Moral Hazard, Wildfires, and the Economic Incidence of Natural Disasters

Patrick Baylis Judson Boomhower*

March 2018

Increased wildfire risk is one of the most salient impacts of climate change in North America. As is the case for many other impacts of climate change, adaptive responses to worsening wildfires include large government investments. These growing public expenditures raise classic public economics questions about moral hazard, distributional impacts, and allocative efficiency. We consider these questions in the context of wildland firefighting expenditures in the United States, which are now several billion dollars per year and are incurred almost entirely by the federal government. We assemble administrative firefighting expenditure data from five federal and state agencies, yielding the most comprehensive database of firefighting costs in existence. We merge this to parcel-level data on the universe of western U.S. homes. We make two main contributions. First, we measure the share of firefighting expenditures that are dedicated to protecting private homes. To do this, we take advantage of natural variation in ignition locations to measure the causal impact of private home presence and density on firefighting costs. Next, we use our data and estimates to calculate parcel-level implicit transfers via firefighting for the entire western U.S. We find that firefighting expenditures are overwhelmingly driven by efforts to protect homes. Costs are strongly non-linear in the number of homes threatened, meaning housing density strongly affects per-home protection costs. Wildland firefighting represents a large transfer of federal revenues to landowners in high-risk, low-density places. For the highest-cost categories of homes, the expected present value of firefighting costs exceeds 10% of the transaction value. We evaluate the moral hazard implications of this implicit subsidy through a back-of-the-envelope exercise using price elasticities for new residential construction.

*(Baylis) Vancouver School of Economics, University of British Columbia; pbaylis@mail.ubc.ca. (Boomhower) University of California, San Diego; jboomhower@ucsd.edu. The authors gratefully acknowledge research support from the Stanford Institute for Economic Policy Research (Boomhower), the Stanford Center on Food Security and the Environment (Baylis), and the Giannini Foundation. We are grateful to seminar participants at Arizona State University, Stanford University, UC Berkeley, University of British Columbia, University of Ottawa, the UC Santa Barbara Occasional Workshop, the Heartland Workshop, and the AERE Summer Conference.

1 Introduction

The burden of natural disasters has increased over the past several decades and is predicted to grow larger as the climate continues to warm. Floods, cyclones, landslides, heat waves, droughts, and wildfires are all predicted to increase in frequency and severity.[1] As with the other impacts of climate change, the costs of these increased natural hazards will depend on how societies respond and adapt. Many important adaptive responses are likely to occur through government investments in public goods like infrastructure, national security, scientific research, public health, emergency response, and other areas. These large public investments may lessen the costs of climate change, but they also raise basic public economics questions about moral hazard, distributional impacts, and allocative efficiency.

We consider these questions in the context of wildland fires in the United States. Increased frequency and severity of wildfires is one of the most salient impacts of climate change in North America. The primary policy response has been a dramatic increase in public spending on wildland firefighting. Over the past 30 years, annual costs have risen from $240 million to over $2 billion in for the federal government alone, and those costs continue to increase rapidly. Wildland fires now consume more than 50% of the U.S. Forest Service's annual budget.[2] Every summer and fall, tens of thousands of men and women and millions of dollars worth of equipment are continuously dispatched throughout the western United States by a complicated combination of federal and state agencies. Almost all of the costs of these efforts are ultimately borne by the federal government, either through direct expenditures by federal agencies or through reimbursements to states and other entities through FEMA grants.

Wildfires are unusual among natural hazards in that it is feasible to prevent private property damage while an incident is ongoing through large investments of manpower

[1] For a review of natural disasters and climate change, see IPCC, 2012: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation. [Field, C.B., V. Barros, T.F. Stocker, D. Qin, D.J. Dokken, K.L. Ebi, M.D. Mastrandrea, K.J. Mach, G.-K. Plattner, S.K. Allen, M. Tignor, and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, UK, and New York, NY, USA.

[2] National Interagency Fire Center. "Federal Firefighting Costs (Suppression Only)". https://www.nifc.gov; and USDA Forest Service 2015. "The Rising Cost of Wildfire Operations: Effects on the Forest Service's Non-Fire Work." .

and equipment. Unlike cyclones or earthquakes, for example, wildfires can often be "stopped in their tracks" to protect homes and other valuable assets. While tragic losses of life and property do occur and (appropriately) receive great attention, a large share, if not the majority, of the costs imposed on society by wildfires come in the form of extremely costly efforts to prevent property damage. At least in the United States, those efforts benefit homeowners in high-risk areas but are paid for out of general government revenues and are thus borne equally by all taxpayers.

In this paper, we provide the first estimates of which we are aware of the implicit transfer to homeowners due to fire protection at the individual parcel level for homes throughout the western United States. We combine parcel-level data on the universe of single family homes in the West with administrative data on historical firefighting expenditures to estimate federal government expenditures dedicated to protecting each home from wildfires. We assembled the firefighting cost data from administrative records of five different federal and state agencies through multiple Freedom of Information Act and public records requests, yielding the most comprehensive dataset on wildland firefighting expenditures in existence. We first take advantage of randomness in ignition locations to measure the share of firefighting expenditures that are solely devoted to protecting homes. We then use these estimates to construct expected protection costs across groups of similar-risk homes. Finally, we apply simple spatial equilibrium reasoning to quantify potential distortions in new residential construction due to moral hazard, and to explore a policy counterfactual where developers pay a fee equal to the expected net present value of fire protection costs at the time of initial home construction.

We find that residential development dramatically increases fire suppression costs. Efforts to protect private homes account for over two-thirds of all firefighting expenditures. Surprisingly, among fires that threaten homes, the number or total value of homes threatened has little effect on firefighting costs. This means that development density is an important determinant of per-home protection cost. Overall, firefighting represents a remarkably large transfer to a few landowners in high-risk, low-density places. In our highest-risk category, the net present value (NPV) of fire protection exceeds 10% of total property value. Because the supply of new homes in these areas is relatively elastic (Saiz, 2010), this large implicit subsidy suggests substantial distortions in new home construction. We also consider the distributional effects of these large transfers.

Our paper contributes to a small literature about natural hazards and location choice. Kousky et al. (2006) and Boustan et al. (2012) examine adaption to hurricanes and floods. A related paper is Kousky and Olmstead (2012), which shows that changes over time in federal firefighting policy affected the number of homes built near public lands. We make several novel contributions. By introducing data on firefighting costs, we are able to quantify the implicit firefighting subsidy. To our knowledge, we are the first to measure this subsidy and to calculate the optimal "fire protection fee" for each home. We also demonstrate a strongly non-linear response of firefighting costs to the number of threatened homes, meaning that housing density affects protection costs. Finally, by using parcel-level data on 18 million western homes, we are able to be geographically precise about risks and costs. This specificity is important since fire and other disaster risks can vary substantially over small distances.

2 Background

2.1 Increasing Costs of Wildland Firefighting

The rapid increase in firefighting cost over the past several decades has been attributed to three primary factors: the lengthening of the fire season as a result of climate change, the buildup of increasingly dangerous fuel loads, and increased human habitation in fire-prone areas. Changes in climate can increase wildland fire activity by either increasing the amount of fuel available for fires or by drying out existing fuel, rendering it more flammable. Prior work estimates that climate change is responsible for an additional 4.2 million acres in burned area between 1984 and 2015, accounting for nearly half of the increase in acres burned (Abatzoglou and Williams, 2016).

However, the increase in available fuels has not been solely driven by climate change. Land use change in beginning in the 19th century and an increase in fire suppression activity in the 20th century have both altered the type and the extent of fuel availability in the Western United States (Stephens et al., 2016). Although the precise impacts of these changing fuels on the cost of fires is the subject of continuing scientific investigation, the majority view is that the suppression of most fires has led to an increase in the severity in the fires that do escape suppression (Kousky and Olmstead, 2012).

Finally, a rapid increase in the number of homes at possible risk from wildland fires has also contributed to the rising costs of wildland firefighting. Between 1990 and 2000, 8 million homes were added to the Wildland-Urban Interface, or WUI (Hammer et al., 2009). Foresters and planners project that new homes will continue to be built at a rapid pace in these high-risk areas. Gude et al. (2008) projects that huge areas of the WUI that are currently totally undeveloped will be converted to residential housing over the next two to three decades.

2.2 The Cost of Protecting Homes During Wildfires

Wildland firefighting efforts have multiple objectives, among them safeguarding human lives, protecting publicly-owned natural resources and endangered species, and preventing damage to private property. The incidence and housing market impacts of wildland firefighting depend on the share of expenditures that are devoted to private property protection. What additional firefighting expenses result from locating homes in the path of wildfires? Our paper builds on previous studies of firefighting expenditures in both forestry and resource economics. Previous case studies and expert introspection indicate that the presence of homes greatly increases firefighting costs, as it requires significantly more manpower and equipment (e.g., air support, bulldozers) to stop a fire in place before it reaches homes, as opposed to letting the fire burn out naturally at a road or ridge or other natural fire barrier (of Inspector General Western Region, 2006). Forest Service personnel, report that, heuristically, between 50 and 95 percent of federal firefighting costs is due to efforts to prevent damage to homes (of Inspector General Western Region, 2006). Case studies of small samples of fires have found econometric results in line with these estimates (Gebert et al., 2007; Liang et al., 2008). Champ et al. (2009); Kousky and Olmstead (2012) consider the relationship between wildfire risk and housing markets.

In this paper, we validate these case study findings for the entire Western United States and extend the analysis to quantify the implicit subsidy to each homeowner. As described above, the dataset we compile covers all 11 western states and combines administrative expenditure records gathered through public records requests to multiple federal and state agencies, and to the best of our knowledges serves as the most comprehensive record of its kind by a wide margin. By matching this dataset to a proprietary record of parcel-level homes deed and tax records for the universe of residences in our sample, we are further able to more accurately measure the locations of individual homes with respect to the location of the fire.[3] These more detailed data allow us to precisely estimate the parcel-level benefits of this firefighting effort and to document striking non-linearities in fire costs as a function of the number and proximity of threatened homes. Finally, in contrast to the existing literature, which relies on regression adjustment to account for possible confounding variables, our empirical strategy identifies the causal impact of homes accounts carefully for both observed and unobserved variation in home and fire locations.

To fix ideas, we next turn to a conceptual framework.

3 Conceptual Framework: Defensive Expenditures and Housing Demand

This section develops an economic framework to guide our empirical analysis. We adapt a simple spatial equilibrium model to demonstrate how subsidizing natural hazard protection affects demand for housing across locations. This analysis clarifies which costs are borne by households, and gives economic context for the parameters that we measure in the following sections.

Our starting point is the well-known Rosen (1979)-Roback (1982) model. We add a spatially differentiated natural hazard, and consider how location decisions respond to a public guarantee of freely-provided protection. The principal alternative policy we have in mind is a policy that would require homeowners to internalize the expected costs of protecting their home – for example, an up-front fee at the time of construction equal to the expected net present value of protection costs.

Rosen-Roback-type models are a standard conceptual tool in urban, labor, and en-

[3]Previous studies rely on publicly-available housing counts data at the Census Block level, which is less precise in our sample of interest because in rural areas Census blocks are often many square kilometers or larger.

vironmental economics.[4] This type of model has the advantages of simplicity and transparency. At the same time, it sacrifices some realism and detail relative to richer models. We believe the more straightforward approach best serves our primary goal in this section, which is to illustrate how government spending on disaster response affects household decisions. We state our assumptions clearly as we go, and at the end of this section we discuss how relaxing assumptions in a more complicated model might affect the model predictions.

3.1 Setup

There are N identical, perfectly mobile households that choose from $I < N$ possible locations. Population in location i is denoted by N_i. Each household supplies one unit of labor inelastically. Labor productivity in each location is fixed, giving location-specific wages w_i. Amenities $A_i(N_i)$ are decreasing in local population. This reflects the amenity value of "open space" or "privacy", in keeping with stylized facts about development in the wildland-urban interface areas. Housing prices r_i are determined by the location-specific housing supply curve. We consider the form of the housing supply function in more detail below. We make the common simplifying assumption that there is a reservation location that offers an exogenous utility level $\bar{u}$ (that is, a location with perfectly elastic housing supply and uncongestible amenities).

We introduce a location-specific natural disaster probability ϕ_i. When a disaster occurs, households may incur rebuilding costs. Defensive expenditures f made in response to the disaster can reduce expected rebuilding costs, which we denote $H_i(f)$. Defensive expenditures (e.g., firefighting) are supplied by the central government. We make the following assumptions about f and $H(f)$, which are consistent with our data and stylized facts about natural hazard response.

1. $H_i'(f) < 0$ and $H_i''(f) > 0$
2. The benefits of defensive expenditures are non-rival within a location.

[4] Gyourko et al. (1999) and Kahn and Walsh (2015) review work related to the environment. A related paper to ours is a theoretical paper by Kousky et al. (2006), which develops a model of government protection and private investment primarily focused on flood risk. Albouy et al. (2016) examines the amenity cost of temperature changes due to climate change using a similar framework.

