two separate steps—whether wildfire occurs and the severity of wildfire. As mentioned in the data section, using the same set of explanatory variables, we made separate estimates for wildfire records from 1991 to 2020 by classifying them into four different periods (i.e., during a 5-year, 10-year, 20-year, and 30-year period) to compare how the wildfire severity varies with the different periods.

3.1 Tobit Model

The standard method to correct a limited outcome variable is to estimate the Tobit model by MLE. Given that our outcome variable, wildfire severity, cannot be negative, and that it is a continuous variable over a specific period, we use the Tobit model censored at zero to explain the heterogeneity of water utility characteristics affecting wildfire severity. We take the log of the outcome variable not only for dealing with skewness and getting closer to a normal distribution but also for linearizing the non-linear relationship between the outcome and explanatory variable. Since wildfire at zero cannot have log transformation, the natural logarithm was taken after adding one to all observations (MaCurdy and Pencavel 1986). By doing this, the value of one is converted to zero again when applying the log, so it does not distort the meaning that no wildfire was observed or that no wildfire occurred in the analysis.

This wildfire severity is dependent on the relevant influencing factors, and the Tobit model shows the effect on the wildfire's occurrence and the effect on its severity by these factors are equal and estimates the wildfire severity as a single function as:

$$y_i^* = x_i' \beta + \varepsilon_i \quad \varepsilon_i \sim N(0, \sigma^2) \\ i = 1, 2, \dots, n \quad (1)$$

where the variables y_i and y_i^* are receptively observable wildfire severity and a latent wildfire severity (i.e., it is an unobservable and random variable whose realized value is hidden)

⁴ The use of this mechanism to explain that expenditure cannot be negative values in family expenditure surveys was introduced by Tobin (1958). Since then, Amemiya (1985) classified this Tobit model into five basic types according to the form of the likelihood function.

⁵ The Heckman (1976) two-step procedure was used in criminology to deal with selection effects, which may cause biased inferences regarding a variety of different criminological outcomes. After that, Hackman model using a two-step procedure was often called "Type-III Tobit, Heckman's two-step Tobit, or Heckman's two-step selection model" among five classifications by Amemiya (1985), since this approach involves estimation of a Probit model for selection, followed by the insertion of a correction factor—the IMR. According to Vella (1998), the Heckman model of using a two-step procedure is more robust to distributional assumptions compared to using MLE introduced by Heckman (1974).

at each utility i (i.e., n has a total of 4,957 water utilities⁶). The variable x_i denotes utility characteristics, such as ownership (local government versus private sector), a primary source (groundwater versus surface water), trade (whether a utility relies on purchased water or its own local sources), geographic location (Northern versus Southern California, and inland versus coastal California), and population size served by utilities (small-populated, medium-populated, and large-populated level). The variable ε_i is normally distributed error term having a mean of 0 and a standard deviation of σ^2 to capture random influences on the relationship between explanatory and outcome variables. The term β represents a parameter to be estimated in Tobit that is censored from below the threshold value τ ($\tau = 0$ in this analysis):

$$y_i = \begin{cases} y_i^*, & \text{if } y_i^* > \tau \\ \tau, & \text{otherwise.} \end{cases} \Leftrightarrow y_i = \begin{cases} y_i^*, & \text{if } y_i^* > 0 \\ 0, & \text{if } y_i^* \leq 0 \end{cases} \quad (2)$$

This Eq. (2) describes the probability that wildfire is greater than 0, the conditional expectation of the latent outcome variable y_i^* is given as:

$$E(y_i^* | y_i^* > 0, x_i) = x_i' \beta + \sigma \frac{\Phi}{\phi} \quad (3)$$

where Φ and ϕ denote the standard normal cumulative distribution function (CDF) on $x_i' \beta / \sigma$ and the corresponding probability density function (PDF), respectively. Accordingly, the parameters β and σ are derived by maximizing the following log-likelihood for each water utility i of the one-limit Tobit model is (Greene 2003):

$$\ln L(\beta, \sigma) = \left[\sum_{y_i > 0} \ln \left(\frac{\phi}{\sigma} \left(\frac{y_i - x_i' \beta}{\sigma} \right) \right) + \sum_{y_i = 0} \ln \left(1 - \Phi \left(\frac{x_i' \beta}{\sigma} \right) \right) \right] \quad (4)$$

This maximum log-likelihood function when the threshold value $\tau = 0$ is the sum of such log-likelihood for each utility obtained by taking the logarithm after combining Eqs. (2) and (3), and by maximizing it through the convergence process by repeatedly operating to get consistent estimates of β and σ .