3. Within a location, homes are geographically homogeneous so that $H_i(f)$ is constant across homes.

Assumption 1 ensures that defensive expenditures reduce expected damages, and do so with diminishing returns. Assumption 2 matches the facts of our empirical setting, where firefighting efforts are focused on protecting entire communities. Assumption 3 abstracts away from local heterogeneity to focus the analysis on community-level interactions.

In the event of a disaster, the government chooses the optimal level of defensive expenditure given local population. This value $f^*(N)$ minimizes the sum of defensive expenditures and total expected rebuilding costs, $f + N_i H_i(f)$.[5] The following section considers how the financing of defensive expenditures affects the population distribution and welfare.

3.2 Incidence and Moral Hazard

If households are required to reimburse the central government for their proportional share of defensive expenditures after a disaster, household utility in place i in the event of a disaster is, $w_i + A_i(N_i) - r_i(N_i) - \frac{f_i}{N_i} - H_i(f_i)$. In the no-disaster state of the world, household utility is $w_i + A_i(N_i) - r_i(N_i)$. Assuming risk-averse households and perfectly competitive insurance markets, households will purchase full insurance to cover their costs in the event of a disaster. Premiums in each place will equal expected losses, $\pi_i = \phi_i[\frac{f_i^*}{N_i} + H_i(f_i^*)]$. Holding population constant, disaster costs are larger in areas with higher disaster risk. At the same time, per-capita expected damages are decreasing in local population. These scale economies are due to the non-rival nature of defensive expenditures.

We normalize the disaster risk in the reservation location to zero. The new spatial equilibrium condition is,

$$\bar{U} = w_i + A_i(N_i) - r_i(N_i) - \pi_i(\phi_i, N_i)$$

To understand how disaster costs affect location decisions, consider expected util-

[5]This result mimics the principle of "least cost plus net value change" in the natural resource economics literature on fire suppression.

ity relative to the reservation location. Let $z_i(N_i)$ and $z_0(N_0)$ represent wages plus amenities minus housing prices in location i and the reservation location. Equilibrium requires that,

$$z_i(N_i) - z_0(N_0) = \pi_i(\phi_i, N_i)$$

Populations are distributed such that the difference in wages, amenities, and home prices between each risky location and the riskless location exactly offsets expected disaster costs in the risky location. This equilibrium will depend on the relative disaster risks across areas (ϕ), the slopes of the housing supply and amenity functions, and the shape of the damages function, $H(f)$.

Now consider an alternative policy where the central government makes defensive expenditures without reimbursement. The government continues to use the same dispatch rule, providing the optimal level of defensive expenditures, $f^*(N_i)$, in the event of a disaster. When the government provides natural disaster defense for free, the expected disaster costs borne by households (and thus the household's insurance premium) include only expected rebuilding costs, $\rho_i H(f_i^*(N_i))$. The compensating differential in terms of real wages and amenities is smaller. For budget balance, all households regardless of location are assumed to pay a flat tax $\tau = \frac{1}{N}\sum_{i=0}^{I} f^*(N_i)$ that just covers total defensive expenditures.

Figure 1 shows these two alternative policies for a single risky location. The horizontal axis shows local population. The downward-sloping black line represents per-capita wages and amenities. Housing prices are given by the housing supply curve labeled S, which is elastic up to a point beyond which geography or regulation constrains further development. When households pay for defensive expenditures, the per-capita benefit of this location follows the solid gray curve, which reflects wages and amenities minus expected rebuilding costs and per capita defensive expenditures. This line approaches $w_i + A(N_i)$ as local population increases, reflecting economies of scale in defensive expenditures.[6] The equilibrium population under this policy is $\hat{N}$, the population level that sets per-capita expected utility equal to $\bar{u}$.

When households do not reimburse the government for defensive expenditures, per-capita private benefit shifts up to the dashed gray line, which includes expected

[6]The online appendix shows that per capita disaster-related costs decrease in population while total disaster-related costs are increasing. The proof follows from the envelope theorem, since $f^*(N)$ is chosen optimally to minimize disaster-related costs.

rebuilding costs but not expected defensive expenditures. Households no longer consider the full costs of natural disasters in their location decisions. There is still a compensating differential for risky areas, reflecting private insurance premiums equal to expected rebuilding costs. However, defensive expenditures do not enter household decisions. Under this policy, local population increases to N'.

Per-capita welfare in every location is lower under the policy that provides defensive expenditures without reimbursement. Household utility is $\bar{u}$ in all locations when reimbursement is required. When reimbursement is not required, household utility in all locations is $\bar{u} - \tau$, where τ represents the per-household cost of all defensive expenditures across all locations.[7]

The incidence of the natural hazard also depends on the payment regime. The benefits of defensive expenditures are entirely captured by homeowners in risky locations under either policy. When households pay for defensive expenditures, the costs of natural hazards are also fully borne by risky locations. When defensive expenditures are centrally provided, residents in all locations, including the risk-less reservation city, bear costs τ per household due to natural disasters in risky locations.[8]

The diagram illustrates two additional important points: The changes in population due to moral hazard are larger in areas with relatively low population levels, since at dense levels of population per-capita costs of protection are small. Additionally, in areas where the population level is high enough to be on the inelastic portion of the housing supply curve, the effects of public provision of defensive expenditures will change housing prices but will have little effect on quantity.

[7]While total welfare under the second policy is unambiguously lower, it is worth emphasizing that neither equilibrium maximizes total welfare. The spatial equilibrium equalizes per-capita utility (average utility) across locations; while the welfare-maximizing population distribution would equalize marginal utility (Glaeser, 1998; Bergstrom, 1986).

[8]In our simple spatial equilibrium setup, all households are equally well off by construction, so that this redistributive effect of natural hazard spending does not affect relative utility (only total welfare). Introducing preference heterogeneity or other moving restrictions into the model would change this.

4 Data

Our primary dataset combines wildfire suppression expenditures data with topographical and environmental conditions at the ignition point and parcel-level assessor data for the universe of western U.S. homes. The unit of observation is a single fire incident. For each incident, we construct measures of the proximity and number of homes within range of the fire, as well as a set of covariates describing the surface, environmental, and atmospheric conditions of the location in which the fire started. We describe the construction of this dataset in detail in this section, and provide additional details on the data cleaning process in the online appendix.

4.1 Fire suppression expenditures

We compile fire suppression cost data from five different sources, including four federal agencies and one state firefighting agency. The federal agencies we include are the United States Forest Service, the National Parks Service, the Bureau of Land Management, and the Bureau of Indian Affairs. The state agency is California's Department of Forestry and Fire Protection (CAL FIRE). We obtain these data either through publicly available sources or through Freedom of Information Act (FOIA) requests, and they represent by-incident spending for these agencies. Our geographical sampling frame is Western United States[9], where wildfires are most frequent and costly to suppress. We discuss each source of data in detail below, as well as the process by which we harmonize these datasets.

4.1.1 Forest Service fire expenditures

The Forest Service (FS) accounts for the largest share of fire suppression expenditures of any federal agency. We obtain historical by-incident suppression costs fires managed by the USDA Forest Service from 1995 to 2014 from the National Interagency Fire Management Integrated Database (NIFMID). These data are compiled by the Kansas City Fire Access Software (KCFAST)[10] and include data on suppression

[9]Specifically, our dataset includes only fires with ignition points in the states of Arizona, California, Colorado, Idaho, Montana, New Mexico, Nevada, Oregon, Utah, Washington, or Wyoming.

[10]Obtained here: `https://fam.nwcg.gov/fam-web/kcfast/html/ocmenu.htm`.

expenditures and fire locations. The forest service is primarily responsible for fighting fires that ignite in or near the boundaries of National Forest areas, displayed in dark green in Figure 1. Suppression expenditures in these data are intended to represent the cost of the suppression effort put forth by the Forest Service during a given wildfire incident. These costs reflect the deployment of personnel and equipment in the service of the suppression effort, but exclude the cost of long-run preventative effort, such as fuel thinning, that occurs separately of a particular suppression effort.

Over the course of our sampling frame, more than 150,000 wildfire incidents are logged in this database. However, since the Forest Service only reports per-fire cost data for fires above 300 acres, we limit this sample to the 2,563 fires in the 11 western states with a size of 300 acres or larger (the smallest size for which suppression expenditures are separately reported) with ignition date and location data available.

4.1.2 Department of Interior fire expenditures

Four separate agencies within the Department of Interior (DOI) are responsible for fire suppression. They are the Bureau of Land Management (BLM), the Bureau of Indian Affairs (BIA), the National Parks Service (NPS), and the U.S. Fish and Wildlife Service (FWS). We obtain data for BLM, BIA,and NPS through FOIA requests. BLM is responsible for fires which ignite on the 248 million acres of public lands they manage,[11] BIA is responsible for fires starting on the 55 million acres of Indian Lands, [12], and the NPS is responsible for fires igniting within its 417 park units across 84 million acres of land.[13] At the time of this writing, we have not yet been able to obtain firefighting suppression costs for FWS. For the remaining three agencies, the DOI data are available from 2003-2016. To match the data available from the Forest Service, we limit this sample to include only fires the affect more than 300 acres and apply similar data quality restrictions as those described for the FS data. Our final dataset includes 3,003 BLM fires, 418 BIA fires, and 240 NPS fires.

[11] BLM public land statistics: https://www.blm.gov/sites/blm.gov/files/PublicLandStatistics2016.pdf.

[12] Map: https://www.nifc.gov/PIO_bb/Agencies/BIA/MapBIARegions.pdf

[13] NPS fact sheet: https://www.nps.gov/orgs/1965/upload/wildland-fire-fact-sheet.pdf.

4.1.3 California fire expenditures

We also collect fire suppression cost data for California, which has the most frequent and costly wildfires of any state in the West. Suppression cost data for California come from a public records request to the California Department of Forestry and Fire Protection (CAL FIRE). In order to combine these costs data with wildfire location information, we merge three sets of administrative records. The first is a complete listing of all reported wildland fire incidents in the CAL FIRE protection area during 2007–2016, regardless of size. This dataset includes the ignition date, acres burned, CAL FIRE geographic unit, and, for incidents after mid-2011, the latitude and longitude of the ignition point. To supplement the location records for these fires, we also obtain shapefile data for a subset of report CalFire incidents.[14] The third dataset is an administrative record of firefighting expenditures at the incident level for 788 incidents during 2011–2016. According to CAL FIRE, these expenditure data are carefully tracked because they are the basis of cross-agency reimbursements for mutual aid expenditures – for example, reimbursements to California by the federal government under the FEMA FMAG program, or by local governments to CAL FIRE for firefighting assistance in incorporated areas.

4.1.4 Fire expenditures harmonization

To ensure consistent data quality, we harmonize the data across the five agencies from which we source suppression expenditures. Specifically, we ensure that ignition date, ignition location, responsible agency, cause of fire, area burned, and suppression cost data are present for all incidents and that the costs reflect values in 2014 dollars. Our final dataset includes 6,422 fires and account for 111 billion dollars of suppression costs.

4.2 Fire covariates

Using the harmonized location data, we obtain elevation, slope, aspect, and fuel model data for the ignition point of each fire from LANDFIRE (United States Department of

[14] Fire perimeters data here: http://frap.fire.ca.gov/data/frapgisdata-sw-fireperimeters_download.

Interior, 2013). The former three products are derived from the high-resolution National Elevation Dataset[15]; elevation represents the land height above sea level and is given in meters, slope represents the angle the land and is given in degrees, and aspect represents the direction of the slope and is given in degrees as well. The fuel model data are the 13 Anderson Fire Behavior Fuel Models (Anderson, 1982) and describe the fire potential of surface fuel components (e.g., the type of foliage in the area) on which the fire starts. We also obtain ignition-day weather (maximum and minimum temperatures, precipitation, and measure of humidity) from the PRISM daily weather dataset (Group, 2004).

4.3 Parcel data

The homes data include information on home locations, values, year built, and other property characteristics for the universe of 17,700,000 single-family homes in the western United States. These data are curated by CoreLogic and represent a compilation of tax assessor data from individual counties. We limit the sample to include only homes in partially vegetated areas that would be threatened by wildland fires, based on wildland-urban interface categories identified in Radeloff et al. (2005) (see appendix for details). Because the federal government controls so much land in the West, and so much residential development is in wildland areas, these sample exclusions are not that restrictive. Our analysis dataset includes 8,046,957 homes (about 47% of all single-family homes in the West). We use the US Forest Service Wildfire Fire Hazard Potential ratings from (Dillon, 2015) to assess physical fire risk at the parcel level.

5 The Cost of Saving Homes During Wildfires

5.1 Empirical Strategy

Even in the absence of any nearby private home development, some amount of resources would likely be devoted to managing and suppressing a fire. Thus, the first

[15]NED: `https://nationalmap.gov/PERS_Jan2002_NED_highlight_article.pdf`

step in our empirical analysis is to measure the share of firefighting expenditures that are dedicated to protecting homes.