3.2 Hackman Model

An alternative method, Hackman's two-step procedure model is expressed as:

$$\begin{aligned} 1^{st} \text{ step} : z_i^* &= w_i' \gamma + u_i, \\ z_i &= \begin{cases} 1, & \text{if } z_i^* > 0 \\ 0, & \text{if } z_i^* \leq 0 \end{cases} \\ u_i &\sim N(0, 1) \quad i = 1, 2, \dots, n \end{aligned} \quad (5)$$

$$\begin{aligned} 2^{nd} \text{ step} : y_i^* &= x_i' \beta + \varepsilon_i, \\ y_i &= \begin{cases} y_i^*, & \text{if } z_i^* > 0 \\ 0, & \text{if } z_i^* \leq 0 \end{cases} \\ \varepsilon_i &\sim N(0, \sigma) \quad i = 1, 2, \dots, n \end{aligned} \quad (6)$$

⁶ We only focus on water utilities with wildfires out of 7,500 utilities in California. Hence, we use subsamples composed of 4,957 water utilities throughout California to analyze the impact of wildfires given each period.

where the variable z_i^* denotes an observable variable on whether a wildfire occurs and y_i^* a latent variable on wildfire severity. The variables w_i and x_i mean on the explanatory variables of utility characteristics as stated in the Tobit model. The variables u_i and ε_i are error terms following a bivariate normal distribution with a mean of 0 and a standard deviation of 1 and σ . The correlation between the two error terms to be estimated denotes ρ (i.e., $\text{corr}(u_i, \varepsilon_i) = \rho$), and it implies that the error terms are not independent.

Since the variable y_i^* can be observed when the variable z_i^* has a positive value, when targeting the cases of $z_i^* = 1$ from the Eq. (5), the conditional expectation of y_i^* is yielded as (Greene 2003):

$$\begin{aligned} E(y_i|z_i^*) &= x_i'\beta + E(\varepsilon_i|w_i'\gamma + u_i > 0) \\ &= x_i'\beta + E(\varepsilon_i|u_i > -w_i'\gamma) \\ &= x_i'\beta + \rho\sigma \left[\frac{\varnothing(-w_i'\gamma)}{1 - \Phi(-w_i'\gamma)} \right] \\ &= x_i'\beta + \rho\sigma \left[\frac{\varnothing(w_i'\gamma)}{\Phi(w_i'\gamma)} \right] \Leftrightarrow x_i'\beta + \rho\sigma\lambda_i(-w_i'\gamma) \end{aligned} \quad (7)$$

where $\Phi(\cdot)$ and $\varnothing(\cdot)$ reactively denote CDF and PDF, and $\lambda_i(\cdot) \equiv \varnothing(\cdot)/\Phi(\cdot)$ is a selection parameter called inverse Mills ratio (IMR) or selection hazard. This represents the instantaneous probability that each observation will be excluded from the sample conditional on being at the hazard (Bushway et al. 2007).

Using Eq. (7) the estimation steps are as follows: In the 1st step, using a Probit regression to model the sample selection process in Eq. (5), estimates $\hat{\gamma}$ and calculates $\hat{\lambda}_i(-w_i'\hat{\gamma})$. In the 2nd step, $\hat{\lambda}_i(-w_i'\hat{\gamma})$ is added to a regression, based on Eq. (6) as a regressor along with the other explanatory variables, x_i by adjusting the standard error and t value. This OLS is used to provide the consistent parameter estimates of both β and in ρ . Thus, despite $\rho \neq 0$, estimating wildfire severity with data considering only $z_i^* = 1$, a sample selection bias of $\rho\sigma\lambda_i(-w_i'\gamma)$ occurs; resultantly, the effect of a specific explanatory variable on the outcome variable may be underestimated or overestimated.

4 Results

4.1 Wildfire Severity and Water Utilities

Maps of calculated wildfire severity by the period in the water utilities' service area boundaries in California are presented in Fig. 1. As indicated before, we define the wildfire severity in each water utility service area by dividing the total land acres burned (i.e., wildfire extent) by the total number of events that occurred (i.e., wildfire exposure frequency). The wildfire severity can have a neutralizing effect on "exposure frequency," which can be used to identify vulnerable locations and the degree of vulnerability to wildfires more effectively. Otherwise, displaying raw counts rather than relative values, such as rates per unit (here, acres burned per event), might cause a misleading emphasis on bigger geographic areas. Thus, we decided that it is better to examine the utilities' vulnerability to wildfires by calculating the proportion of the phenomenon measured in severity. Each map illustrates the severity of exposure to wildfire risk within and near the utility's service area boundaries through the severity that an area burned over a 5- to 30-year period. The severity of