In order to isolate spending dedicated to protecting homes, we estimate the causal effect of nearby homes on firefighting costs relative to a no-development counterfactual. A number of observable and unobservable factors should be expected to affect firefighting costs, including ecological characteristics, local weather trends, and the typical response behavior of local fire managers. Our empirical strategy addresses this identification challenge by taking advantage of randomness in ignition locations within U.S. national forests. Each of the national forests in our dataset experienced multiple large fires during our study period. We compare suppression costs for fires within the same national forest that happened to start at different distances from homes. Some fires start far away from private homes, for example deep inside the national forest, while other fires start nearer to homes, because the ignition point is closer to the national forest boundary or to a privately-owned "inholding", or because new homes have been built near the boundary. Figure 2 illustrates this variation for four example national forests. In each panel, the area of the national forest is shown in green. Fires are shown as x's and are colored by the distance from the ignition point to the nearest home. Fires that started more than 10 kilometers away from any home are shown in dark blue. Black markers indicate homes.

We take advantage of this variation in ignition locations using a fixed-effects estimation strategy. We model the effect of homes on fire suppression costs as,

$$ln(\text{Cost}_{ift}) = \beta f(\text{Homes}_{it}) + X_{ift}\rho + \delta_f + \omega_{st} + \eta_{ift} \tag{1}$$

$Cost_{ift}$ is the suppression cost for fire i in national forest f in month-of-sample t. We are primarily interested in how this cost depends on the potential threat posed by the fire to private homes, Homes_{it}. We begin in Section 5.2 by parameterizing Homes_{it} as the distance from the ignition point of the fire to the nearest home. In Section 5.3, we consider the total number of homes near the ignition point. In either case, our preferred model approximates $f()$ with a binned step function to allow a flexible response of costs to threatened homes (although our estimates are robust to a variety of functional forms).

The identifying assumption in this analysis is that unobserved determinants of fire cost, η_{ift}, are independent of the distance to the nearest home, conditional on national forest fixed effects and our other controls. This panel data approach addresses a number of omitted variables concerns. The national forest fixed effects δ_f control for unobservable determinants of firefighting cost that are constant at the national forest level. We also include time fixed effects ω_{st} that control flexibly for unobserved changes in firefighting costs over time. Our preferred specification includes state by month-of-year fixed effects and state by year fixed effects. Intuitively, this identification strategy amounts to comparing fires in the same national forest during the same month of the year and the same year of the sample (e.g., Flathead National Forest in September, 2003).

One potential remaining issue with this approach is that the locations of private homes are not randomly assigned. Even within a given national forest, areas near homes may differ systematically from areas far from homes in ways that affect firefighting cost. To address this possibility, we include additional control variables X_{ift}. These include the slope of the terrain at the ignition site, the geographic aspect, the vegetation type (fuel model), and weather conditions at the point of ignitions. We argue that the combination of this selection on observables approach with the fixed effects strategy described above provides the most credible causal identification strategy available in this setting. To ensure that fire timing and location is truly random (i.e., not driven by the presence of people), we also estimate a specification where we limit the sample to fires caused by lightning.

Our identification assumptions would fail if there are unobservables which increase firefighting costs disproportionately for fires near homes but are actually driven by the distance between the fires and nearby homes. To the extent that there are remaining unobservables of this nature, we argue that these differences likely bias our estimates towards zero, meaning that our approach conservatively measures the effect of homes on firefighting costs. Areas that are near homes are likely to lie in flatter regions that are more easily accessible by road, both factors which would make firefighting less costly rather than more. The online appendix includes a detailed examination of covariate overlap and balance according to distance from homes.

5.2 Proximity to homes

We begin by considering a version of Equation 1 where the threat to private homes, Homes_{it}, is proxied by the distance from the ignition point to the nearest home that existed at the time of the fire. We calculate this variable by merging ignition point data from the firefighting data to the geographic coordinates of all the homes in the real estate dataset. If, in the absence of suppression effort, wildfires are more likely to destroy homes that are close by, we might expect to find that firefighting effort is higher for fires that start near areas of private development.

Figure 3 shows regression estimates. We consider the total cost of Forest Service fires as a function of the distance from the fire's ignition point to the nearest home. The figure includes three different regression specifications. Each regression includes national forest fixed effects, state by month-of-year fixed effects, and state by year fixed effects. The solid black line shows the estimated marginal effect of distance from a regression of log costs on a cubic polynomial of distance to homes. The shaded gray area is the 95% confidence interval. The dashed black line shows a linear spline in distance to homes, with knots placed every 10 kilometers. Finally, the black dots report coefficients from a binned step function specification. These coefficients correspond to indicator variables for 5-kilometer bins of distance to homes. The omitted category is fires that start more than 50 kilometers from any home. Regardless of the functional form that we choose, there is a clear, steep gradient in firefighting costs with distance. The relationship is steep (noting that the y axis is in logs), monotonic and close to linear. Relative to a fire that starts 45 kilometers from any home, the log costs of a fire less than five kilometers from homes are higher by about 3. Taken literally, these estimates imply that a fire that starts less than 5 km from homes would cost 75% less if there were no homes within 25 km, and 93% less if there were no homes within 40 km.[16]

Table 2 estimates alternative models. Column (1) matches the figure. Column (2) adds additional controls for pre-determined fire characteristics. In general the signs and magnitudes of the included covariates match expectation. Firefighting costs are higher where the terrain slopes more steeply, reflecting difficulty of access. Costs

[16] These percentage changes are calculated using the binned specification. Halvorsen and Palmquist (1980) and Kennedy (1981) show that the percentage effect of an indicator variable in a semi-log regression can be approximated as $e^{\beta - 0.5V(\beta)} - 1$, where β is the regression coefficient.

also increase with wind speed on the ignition day, consistent with the importance of wind in fire spread. Vapor pressure differential (VPD) is a measure of atmospheric dryness, where higher values imply drier air; as expected, high VPD increases firefighting costs.[17] Costs are also higher for fires on south- or southwest-facing slopes, which receive additional sun exposure and thus tend to have more readily combustible vegetation. While many of these covariates have meaningful effects on firefighting costs, including them in the regression has little effect on our estimated distance gradient. This supports the logic of our empirical design, which is that while a number of factors likely affect firefighting costs, those factors do not appear to vary systematically with distance from homes after accounting for national forest fixed effects.

The remaining columns show three robustness checks. Column (3) replaces the time fixed effects with more granular week-of-sample by state fixed effects, which allow for arbitrary shocks to firefighting costs in each week of the dataset in each state. Designed to absorb any high-frequency local cost fluctuations that might be caused by weather patterns or other factors, this alternative specification produces a similar distance gradient. Because some state-week cells include only one fire, this model uses about 700 fewer observations than our preferred specification. Column (4) restricts the sample to fires started by lightning. Some types of human-caused fires are more likely to occur near populated areas, introducing a potential identification concern if fires due to arson or campfires or other causes vary systematically in their difficulty to extinguish. The locations of lightning strikes are plausibly random and thus purged of this potential bias. If anything, the estimated distance gradient is steeper when this restriction is applied, though the estimates are not different in a statistical sense. Column (5) restricts to fires occurring in timber areas, since developed areas are also less likely to be heavily wooded than more remote areas. As before, the estimated distance gradient steepens slightly under this restriction. This is consistent with our expectation that any omitted variables that might persist after our empirical design and control variables would bias our estimated effects downwards.

[17] VPD is the deficit between the observed vapor pressure and the vapor pressure at the current temperature if the air were fully saturated with water. Meteorologists have shown VPD to be an important measure of dryness and predictor of fire severity (Anderson, 1936; Seager et al., 2015).

5.3 Total Number of Homes

The results in the previous section imply that the *presence* of nearby private homes strongly affects firefighting costs. In this section we consider how this effect varies with the *density* of development. To do this, we fix a radius around each fire and estimate a version of Equation 1 that parameterizes Homes_{it} as the total number of homes within that radius. We use a 30 kilometer radius in our baseline specification. The online appendix shows results for alternative radii.

Table 3 shows the effect of home density on fire costs. We report results from the binned step function specification.[18] The reference bin is fires with zero homes within 30 km, and the other bins evenly divide the remaining fires.[19] We define bins by quantiles instead of equal intervals for this table because of the long right tail of the number of homes variable; however, our results are robust to the use of equal intervals as well . Column (1) shows our baseline results. The presence of up to 114 homes increases log costs by 0.87. For up to 625 homes, the cost effect increases to 1.41. Beyond that, costs increase very little with additional homes, even for fires threatening thousands or tens or thousands of homes. Column (2) includes distance to nearest home as a control, to account for the fact that fires occurring near more homes also occur slightly closer to homes on average (as we show in the online appendix) . To interpret the Column (2) estimates, consider the cost of a fire that ignites 11 kilometers from homes, the mean distance conditional on any homes within 30 km. The bracketed numbers to the right of Column (2) show the cost of this fire relative to a fire starting 30 kilometers from the nearest home. The estimates are similar to Column (1), with a slight reduction in the marginal cost of homes at high levels of density. This reinforces the sharp nonlinearity of the cost-density relationship.

Figure 4 shrinks the bins and focuses on the region over which costs are increasing.

[18] An alternative specification is a constant elasticity model (a log-log specification). The average elasticity from a regression of log costs on log threatened homes is 0.14, with a t-statistic of 3.9. This linear relationship in log-log space maps closely to the concave relationship that we measure in log-linear space. We focus on log-linear models in the text because we are interested in the effects of one additional home, as opposed to the effects of proportional increases in homes. The log-log model also does not accommodate fires with zero nearby homes.

[19] To reconcile these results with the proximity results, note that the average of the coefficients in the first five rows of Table 3 gives the average cost difference between fires with homes within 30 km, and those without. This is conceptually similar to the sample-weighted average of the coefficients for the 0-10 , 10-20 , and 20-30 km bins minus the sample-weighted average of the 30-40 and 40+ km bins in Table 2. Both exercises yield about 1.5.

The sample in this figure includes fires with fewer than 2,500 homes within 30 km of the ignition point. We plot predicted effects from three different regressions. Each regression controls for national forest fixed effects, state by month of year fixed effects, and state by year fixed effects. Across the three statistical models, costs increase quickly in density and then level off. The binned specification expands each of the bins from Table 3 into three separate bins, still based on quantiles of number of homes. These fine bins show that costs increase rapidly in the first homes. The first non-zero bin spans 1–19 nearby homes; the next two bins are 20–57 and 58–114. Even a small number of nearby homes raises costs substantially, almost as much as thousands or tens of thousands of homes.

5.4 Additional Results and Robustness Checks

In addition to the checks described above, we include a more detailed set of additional results and robustness checks in the online appendix, which we describe here in brief. The estimates of the impact of population density on firefighting cost are robust to the same control variables and checks shown in Table 2, such as controlling for weather or limiting to lighting-caused fires. Using the total transaction value of nearby homes yields similar results to the number of homes. We also show that the implied marginal effect of homes on fire costs depends intuitively on the radius within which we count homes, where smaller radii imply larger per-home marginal effects, but that the strong non-linear response of costs to number of homes exists for any reasonable choice of radius.

Because our baseline estimates are not suitable to consider the impact of homes on the *frequency* of fires in an area, we conduct a separate analysis to investigate how this might impact our findings. As some wildland fires are ignited by humans, increased human population may create more ignitions. On the other hand, new homes could be accompanied by greater fire prevention efforts. We explore this relationship using panel variation in new home construction near each of the national forests in our federal sample. We find weak evidence of a small positive effect of new home construction on the number of large fires each year in places that start from a low level of development. Adding an additional 1,000 homes in a relatively undeveloped area is associated with about a 3.5% increase in the number of fires each year, or about 0.06 additional fires per year. The finding that human presence increases fire frequency is consistent with work by ecologists (Syphard et al., 2007; Massada et al., 2012; Faivre et al., 2014). This implies that we slightly underestimate the additional firefighting cost created by new homes.

Finally, in the appendix we also use state of California (CalFire) firefighting cost data to validate some of our results from this section. Forest Service fires provide useful natural variation, since these public lands include some areas that are very far from homes. Fires fought by CalFire, on the other hand, almost all occur within 10 kilometers of homes. However, we are still able to use the CalFire data to confirm that the marginal cost of additional homes within 30 km is very small beyond low levels of development. Our CalFire density results are similar to our Forest Service density results.

6 The Implicit Subsidy To Homeowners

The results in the previous section show that a large share of wildland firefighting expenditures are dedicated to protecting private homes. In this section, we estimate the incidence of this implicit subsidy to homeowners. For every individual home in the western United States, we calculate an actuarial measure of the expected net present value of the government's cost of protecting the home during future wildfires.

6.1 Calculating Realized and Expected Protection Costs

In the first step of this calculation, we use the estimated model in Equation 1 to predict the amount of firefighting expenditures on each historical fire that were due to the presence of homes. To do this, we take the difference between the predicted costs for each fire from that regression ($\hat{cost_i}$), and the predicted costs implied by the estimated model for the fire if the nearest home had been located 40 km away ($\tilde{cost_i}$). This difference yields the expenditures due to homes on each fire, which we call Δ_i.

For each fire i, we allocate Δ_i over homes within a fixed radius of the ignition point that were potentially threatened by the fire. Our definition of potentially threatened homes includes homes located within 40 kilometers of the ignition point in areas with wildland vegetation. Our classification of wildand vegetation categories follows Radeloff et al. (2005) and is described in detail in the appendix. Within the set of homes potentially threatened by each fire, we assign a larger share of Δ_i to homes closer to the ignition point. We use two approaches to this weighting, an inverse-distance weighting (IDW) algorithm and an empirical estimate based on the results in section 5.1. For the IDW algorithm, houses within 40km are assigned a weight of $\frac{1}{d}$, weights are normalized to one within each fire, and home protection expenditures by fire are divided using the normalized weights. Our second and preferred approach is identical except that the weights assigned to each fire-parcel combination are the estimated coefficients from Equation 1 for distance between the ignition point and the parcel location, normalized to sum to one for each fire. This exercise divides Δ_i across j potentially threatened homes, yielding costs δ_{ij} for each home, where $\sum_{j=1}^{J} \delta_{ij} = \Delta_i$.

The next step of this calculation sums up the total costs associated with each home during 1995–2014. For each home j, we add up that home's costs for each fire during the study period, $\rho_j = \sum_{i=1}^{I} \delta_{ij}$. We call this quantity the *realized protection cost* for home j because it represents the amount of firefighting expenditure associated with the home during the study period.

Our estimate of interest is not past expenditures, but expected future expenditures for each in the dataset. The observed history of firefighting costs is 20 years or less, which in many regions may not be a long enough period to accurately describe the underlying fire risk. To estimate expected firefighting costs, we group regions with similar ecological and fire risk characteristics together into actuarial groups, much like a private insurer would be expected to do when calculating risk. We calculate expected cost for homes in each group as,

$$\mathop{\mathbb{E}}_{h,d,s} [\rho_j]$$

This calculation takes expectations over bins of wildfire hazard h, housing density d, and geographic region g. Wildfire hazard is defined at the parcel level using the spatially-explicit wildfire hazard potential scores provided by Dillon (2015), which are a physical measure of wildfire risk taking into account ecological and geological factors. The appendix includes more information on this physical risk measure as well as a map of wildfire hazard potential. Housing density (population per square meter) comes from the Gridded Population of the World dataset (GPWv4), which reports population density within 1 km grid cells (Doxsey-Whitfield et al., 2015). We define geographic regions based on the boundaries of the seven Geographic Area Coordinating Centers (GACCs) that coordinate regional firefighting operations in the West. To reflect the ongoing nature of the firefighting guarantee, we calculate the net present value of the expected annual costs for each group of homes. We call this quantity the *expected parcel protection cost*. It represents the present value of the expected government expenditures for fire protection associated with each home.

6.2 Exploring Expected Protection Costs

Figure 5 shows the distribution of expected parcel protection costs. These costs were calculated using 210 actuarial groups of homes created by crossing six bins of physical fire risk, five bins of housing density, and the seven wildland firefighting regions. The sample of homes in this figure includes all 8 million homes in the western U.S. located near areas of wildland vegetation (about 47% of homes). The black line shows the cumulative distribution function of the expected present value of firefighting costs. The dashed gray line shows firefighting costs divided by the transaction value of the property. Most western homes have expected protection costs of a few hundred dollars or less, while the highest-risk homes have costs that are much larger. Five percent of homes have expected protection costs exceeding $5,600, or about 4.4% of home value. These homes belong to 34 separate actuarial groups throughout the West. One percent of homes have expected protection costs exceeding $13,900, or about 10.3% of home value. These homes belong to 11 actuarial groups.

Figure 6 shows the broad geographic distribution of expected protection costs. This map shows the average expected protection cost for homes in each 20 kilometer hexagonal cell. The color scale corresponds to increasing costs. The scale is top-coded, so that the darkest red corresponds to homes with expected protection costs of $15,000 or more. Gray areas represent unpopulated regions and populated regions with no wildland vegetation (e.g., cities). Average expected protection costs are highest in

Northern California, central Oregon and Washington, and Idaho and western Montana.

Figure 7 explores this variation in expected protection costs in more detail. The four panels in the figure show how protection costs vary along four different margins. Panel A shows that protection costs are increasing in our physical measure of underlying fire risk. On average, expected protection costs for homes in the highest category of fire risk are about six times higher than for the lowest category. This relationship is intuitive, but it is also a reassuring validity test on our calculations. Panel B shows that expected parcel protection costs are strongly decreasing in housing density. This somewhat more surprising result is likely due to the nonlinear relationship between firefighting costs and housing density that we documented in Section 5.3. Increases in density are strongly associated with decreasing per-home costs, with expected costs in the lowest decile of density higher than the highest decile by a factor of ten or more.

Panels C and D consider the distributional effects of firefighting expenditures. A frequently-repeated claim about wildfire suppression in the United States is that it primarily benefits the rich (see, for example, "A Case for Letting Malibu Burn" (Davis, 1995)). The opposite appears to be true. Panel C shows that homes in low-income areas receive substantially more benefit from government firefighting efforts on average, compared to homes in high-income areas. This likely reflects the fact that the areas with the highest per-home expected protection costs are low-density rural and semi-rural areas. Wildfire protection costs are lowest in cities, where incomes are higher. Panel D considers an alternative measure of wealth, which is the transaction value of the home. For most American homeowners, the asset value of the home is a strong predictor of overall wealth. Again, the highest protection costs on average are associated with low-value homes. The relationship between average expected cost and home value is U-shaped, with increasing costs for high-value homes. This may reflect greater government efforts to protect high-value homes during wildfires, or high-value second homes located in areas where permanent residents have low incomes ("tourist towns").

Significant local variation in wildfire risk and development density in the West means that expected protection costs vary substantially over small distances. Figures 8A and 8B illustrate this local variation for two areas in California. These maps show the net present value of per-home expected protection costs, averaged at the Census block level for plotting. Figure 8A shows Shasta and Tehama counties in Northern California. This part of California experiences frequent wildfires every summer. Expected protection costs are several hundred dollars per home or less in the more densely-developed areas of Redding and Anderson. Outside of these urban areas, protection costs increase quickly. In some of the more remote Census blocks that border national forest lands or other public wildlands, costs reach $15,000 or more per home. These areas have a high underlying physical risk of fire, meaning that homes built here are likely to repeatedly require costly firefighting efforts to avoid destruction. In addition, these areas include fewer total homes, raising the per-home cost of firefighting. Figure 8B shows San Diego County in Southern California. Again, fire protection costs per home are low in the densely developed areas of San Diego, and increase in the high fire-risk, low-housing-density areas that border federal- and state-owned lands in the eastern part of the County.

7 Conclusion

The federal government spends billions of dollars each year to protect private homes from wildfires. We find that efforts to protect private homes account for about two-thirds of this spending. Interestingly, expenditures vary only slightly with the total number or value of homes threatened, conditional on any homes being threatened. This means development density is an important predictor of per-home protection costs.

Fire suppression spending represents a remarkably large transfer of federal revenues to a small number of landowners in high-cost places. In our highest-risk group, the expected NPV of the implicit subsidy is over 10 percent of total property value. This spending will continue to increase as climate change worsens the fire problem. Meanwhile, in the absence of policies to make homeowners internalize fire costs, the rate of new home construction in high-risk places is likely to continue unabated, implying substantial under-adaptation to this particular impact of climate change. For policymakers interested in addressing this incentive conflict, our empirical analysis provides a road map for calculating the optimal "fire protection fee" for new housing construction in currently undeveloped areas. More broadly, our results emphasize the importance of considering incidence and moral hazard effects of public expenditures on climate change adaptation, including for flooding, sea level rise, drought, and other impacts.

References

Abatzoglou, John T and A Park Williams, "Impact of anthropogenic climate change on wildfire across western US forests," *Proceedings of the National Academy of Sciences*, 2016, *113* (42), 11770–11775.

Albouy, David, Walter Graf, Ryan Kellogg, and Hendrik Wolff, "Climate amenities, climate change, and American quality of life," *Journal of the Association of Environmental and Resource Economists*, 2016, *3* (1), 205–246.

Anderson, Donald B., "Relative Humidity or Vapor Pressure Deficit," *Ecology*, 1936, *17* (2), 277–282.

Anderson, Hal E, "Aids to determining fuel models for estimating fire behavior," 1982.

Bergstrom, Theodore, "Soldiers of Fortune?," in "in," Vol. 2: Equilibrium Analysis of *Essays in Honor of Kenneth J. Arrow*, Cambridge University Press, 1986.

Boustan, Leah Platt, Matthew E Kahn, and Paul W Rhode, "Moving to higher ground: Migration response to natural disasters in the early twentieth century," *The American Economic Review*, 2012, *102* (3), 238–244.

Champ, Patricia Ann, Geoffrey H Donovan, and Christopher M Barth, "Homebuyers and wildfire risk: a Colorado Springs case study," *Society & Natural Resources*, 2009, *23* (1), 58–70.

Davis, Mike, "The Case for Letting Malibu Burn," *Environmental History Review*, 1995, *19* (2), 1–36.

Dillon, Gregory K., "Wildfire Hazard Potential (WHP) for the coterminous United States (270-m GRID), version 2014 classified," Forest Service Research Data Archive 2015.

Doxsey-Whitfield, Erin, Kytt MacManus, Susana B. Adamo, Linda Pistolesi, John Squires, Olena Borkovska, and Sandra R. Baptista, "Taking Advantage of the Improved Availability of Census Data: A First Look at the Gridded Population of the World, Version 4," *Papers in Applied Geography*, 2015, *1* (3), 226–234.

Faivre, Nicolas, Yufang Jin, Michael L Goulden, and James T Randerson, "Controls on the spatial pattern of wildfire ignitions in Southern California," *International Journal of Wildland Fire*, 2014, *23* (6), 799–811.

for International Earth Science Information Network CIESIN Columbia University, Center, "Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 10," 2017.

Gebert, Krista M, David E Calkin, and Jonathan Yoder, "Estimating suppression expenditures for individual large wildland fires," *Western Journal of Applied Forestry*, 2007, *22* (3), 188–196.

Glaeser, E.L., "Should transfer payments be indexed to local price levels?," *Regional Science and Urban Economics*, 1998, *28* (1), 1 – 20.

Group, PRISM Climate, "PRISM Database," Online Feb 2004. http://prism.oregonstate.edu.

Gude, Patricia, Ray Rasker, and Jeff van den Noort, "Potential for Future Development on Fire-Prone Lands," *Journal of Forestry*, 2008, *106* (4), 198–205.

Gyourko, Joseph, Matthew Kahn, and Joseph Tracy, "Chapter 37 Quality of life and environmental comparisons," in "Applied Urban Economics," Vol. 3 of *Handbook of Regional and Urban Economics*, Elsevier, 1999, pp. 1413 – 1454.

Halvorsen, Robert and Raymond Palmquist, "The Interpretation of Dummy Variables in Semilogarithmic Equations," *American Economic Review*, 1980, *70* (3), 474–75.

Hammer, Roger B, Susan I Stewart, and Volker C Radeloff, "Demographic trends, the wildland–urban interface, and wildfire management," *Society and Natural Resources*, 2009, *22* (8), 777–782.

Kahn, Matthew E and Randall Walsh, "Cities and the Environment," *Handbook of Regional and Urban Economics*, 2015, *5*, 405–465.

Kennedy, Peter E, "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations [The Interpretation of Dummy Variables in Semilogarithmic Equations]," *American Economic Review*, September 1981, *71* (4), 801–801.

Kousky, Carolyn and Sheila Olmstead, "Induced Development in Risky Locations: Fire Suppresion and Land Use in the American West," Working Paper 2012.

_ , Erzo F. P. Luttmer, and Richard J. Zeckhauser, "Private investment and government protection," *Journal of Risk and Uncertainty*, 2006, *33* (1), 73–100.

Liang, Jingjing, Dave E Calkin, Krista M Gebert, Tyron J Venn, and Robin P Silverstein, "Factors influencing large wildland fire suppression expenditures," *International Journal of Wildland Fire*, 2008, *17* (5), 650–659.

Massada, Avi Bar, Alexandra D Syphard, Susan I Stewart, and Volker C Radeloff, "Wildfire ignition-distribution modelling: a comparative study in the Huron–Manistee National Forest, Michigan, USA," *International Journal of Wildland Fire*, 2012, *22* (2), 174–183.

of Inspector General Western Region, Office, "Large Fire Suppression Costs," Technical Report, United States Department of Agriculture nov 2006.

of Interior, Geological Survey United States Department, "LANDFIRE Database," Online may 2013. http://landfire.cr.usgs.gov/viewer/.

Radeloff, Volker C, Roger B Hammer, Susan I Stewart, Jeremy S Fried, Sherry S Holcomb, and Jason F McKeefry, "The wildland–urban interface in the United States," *Ecological applications*, 2005, *15* (3), 799–805.