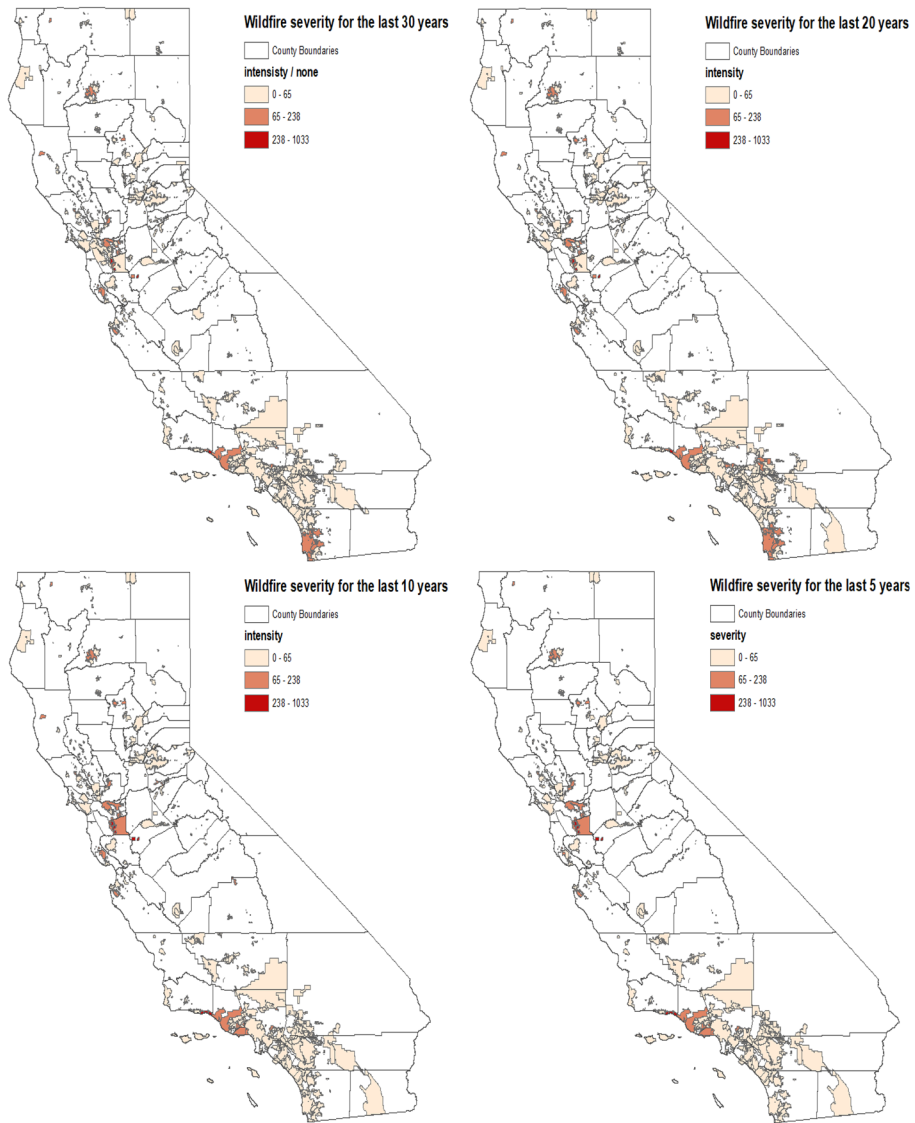


Fig. 1 Wildfire severity (wildfire extent/wildfire exposure frequency) maps by period

wildfires in the last 5 years has increased compared to the last 10, 20, and 30 years. This results from a combination of high temperatures and drought due to the increasing influence of climate change (Bladon et al. 2014). These maps will assist the state's water managers, policymakers, and water utilities in determining the locations that need to implement special management to mitigate wildfire risk.

4.2 Overall Regression Results

Table 3 shows the estimates obtained from the OLS, Tobit, and Heckman models for the association of water utility characteristics and wildfire severity. Statistical signs of all explanatory variables except population size from the OLS model are consistent for the three models, but their statistical coefficients and levels are different.

In summary, all coefficient estimates from the Tobit model, handling ordinary statistical problems in which the data is subject to censoring, are statistically significant and have the expected signs at either the 10% or 1% level. All coefficients of each explanatory variable from the OLS model have smaller absolute terms and are statistically less significant (i.e., government-owned utility, 0.205 to large-populated level utility, 0.305) than those of the Tobit model (i.e., government-owned utility, 1.225 to large-populated level utility, 1.638). Likewise, the Heckman model as a treatment for the sample selection bias also has expected signs and is statistically more significant in the selection equation but less significant in the case of the outcome equation than the OLS model. All coefficients from the Heckman model have relatively greater absolute terms than the OLS model.

All coefficients of the selection equation determining whether or not there are wildfires in the Heckman model are estimated to be smaller than those of the Tobit model at a high significance level—government-owned utility, 0.329 (Tobit 1.225) to large-populated level utility, 0.779 (Tobit 1.638). Yet, the outcome equation implying wildfire severity was instead the opposite. All coefficients are estimated to be larger than those of the Tobit model. None were statistically significant—government-owned utility, 2.192 (Tobit 1.225) to utility size, 3.620 (Tobit 1.638). This indicates that compared to the Tobit model, the Heckman model is valid for estimating the effect of the explanatory variables determining the presence or absence of a wildfire. However, it is not valid for estimating the explanatory variables' impact for determining a wildfire's severity. The following subsections explain the Tobit and Heckman model results, respectively.