Roback, Jennifer, "Wages, rents, and the quality of life," *Journal of political economy*, 1982, *90* (6), 1257–1278.

Rosen, Sherwin, "Wage-based indexes of urban quality of life," *Current issues in urban economics*, 1979, *3*, 490.

Saiz, Albert, "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 2010, *125* (3), 1253–1296.

Seager, Richard, Allison Hooks, A. Park Williams, Benjamin Cook, Jennifer Nakamura, and Naomi Henderson, "Climatology, Variability, and Trends in the U.S. Vapor Pressure Deficit, an Important Fire-Related Meteorological Quantity," *Journal of Applied Meteorology and Climatology*, 2015, *54* (6), 1121–1141.

Stephens, Scott L, Brandon M Collins, Eric Biber, and Peter Z Fulé, "US federal fire and forest policy: emphasizing resilience in dry forests," *Ecosphere*, 2016, *7* (11).

Syphard, Alexandra D, Volker C Radeloff, Jon E Keeley, Todd J Hawbaker, Murray K Clayton, Susan I Stewart, and Roger B Hammer, "Human influence on California fire regimes," *Ecological applications*, 2007, *17* (5), 1388–1402.

Figure 1: Defensive Expenditures and the Housing Market

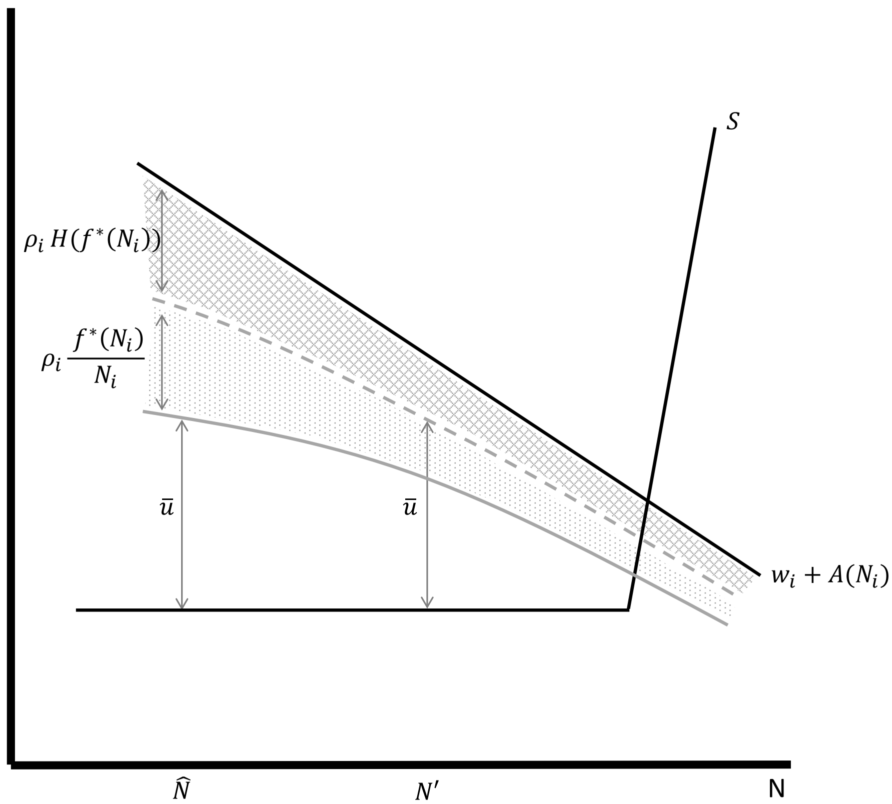
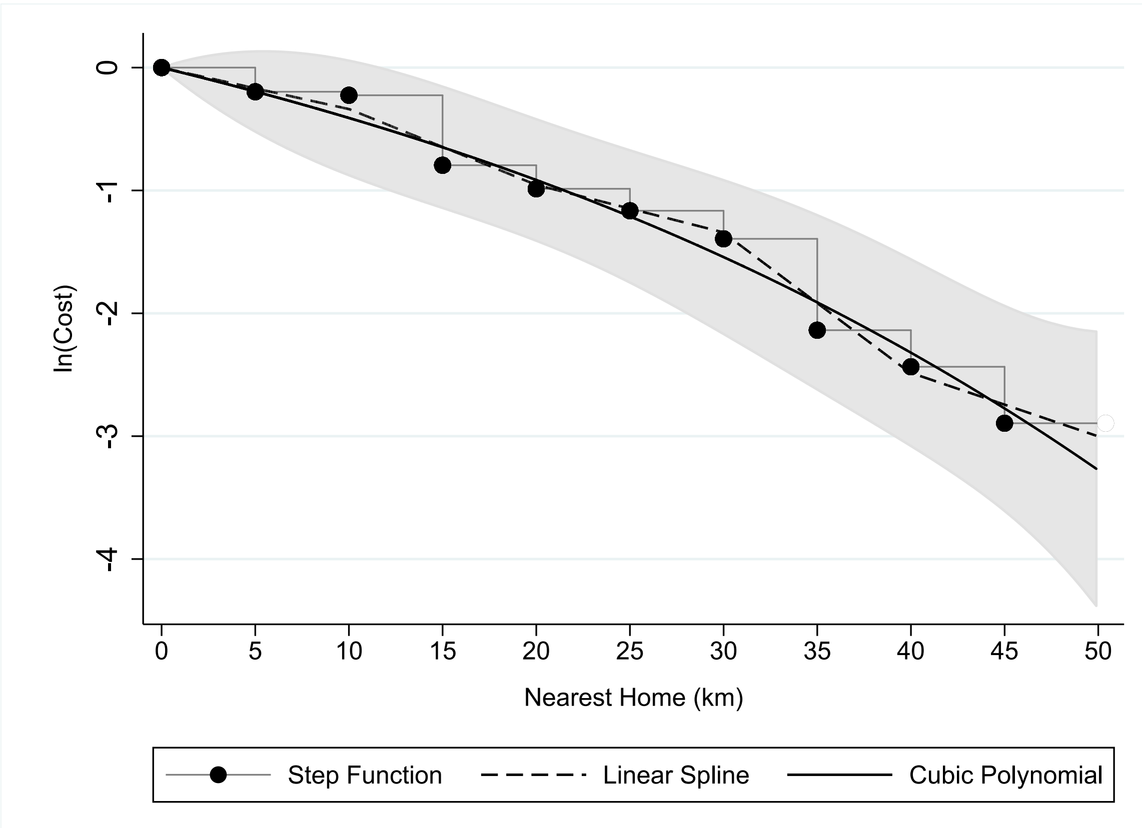


Figure 2: Example National Forest Units

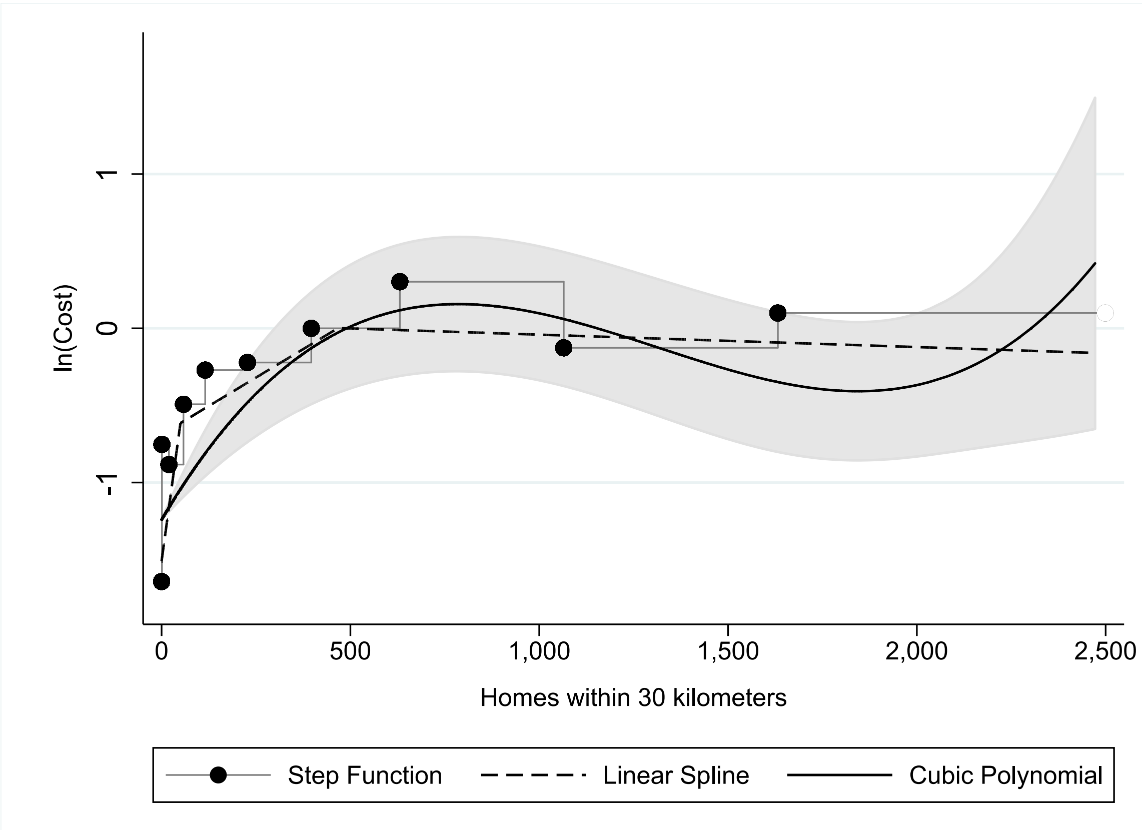
Each panel shows a single national forest area in green. The X's represent individual wildfires, colored according to the distance to the nearest home. Black dots indicate private homes. Clockwise from upper left, the forests are Shasta Trinity National Forest (California), Los Padres National Forest (California), Okanogan-Wenatchee National Forest (Washington), and Flathead National Forest (Montana).

Figure 3: The Effect of Homes on Firefighting Costs



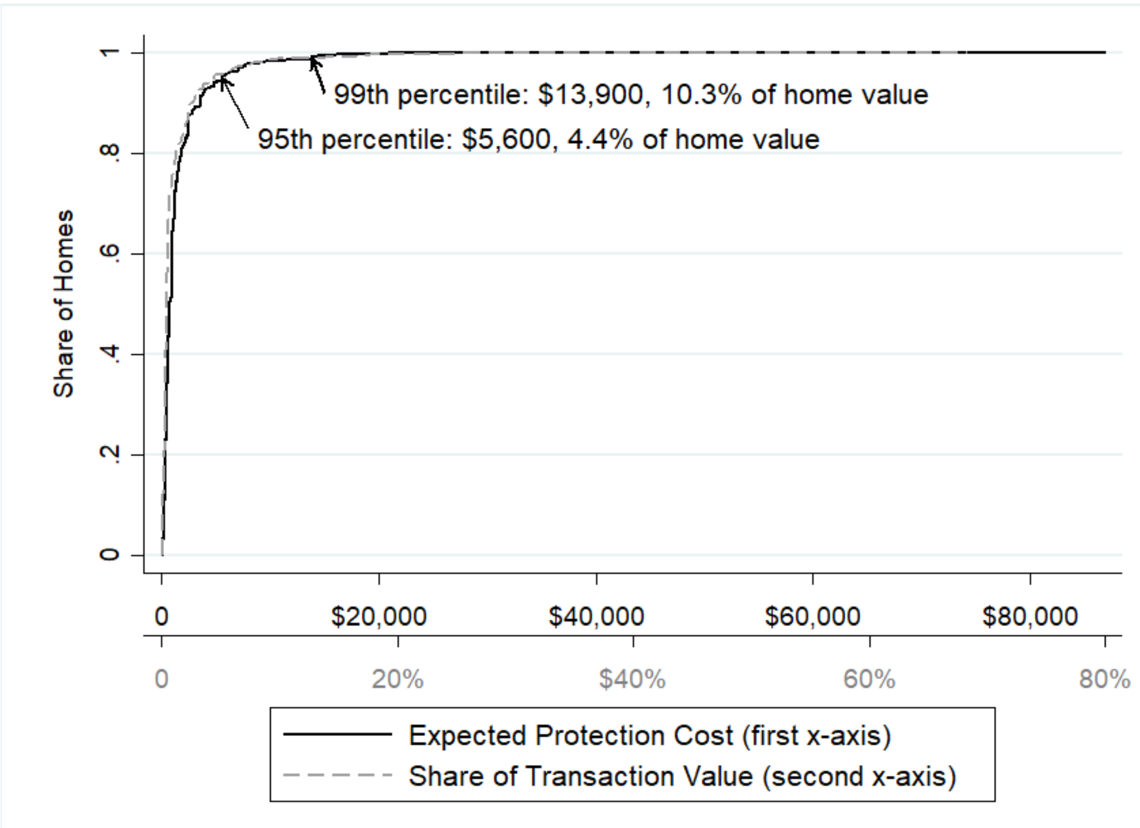
This figure reports results of three separate regressions of log firefighting cost on distance from the ignition point to the nearest home. The step function plots coefficients from a regression of log costs on indicators for 5 km distance bins. The linear spline is piecewise linear regression with knots every 10 km. The gray shaded area around the cubic polynomial is the 95% confidence interval for that model. Standard errors are clustered by national forest. Each regression includes national forest fixed effects, state by month-of-year fixed effects, and state by year fixed effects.

Figure 4: Non-linear effects of the number of nearby homes



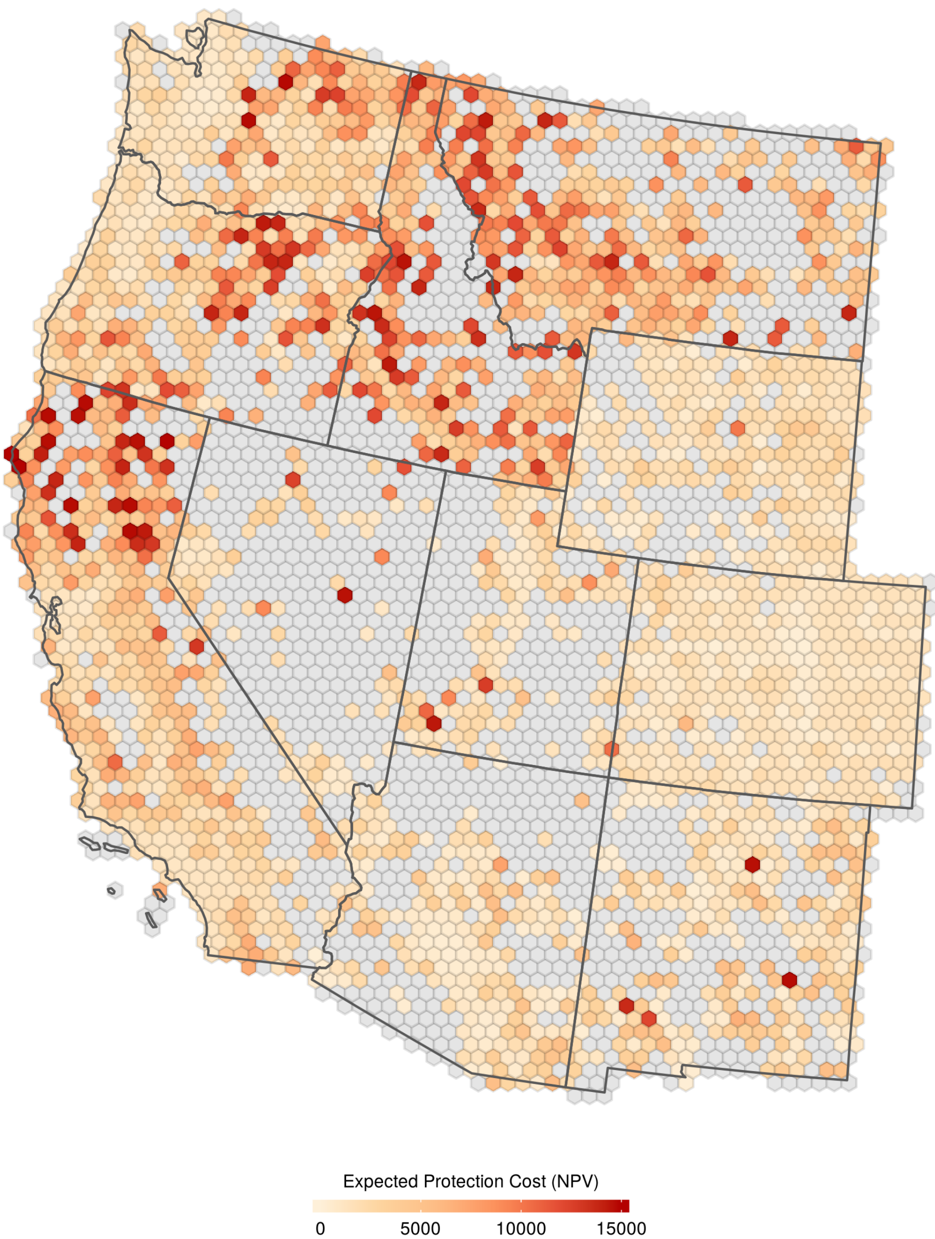
Notes: The sample for this figure includes 1,503 fires that had fewer than 2,500 homes within 30 kilometers of the ignition point. The lines show three separate regressions. The step function plots coefficients from a regression of log suppression cost on indicator bins for nine equal-observation groups plus a separate bin for fires near zero homes. The linear spline is a piecewise linear regression with knots at the 33rd and 66th percentiles of nearby homes. The gray shaded area around the cubic polynomial is the 95% confidence interval for that model. Standard errors are clustered by national forest. Each regression includes national forest fixed effects, state by month of year fixed effects, and state by year fixed effects. Predicted costs for each model are normalized to zero at the sample average number of nearby homes.

Figure 5: Distribution of Expected Parcel Protection Costs for 8 Million Western Homes



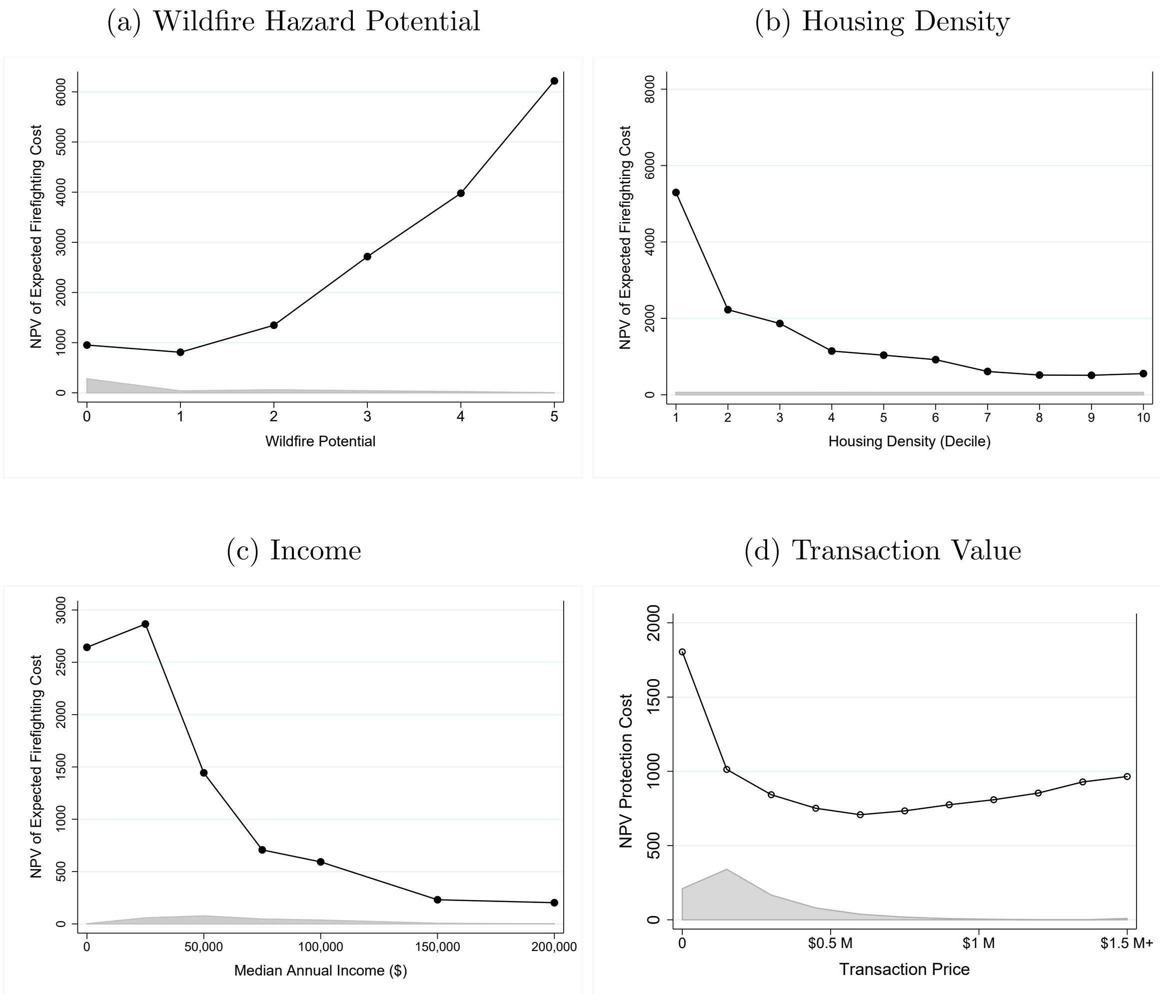
This figure describes the distribution of expected future firefighting costs for 8,046,957 homes in the western United States. These costs are calculated using 210 actuarial groups of similar-risk homes. Actuarial groups were created by crossing six categories of landscape fire risk, five categories of housing density, and seven wildland firefighting dispatch regions (GACC regions). The black line shows the cumulative distribution function for net present value firefighting costs, using a 5% discount rate. The dashed gray line shows the cumulative distribution function of firefighting costs divided by transaction value of the home, and corresponds to the second horizontal axis labels.

Figure 6: Expected Protection Cost by Region



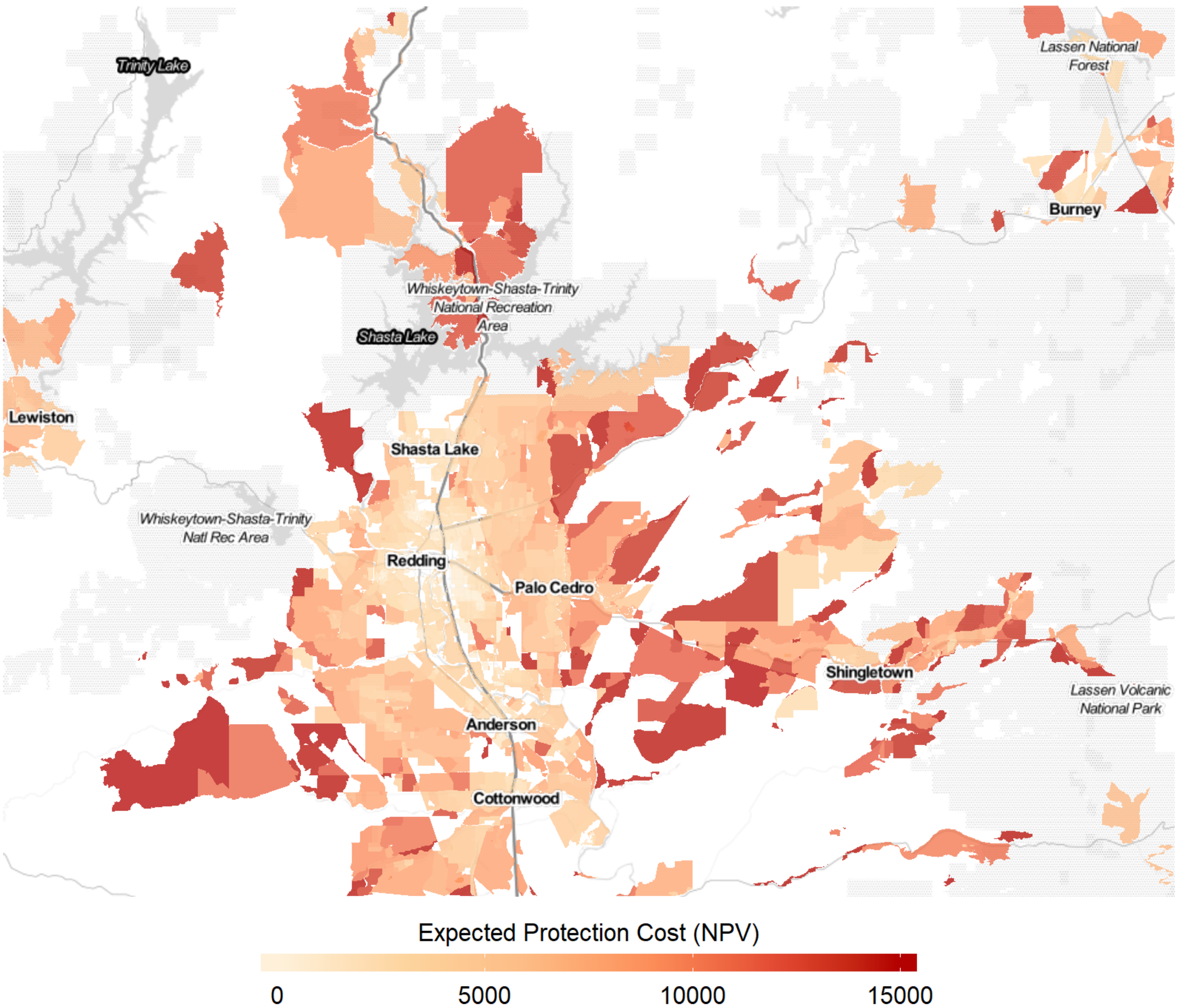
Notes: This figure shows the net present value of the expected future cost of protecting a home during wildfires, averaged across 20 km hex cells. The sample includes 8 million homes near wildland vegetation areas (47% of all western homes), and includes fire suppression costs from all agencies in our sample. See section 6 for a detailed description of the construction of this measure.