4.3 Estimated Results for Tobit Model

First, in reviewing the results from censored Tobit regression estimation in Table 3, the coefficients for characteristics of a water utility such as ownership—local government versus privately owned utilities—reveal a comparison of the vulnerability of wildfires according to ownership. This indicates that while holding other variables in the model constant, water utilities that the government owns compared to those that are privately owned have a greater wildfire severity. This result is statistically significant at the 1% level.

Second, our results show that utilities relying on groundwater as their primary source have a statistically significantly lower rate of wildfire severity (i.e., less vulnerable to wildfires) than those depending on surface water, all else equal. Degradation of water quality from wildfire has the potential to have adverse effects on both surface and groundwater. However, their impacts on surface water are more evident than those on groundwater (Mansilha et al. 2020). This is due to groundwater's hydrogeographic position, which is hard to see and measure. Wildfires can produce significant water quality changes that may impact many of California's reservoirs and associated water infrastructure supplying community drinking water. These impacts are created and accumulated due to pollutants after the wildfire, including black ashes, debris, sediment, and chemicals used to fight the fire. Post-fire runoff provides a

Table 3 Results of OLS model with Tobit and Heckman models

Variables	OLS	Tobit		Heckman
	Coefficient	Coefficient	ME	Coefficient
<i>Outcome equation</i>				
Government	0.205***	1.225***	0.228***	2.192
(Base: Private)	(0.047)	(0.207)	(0.041)	(1.765)
Groundwater	-0.459***	-2.189***	-0.480***	-4.132
(Base: Surface water)	(0.071)	(0.243)	(0.065)	(2.987)
Purchased	-0.256*	-1.529***	-0.214***	-3.057
(Base: Owned)	(0.105)	(0.337)	(0.037)	(2.432)
Southern CA	0.394***	2.106***	0.447***	3.866
(Base: Northern CA)	(0.052)	(0.196)	(0.049)	(2.970)
Coastal	0.266***	1.109***	0.201***	2.124
(Base: Inland)	(0.042)	(0.189)	(0.036)	(1.298)
Size-Medium	-0.065	0.635*	0.119*	1.239
(Base: Small)	(0.070)	(0.255)	(0.051)	(1.731)
Size-Large	0.305	1.638***	0.362**	3.620
(Base: Small)	(0.177)	(0.419)	(0.115)	(3.482)
<i>Selection equation</i>				
Government				0.329***
(Base: Private)				(0.051)
Groundwater				-0.576***
(Base: Surface water)				(0.063)
Purchased				-0.449***
(Base: Owned)				(0.094)
Southern CA				0.570***
(Base: Northern CA)				(0.053)
Coastal				0.243***
(Base: Inland)				(0.047)
Size-Medium				0.324***
(Base: Small)				(0.069)
Size-Large				0.779***
(Base: Small)				(0.156)
Observations	4,478	4,478		4,478
R2	0.054	-		-
Pseudo R2	-	0.052		-
σ	-	18.439***		9.633
ρ	-	-		1.000
λ_i	-	-		9.633

The coefficient of IMR is reported as λ_i Robust standard errors in parentheses

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

route to carry these pollutants to surface water, which may lead to the substantial water quality of water available impacts. Loaded pollutants along the route can fill drainage channels, reduce reservoir holding capacities, and add contaminants to the existing water infrastructure (Moody et al. 2013; Abraham et al. 2017; Murphy et al. 2020).

Third, provided the other variables are held constant, we find utilities mainly dependent on purchased water are less vulnerable to wildfires than utilities dependent on their local sources. These results are statistically significant at the 1% level. Fourth, utilities in Southern California have a higher severity of wildfires relative to those in Northern California. Predictably, the result is due to different weather conditions that result in less rain and hot and dry weather in Southern California. Higher temperatures and drier conditions, such as heatwaves, often exacerbate wildfires. Usually, in response to wildfires that threaten life, property, and high-value resources, fire retardants are used to slow the spread of wildfires, thus allowing time for firefighters to respond to the incident. However, it is difficult to expect these retardants to prevent the spread of parched land due to high temperatures and severe drought. Consequently, it is more likely that Southern California, with landscapes that are largely semi-arid with Mediterranean ecosystems and significant chaparral shrublands,⁷ is a more wildfire-prone environment than Northern California (Barro and Conard 1991; Syphard et al. 2018).

Fifth, utilities in coastal California have a higher severity of wildfires than those in inland California. The more intense wildfire outbreaks in coastal California are due to climate trends that make it easy for wildfires to spark. That is because offshore, downslope winds with low humidity blow from inland areas toward the coast. Burnt trees are like charcoal, so the embers last for a long time, and the wind spreads sparks from those embers into neighboring areas. The lower the humidity and the stronger the wind, the faster this diffusion and the greater the distance the fire sparks will travel (Goss et al. 2020). Large-scale wildfires in coastal California have increased by about 10% per decade since 1984 due to climate trends (Zhongming et al. 2020). These trends push existing fires into urban communities. This phenomenon causes more damage when wildfires occur near densely populated areas (e.g., Bay Area, Los Angeles, San Bernardino, and San Diego counties).

In terms of population size served by a utility, results show how much more (or less) likely utilities serving a larger population are vulnerable to wildfires than utilities serving a smaller population. The increase in population size causes an incremental increase in wildfire severity from each utility. Specifically, the increase of population size (i.e., serving less than 100,000 people) from a utility supporting a medium-population, and a utility supporting a large population (i.e., serving more than 100,000 people) result in an incrementally growing severity, compared to a utility supporting a small-population (i.e., serving less than 3,300 people). Such results are all statistically significant at either the 5% or 1% level. Utilities that provide services to a larger population are often large and government-owned. These larger-population areas may also increase the number of wildfires caused by inadvertent human actions.

Lastly, σ value shows that it is reasonable to choose the Tobit model to handle the case in which the severity of the wildfire is less than or equal to zero. The value was found to be statistically significant at 18.439 ($p < 0.001$), indicating that the Tobit model is a valid

⁷ Chaparral vegetation is widely distributed and dominant in a region of Mediterranean-type climate, which has hot dry summers and mild wet winters because it has characteristics to be well adapted to extreme dry-hot conditions and interannual precipitation variability (Molinari et al. 2018).

method. The Tobit model indicates the effects of all explanatory variables on wildfire severity and remains statistically significant.

The coefficients from the OLS model show the marginal effect in which average changes of the outcome variable are given with explanatory variable changes. However, the coefficients from the Tobit model have average changes in the outcome variable and the probability that the changes will be observed. Thus, to get the marginal effect of an explanatory variable on the observed outcome variable, it is necessary to multiply the coefficient by the probability that the observed variable is greater than zero (Amore and Murtinu 2021). The value of the marginal effect is also presented alongside the coefficient in Table 3.

4.4 Estimated Results of Heckman Model

We estimated two equations in Heckman's two-step procedure, with the estimations from the Probit model as a selection equation at the first step and the OLS model as an outcome equation at the second step. The overall results of the Heckman model in Table 3 show no difference in the direction of the effect explanatory variables have on the outcome variable between the first Probit selection model and the second linear probability model.

The selection equation determines whether or not there are wildfires and its coefficients' signs on all explanatory variables are consistent with those from Tobit's outcomes, but they are estimated smaller than the Tobit model and are more statistically significant. Thus, the effects of heterogeneous characteristics of each water utility on the severity of wildfires would be the same as what the Tobit model has. The outcome equation describes the effect of the explanatory variables on the wildfire's severity. The statistical signs from the coefficients on all explanatory variables are still equivalent to the Tobit model, yet they were not statistically significant. This means that water utilities' characteristics have a significant impact on whether a wildfire occurs, but not on its severity.

Notice that the coefficient of Lambda (λ_i), denoting IMR, is insignificant with t -value = 1.40 (at $p > 0.1$). This suggests for our analysis that selection bias is not a significant issue. Given the meaning of λ_i as a self-selectivity correction factor, this result shows that selection bias is not a significant issue for our analysis, indicating that the Tobit model would give us a better explanation in the wildfire severity than the two-step Heckman model. Furthermore, the Tobit model is a less-complicated statistical approach. For another parameter ρ , when it is positive, this indicates the unobservable factors that cause a wildfire's severity are positively correlated with one another; conversely, negative ρ indicates the unobservable factors are negatively correlated. The error terms for ρ in Table 3 show the unobserved factors that cause wildfire severity are positively correlated.

4.5 Estimated Results for Different Time Periods

Tobit model estimates for wildfire severity presented in Table 4 show changes in wildfire exposure over time. These variations reflect the base period vulnerability of water utilities to wildfire. Since all estimates for each period have the same statistical signs, the direction of the effects—either increase or decrease—from the characteristics of water utility on the severity of the wildfires will be the same. These results are the same as those described in the sections above. These results are the same in statistical significance at either the 5% or 1% level, except for features about population size. However,