Figure 7: Expected Parcel Protection Cost by Fire Risk, Housing Density, and Wealth



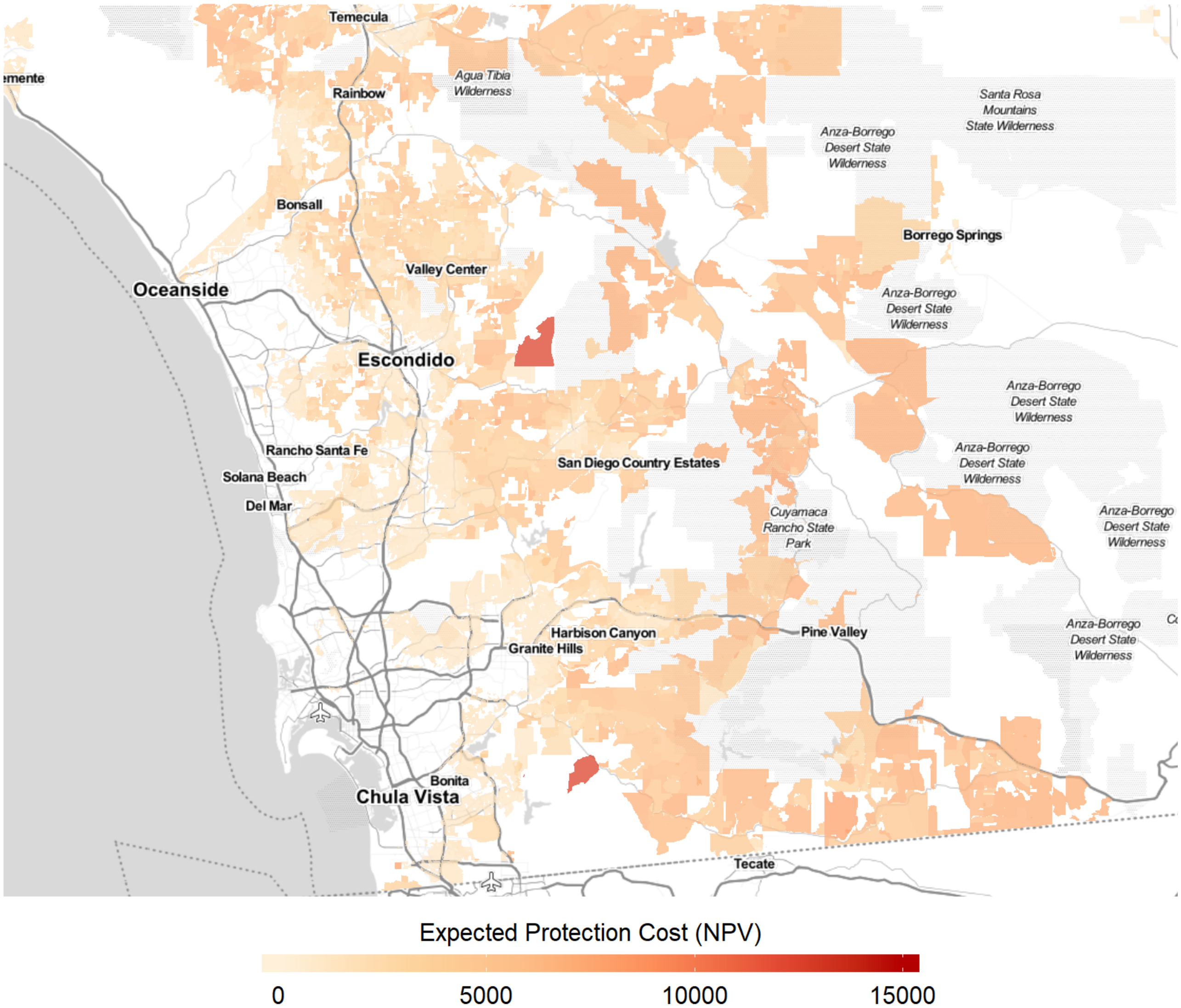
Each panel shows the variation in the net present value of expected protection costs along a single margin of interest. The black line in each panel shows average expected protection costs. The gray density shows the distribution of homes. Panel (a): The six categories correspond to wildfire hazard potential risk categories in Dillon (2015). Panel (b): Costs are plotted according to deciles of pixel-level population density for the study area from the Gridded Population of the World database Columbia University CIESIN, 2017. Panel (c): Each home is assigned the median annual income for its Census block group from the 2015 American Community Survey.

Figure 8A: Local variation in Expected Cost



This map shows expected protection costs averaged by Census block for Shasta and Tehama counties in Northern California. The color scale indicates the average expected NPV of a home's protection cost, and is top-coded at $15,000. Crosshatched areas are public lands. White areas have no wildland vegetation (e.g., urban areas) or no homes. The online appendix includes example maps for additional areas throughout the West.

Figure 8B: Local variation in Expected Cost, Continued



This map shows expected protection costs averaged by Census block for San Diego County, California. The color scale indicates the average expected NPV of a home's protection cost, and is top-coded at $15,000. Crosshatched areas are public lands. White areas have no wildland vegetation (e.g., urban areas) or no homes. The online appendix includes example maps for additional areas throughout the West.

Table 1: Descriptive statistics

Panel A: Pooled fire characteristics

	Mean	P10	P50	P90
Area burned	7,977	376	1,408	16,613
Aspect	137	-1	126	305
Elevation	1,503	656	1,507	2,328
Fire cost	1,511,356	4,500	124,279	3,069,388
Nearest home distance	13	1	9	32
Parcels in 10km	764	0	0	1,230
Parcels in 20km	3,340	0	96	7,391
Parcels within 5km	168	0	0	125
Value in 10km	181,714	0	0	168,827
Value in 20km	800,259	0	11,044	1,264,275
Value in 5km	42,957	0	0	18,026
Precipitation	0	0	0	1
Slope	11	1	9	28
Temperature	21	13	21	27
Vapor Pressure Deficit	21	11	21	32

Panel B: Fire characteristics by agency

	USFS	BLM	BIA	NPS	CAL FIRE
Number of fires	2,563	3,003	418	240	198
Area burned	7,966	7,904	5,982	10,565	10,306
Fire cost	2,930,690	222,295	728,850	456,903	5,619,710

This table report descriptive statistics for the 6,422 fires with area greater or equal to 300 acres in our sample. P10, P50, and P90 indicate the 10th, 50th (median), and 90th percentile of values. Aspect is given in degrees, elevation is in meters above sea level, fire cost is in 2014 US $, nearest home distance is in kilometers, parcels is the number of parcels within the given distance, value is the total parcel value (land and improvements), precipitation is in mm, slope is in degrees, temperatures is in Celsius, and Vapor Pressure Deficit is in millibars.

Table 2: The Effect of Proximity to Homes on Firefighting Costs

	(1)	(2)	(3)	(4)	(5)
Distance to Homes (km)					
10–20	-0.24	-0.31**	-0.25	-0.32	-0.42
	(0.15)	(0.14)	(0.24)	(0.20)	(0.27)
20–30	-0.92***	-0.99***	-0.92***	-1.04***	-1.41***
	(0.25)	(0.26)	(0.32)	(0.34)	(0.28)
30–40	-1.54***	-1.60***	-1.36***	-1.57***	-1.86***
	(0.43)	(0.44)	(0.40)	(0.48)	(0.43)
40+	-1.94***	-1.96***	-1.74***	-2.00***	-2.35***
	(0.36)	(0.41)	(0.28)	(0.41)	(0.35)
Additional Controls					
Terrain Slope		0.007**	0.006	0.008*	0.006
		(0.003)	(0.005)	(0.004)	(0.005)
Wind (mph)		0.026	0.004	0.014	0.009
		(0.018)	(0.034)	(0.027)	(0.039)
Vapor Pressure Differential		0.021*	0.012	0.010	0.021
		(0.011)	(0.016)	(0.014)	(0.020)
South/southwest-facing		0.24*	0.32	0.30	0.10
		(0.14)	(0.22)	(0.20)	(0.28)
Fuel Model FE		X	X	X	
National Forest FE	X	X	X	X	X
Year by State FE	X	X		X	X
Month-of-Year by State FE	X	X		X	X
Week-of-Sample by State FE			X		
Lightning fires only				X	
Timber Fuels only					X
N	2,069	2,069	1,365	1,437	1,018

This table reports the results of five separate OLS regressions. The sample includes western U.S. fires managed by the Forest Service during 1995-2014. In each regression the dependent variable is the natural log of suppression cost. The table rows report coefficients and standard errors on dummy variables corresponding to distance to the nearest home. The omitted category is 0–10 kilometers. Terrain slope is the linear slope of the ground surface. Wind speed is average speed on the day of ignition at the reference weather station listed in NIFMID. Vapor pressure deficit is for the ignition location and day, from PRISM, and measured in hectopascals (millibars). Fuel model fixed effects include four categories corresponding to NFDRS fuel codes for brush, grass, slash, and timber. Forest unit fixed effects include the 88 national forests in the Western U.S. Standard errors are clustered at the national forest level.

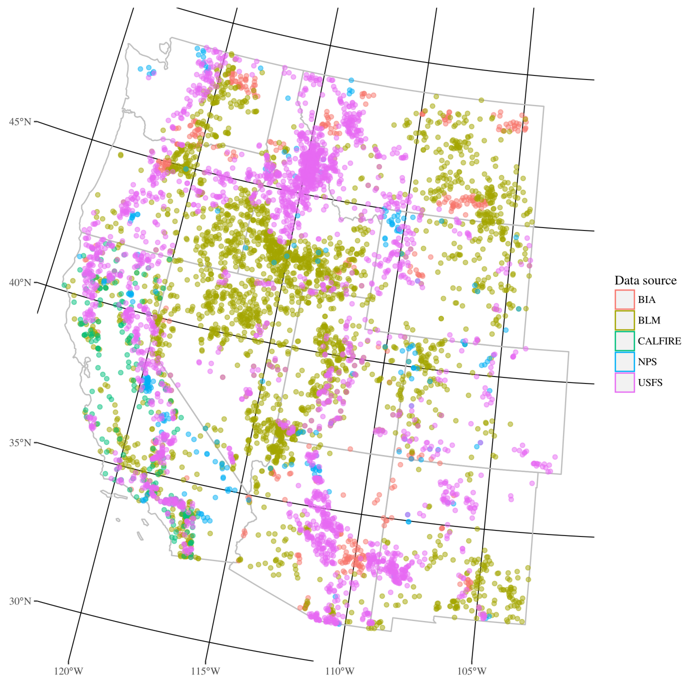
Table 3: Firefighting Costs by Number of Nearby Homes

	(1)	(2)	[Column 2 implied effect]
Number of homes			
1-114	0.87***	0.39	[0.93]
	(0.28)	(0.25)	
115-625	1.41***	0.80**	[1.34]
	(0.34)	(0.31)	
626-2,498	1.59***	0.88**	[1.41]
	(0.41)	(0.36)	
2,499-8,523	1.71***	0.91***	[1.44]
	(0.34)	(0.31)	
8,524+	1.76***	0.91**	[1.45]
	(0.40)	(0.35)	
distance		-0.0439**	
		(0.0168)	
distance^2		0.0003	
		(0.0002)	
Fires	2,069	2,069	

This table reports the results of two separate regressions. The sample includes western U.S. fires managed by the Forest Service during 1995-2014 larger than 300 acres. In each regression the dependent variable is the natural log of suppression cost. We report coefficients and standard errors on dummy variables corresponding to equal-observation bins of number of homes within 30 km. The omitted category is fires with zero homes within 30 km. "Distance" is the distance from the ignition point to the nearest home, in kilometers. The bracketed numbers to the right of the Column (2) estimates show the implied difference in log costs between a fire 11 km from the given number of homes and a fire with the nearest home 30 km away (for comparison with Column (1)). Both regressions include national forest fixed effects, state by month of year fixed effects, and state by year fixed effects. Standard errors are clustered at the national forest level.