Table 4 Estimated results of Tobit model for wildfire severity by period

Variables	Last 30-years (1991–2020)		Last 20-years (2001–2020)		Last 10-years (2011–2020)		Last 5-years (2016–2020)	
	Coefficient	M.E	Coefficient	M.E	Coefficient	M.E	Coefficient	M.E
Government	1.225***	0.228***	1.173***	0.195***	1.280***	0.166***	1.266***	0.130***
(Base: private)	(0.207)	(0.041)	(0.230)	(0.041)	(0.268)	(0.037)	(0.320)	(0.035)
Ground-water	-2.189***	-0.480***	-2.302***	-0.455***	-2.027***	-0.303***	-2.320***	-0.282***
(Base: surface water)	(0.243)	(0.065)	(0.268)	(0.065)	(0.317)	(0.057)	(0.372)	(0.055)
Purchased	-1.529***	-0.214***	-1.434***	-0.184***	-1.327**	-0.134***	-1.545**	-0.122***
(Base: owned)	(0.337)	(0.037)	(0.372)	(0.038)	(0.434)	(0.036)	(0.503)	(0.032)
Southern CA	2.106***	0.447***	2.049***	0.386***	0.747**	0.097**	0.560	0.057
(Base: Northern CA)	(0.196)	(0.049)	(0.219)	(0.048)	(0.253)	(0.035)	(0.309)	(0.033)
Coastal	1.109***	0.201***	1.318***	0.216***	1.411***	0.180***	1.989***	0.206***
(Base: Inland)	(0.189)	(0.036)	(0.208)	(0.036)	(0.240)	(0.032)	(0.283)	(0.032)
Size-Medium	0.635*	0.119*	0.742**	0.126*	1.168***	0.161**	1.327***	0.147**
(Base: Small)	(0.255)	(0.051)	(0.283)	(0.052)	(0.328)	(0.052)	(0.392)	(0.050)
Size-Large	1.638***	0.362**	1.886***	0.384**	2.371***	0.398**	3.167***	0.469***
(Base: Small)	(0.419)	(0.115)	(0.478)	(0.124)	(0.548)	(0.124)	(0.615)	(0.131)
Observations	4,478		4,478		4,478		4,478	
Pseudo R^2	0.052		0.051		0.038		0.042	
σ	18.439***		21.323***		24.972***		30.880***	

Robust Standard errors in parentheses

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

it can be seen that the degree of the effect differs according to the period, as the coefficients are different by each period.

Although the point estimates of the effects vary according to the period, it is clear that the overall effects are larger in a short period (the largest point estimates are from the 5 years). The longer the period, the less fluctuation degree in severity estimated despite the greater frequency of accumulated wildfires, implying that the area burned by the wildfire is not that large. In other words, since the number of occurrences of small-scale wildfires accounts for a large proportion, the increase in total area burned is not large compared to a short period. This indicates that large-scale wildfires have mainly occurred in a relatively short period (the last 5 years). A large body of research suggests that burning areas have increased within the last 5 years, and the top 20 largest California

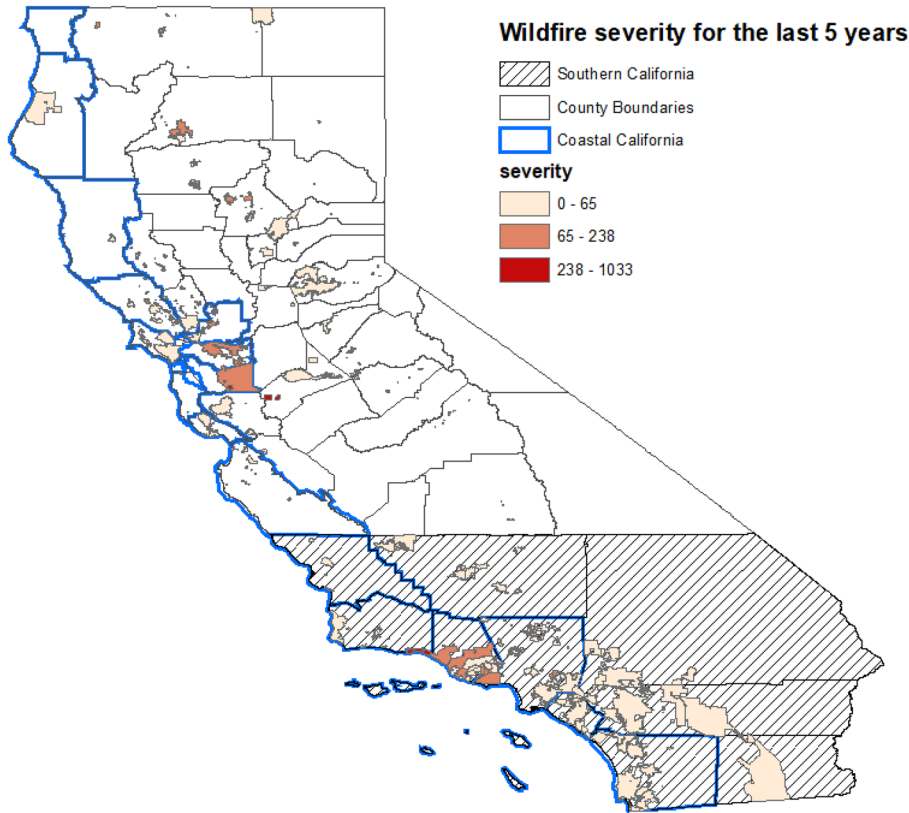


Fig. 2 Wildfire severity map for coastal and southern California

wildfires since 1932 mainly include the recent 5 years (Hoover and Hanson 2021; Keeley and Syphard 2021; California Department of Forestry and Fire 2021). Figure 2 focuses on the last wildfire severity in the last 5 years by geographic location in California. As shown, southern and coastal California utilities are more vulnerable to wildfires than those in northern or inland California.

5 Conclusions

Growing wildfire size and frequency in California concerns the water utilities and imposes strains on water utilities regarding increased costs for the existing water infrastructures' management and operations. Wildfires impact both post-wildfire water quality and water infrastructure management and operation. Accordingly, identifying the water utilities at a higher risk of wildfires is necessary to develop plans and strategies for effective water management and operation. Our study contributes to the literature by measuring the vulnerability of each utility to wildfires based on the exposure frequency and the extent of area burned by wildfires that occurred in each water utility service area. Our quantitative models take into account the nature of the censoring and selection biases on wildfire data

and show an association between water utility characteristics and the level of vulnerability to wildfire risks.

To summarize, we find that wildfire vulnerability is higher in government-owned utilities than in private ones, utilities primarily relying on surface water than groundwater, utilities using local-sourced water than purchased water, utilities located in Southern California than those in Northern California, utilities located in coastal California than those in inland California, and utilities serving a larger population. These findings provide useful information on water utilities that are likely to be more vulnerable to wildfires based on their characteristics, which may require more caution in their management and operation. After a wildfire, water utilities tend to experience increased infrastructural damages (e.g., water treatment, replacement or repair of the broken water system, water storage or distribution system) and operational challenges (e.g., emergency operation/response plans, disinfection, and boil water orders). The introduction of funding mechanisms to support water utilities with wildfire-vulnerable characteristics to respond fiscally and flexibly to wildfires. Such mechanisms might include insurance, grants, subsidies, and other state and federal aid programs; resultantly, water agencies may be able to recover some of the costs of water management and operation post-disaster assistance through the mechanisms. Given that adaptation and mitigation of the risk of wildfire are inextricably linked to the provision of a safe water supply, it is essential to constitute partnerships across the diverse forest- and water entities at all levels: federal, state, local municipalities, communities, and nongovernment.

Acknowledgements Partial support from the W4190 Multistate NIFA-USDA-funded project, Management, and Policy Challenges in a Water-Scarce World, and the Giannini Foundation of Agricultural Economics, is greatly appreciated.

Authors Contributions All authors contributed to the study's conception and design. JL and MN performed material preparation, data collection, and analysis. JL wrote the first draft of the manuscript and all authors commented on previous versions. All authors read and approved the final manuscript.

Funding Partial funding was provided for this project from the W4190 Multistate NIFA-USDA-funded project, Management and Policy Challenges in a Water-Scarce World, and the Giannini Foundation of Agricultural Economics.

Availability of data and materials Data associated with this research will be available from the corresponding author upon request.

Declarations

Ethical Approval This study does not involve human subjects and/or animals.

Consent to Participate This study does not involve human subjects.

Consent to Publish This study does not involve human subjects.

Conflict of Interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abraham J, Dowling K, Florentine S (2017) Risk of post-fire metal mobilization into surface water resources: A review. *Sci Total Environ* 599:1740–1755
- Amemiya Y (1985) Instrumental variable estimator for the non-linear errors-in-variables model. *Journal of Econometrics* 28(3):273–289
- Amore MD, Murtinu S (2021) Tobit models in strategy research: Critical issues and applications. *Glob Strateg J* 11(3):331–355
- Barro SC, Conard SG (1991) Fire effects on California chaparral systems: an overview. *Environ Int* 17(2–3):135–149
- Batellis SC, Nalbantis I (2014) Potential effects of forest fires on streamflow in the Enipeas river basin, Thessaly. *Greece Environmental Processes* 1(1):73–85
- Bladon KD, Emelko MB, Silins U, Stone M (2014). *Env Sci Tech*. 48(16), 8936–8943. <https://doi.org/10.1021/es500130g>
- Bladon KD (2018) Rethinking wildfires and forest watersheds. *Science* 359(6379):1001–1002
- Bushway S, Johnson BD, Slocum LA (2007) Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology. *J Quant Criminol* 23(2):151–178
- California Department of Forestry and Fire (CAL FIRE) (2021) Historical Wildfire Activity Statistics. Available from <https://www.fire.ca.gov/stats-events/>
- DeBano LF, Ritsema CJ, Dekker LW (2003) The role of fire and soil heating on water repellency. *Soil water repellency: Occurrence, consequences, and amelioration* pp 193–202
- DeBano LF (2000) The role of fire and soil heating on water repellency in wildland environments: a review. *J Hydrol* 231:195–206
- Goss M, Swain DL, Abatzoglou JT, Sarhadi A, Kolden CA, Williams AP, Diffenbaugh NS (2020) Climate change is increasing the likelihood of extreme autumn wildfire conditions across California. *Environ Res Lett* 15(9):094016
- Greene WH (2003) *Econometric analysis*: Pearson Education India
- Hallema DW, Robinne FN, Bladon KD (2018) Reframing the challenge of global wildfire threats to water supplies. *Earth's Future* 6(6):772–776
- Heckman J (1974) Shadow prices, market wages, and labor supply. *Econometrica* 679–694
- Heckman J (1976) The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. *Ann Econ Soc Meas* 5(4):475–492. NBER
- Hohner AK, Cawley K, Oropeza J, Summers RS, Rosario-Ortiz FL (2016) Drinking water treatment response following a Colorado wildfire. *Water Res* 105:187–198
- Hoover K, Hanson LA (2021) Wildfire statistics. Congressional Research Service p 2
- Ice GG, Neary DG, Adams PW (2004) Effects of wildfire on soils and watershed processes. *J Forest* 102(6):16–20
- Keeley JE, Syphard AD (2021) Large California wildfires: 2020 fires in historical context. *Fire Ecology* 17(1):1–11
- Kim CG, Shin K, Joo KY, Lee KS, Shin SS, Choung Y (2008) Effects of soil conservation measures in a partially vegetated area after forest fires. *Sci Total Environ* 399(1–3):158–164
- Liu JC, Mickley LJ, Sulprizio MP, Dominici F, Yue X, Ebisu K, Anderson GB, Khan RF, Bravo MA, Bell ML (2016) Particulate air pollution from wildfires in the Western US under climate change. *Clim Change* 138(3–4):655–666
- MaCurdy TE, Pencavel JH (1986) Testing between competing models of wage and employment determination in unionized markets. *J Polit Econ* 94(3, Part 2):S3–S39
- Mansilha C, Duarte CG, Melo A, Ribeiro J, Flores D, Marques JE (2019) Impact of wildfire on water quality in Caramulo Mountain ridge (Central Portugal). *Sustainable Water Resources Management* 5(1):319–331
- Mansilha C, Melo A, Martins ZE, Ferreira IM, Pereira AM, Espinha Marques J (2020) Wildfire effects on groundwater quality from springs connected to small public supply systems in a peri-urban forest area (Braga Region, NW Portugal). *Water* 12(4):1146
- Molinari NA, Underwood EC, Kim JB, Safford HD (2018) Climate change trends for chaparral. *Valuing Chaparral*. Springer, Cham, pp 385–409
- Moody JA, Shakesby RA, Robichaud PR, Cannon SH, Martin DA (2013) Current research issues related to post-wildfire runoff and erosion processes. *Earth Sci Rev* 122:10–37
- Murphy SF, McCleskey RB, Martin DA, Holloway JM, Writer JH (2020) Wildfire-driven changes in hydrology mobilize arsenic and metals from legacy mine waste. *Sci Total Environ* 743:140635

- Neary DG, Ryan KC, DeBano LF (2005) Wildland fire in ecosystems: effects of fire on soils and water. Gen. Tech. Rep. RMRS-GTR-42-vol. 4. Ogden, UT: US Department of Agriculture, Forest Service, Rocky Mountain Research Station 250:42
- National Interagency Fire Center (NIFC) (2021) Fire Information: Wildfires and Acres. Available from <https://www.nifc.gov/fire-information/statistics/wildfires>
- Prior T, Eriksen C (2013) Wildfire preparedness, community cohesion and social–ecological systems. *Glob Environ Chang* 23(6):1575–1586
- Robinne FN, Bladon KD, Silins U, Emelko MB, Flannigan MD, Parisien MA, Wang X, Kienzle SW, Dupont DP (2019) A Regional-Scale Index for Assessing the Exposure of Drinking-Water Sources to Wildfires. *Forests* 10(5):384
- Robinne FN, Miller C, Parisien MA, Emelko MB, Bladon KD, Silins U, Flannigan M (2016) A global index for mapping the exposure of water resources to wildfire. *Forests* 7(1):22
- Schulze SS, Fischer EC (2020) Prediction of water distribution system contamination based on wildfire burn severity in wildland urban interface communities. *ACS ES&T Water* 1(2):291–299
- Seibert J, McDonnell JJ, Woodsmith RD (2010) Effects of wildfire on catchment runoff response: a modeling approach to detect changes in snow-dominated forested catchments. *Hydrol Res* 41(5):378–390
- Sham CH, Tuccillo ME, Rooke J (2013) Effects of wildfire on drinking water utilities and best practices for wildfire risk reduction and mitigation: Water Research Foundation
- Syphard AD, Brennan TJ, Keeley JE (2018) Chaparral landscape conversion in southern California. *Valuing chaparral*. Springer, Cham, pp 323–346
- Fire and Resource Assessment Program (FRAP) (2021) Fire Perimeters. Available from <https://frap.fire.ca.gov/frap-projects/fire-perimeters/>
- Tobin J (1958) Estimation of relationships for limited dependent variables. *Econometrica: journal of the Econometric Society*:24–36
- Vella F (1998) Estimating models with sample selection bias: a survey. *J Hum Resour* 127–169
- Zhongming Z, Linong L, Wangqiang Z, Wei L (2020) Wildfire on the rise since 1984 in Northern California's coastal ranges

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.