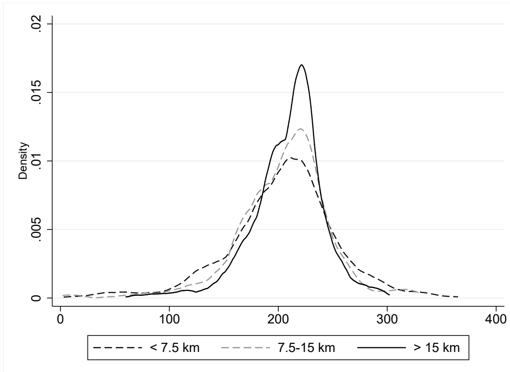
1 Additional Results and Robustness Checks

Appendix Figure 1: Western Wildfires, 1995–2004

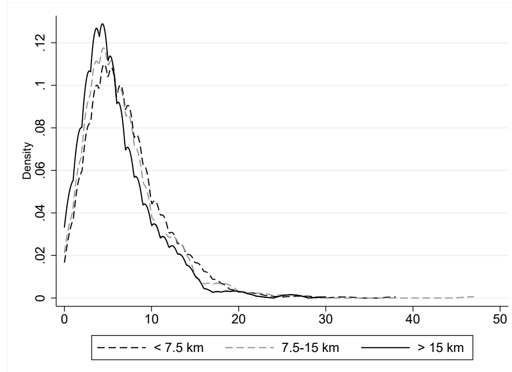


Appendix Figure 2: Covariate Overlap by Distance from Ignition Point to Nearest Home

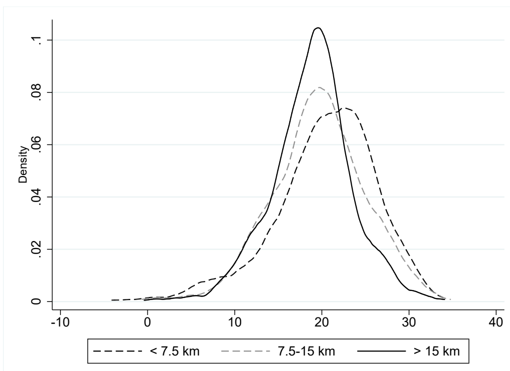
(a) Day of Year (Ignition)



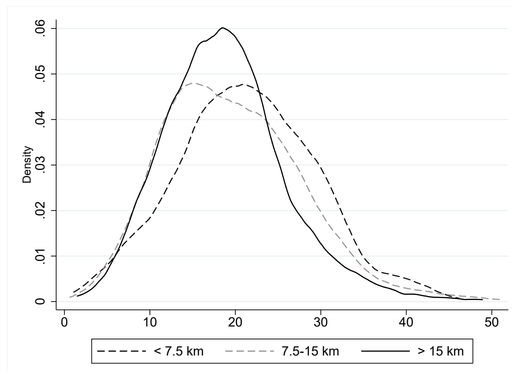
(b) Wind Speed (mph)



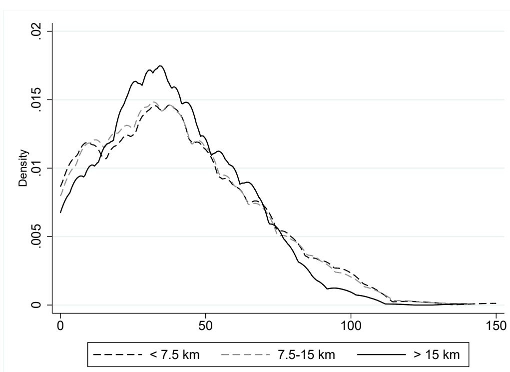
(c) Temperature (F)



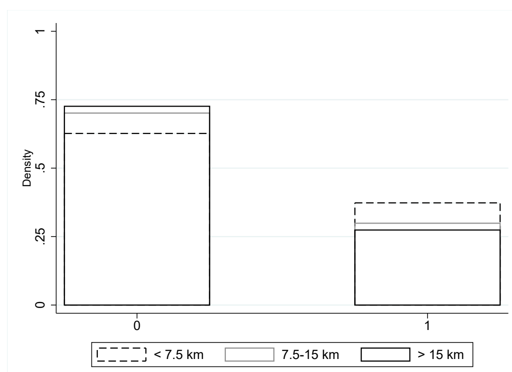
(d) Vapor Pressure Differential



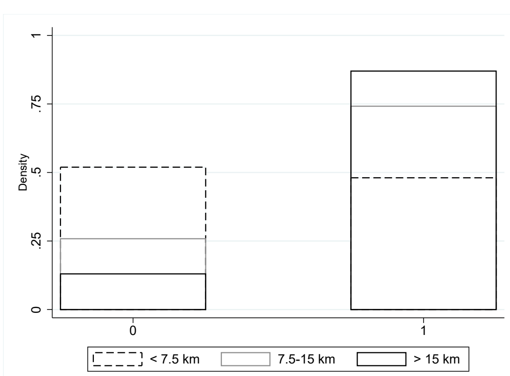
(e) Terrain Slope



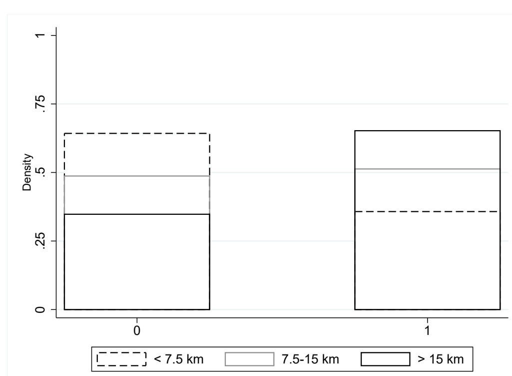
(f) South/southwest-facing



(g) Lightning-caused

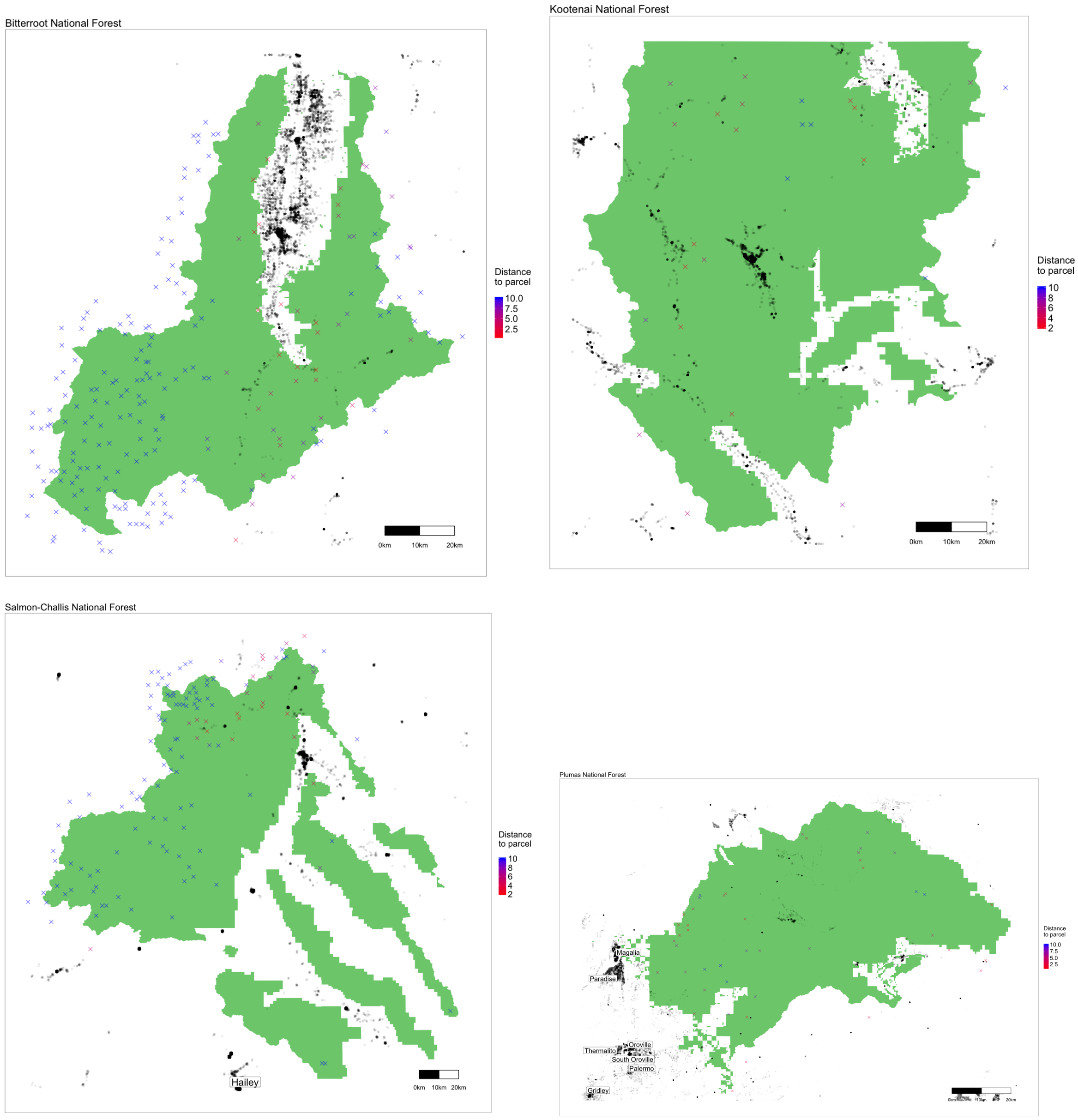


(h) "Timber" fuel model



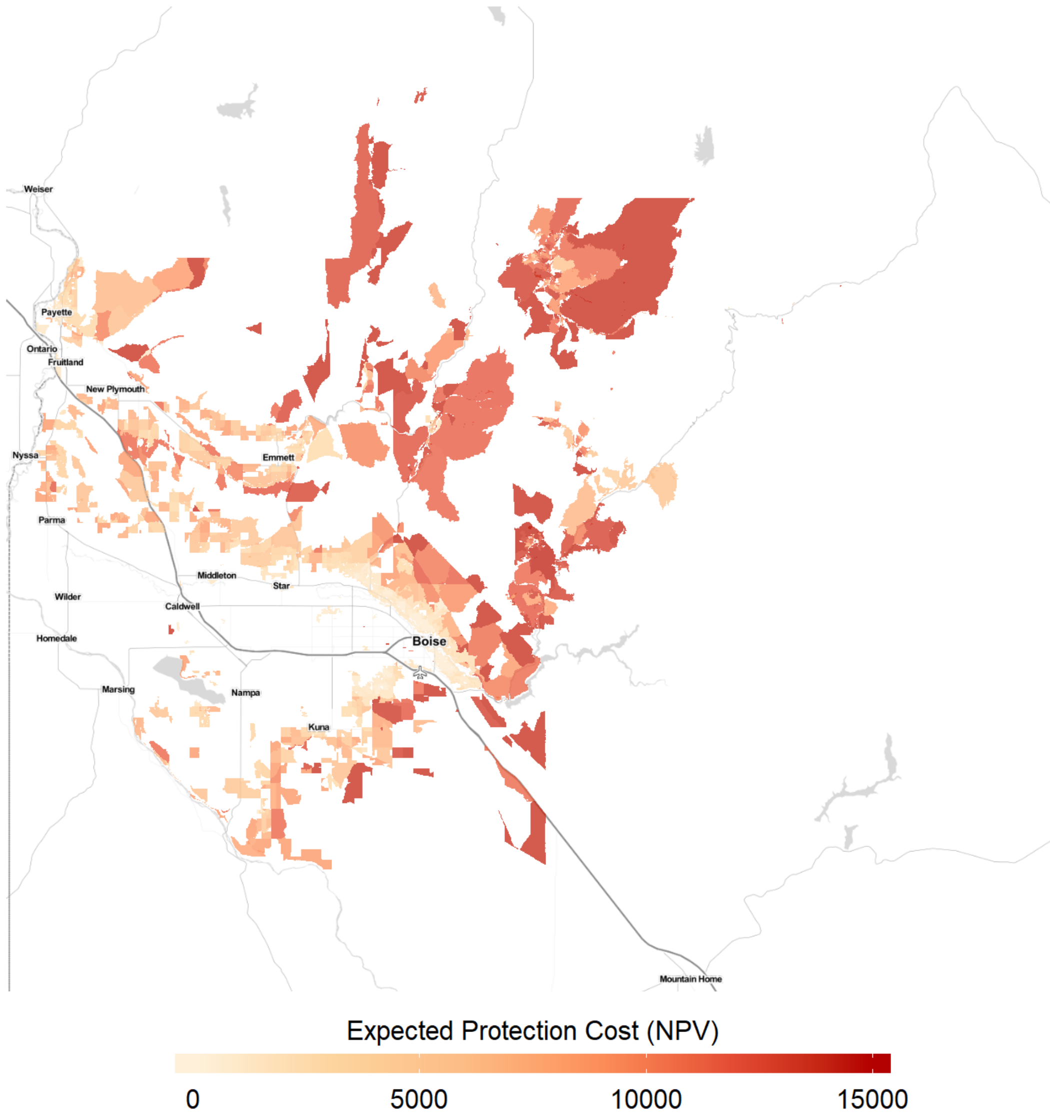
This figure shows covariate distributions for the US Forest Service fires analyzed in Tables 2 and 3. Panels (b), (c), and (d) report weather on the day of ignition. Wind speed is average wind speed from the reference weather station reported in NIFMID. Temperature and vapor pressure differential are mean daily values from PRISM. Terrain slope is the slope percentage, where 100 corresponds to a slope of 1 (i.e, a 45-degree line). "Timber" fuel models are National Fire Danger Rating System fuel models E, G, H, P, R, and U.

Appendix Figure 3: Additional National Forest Examples



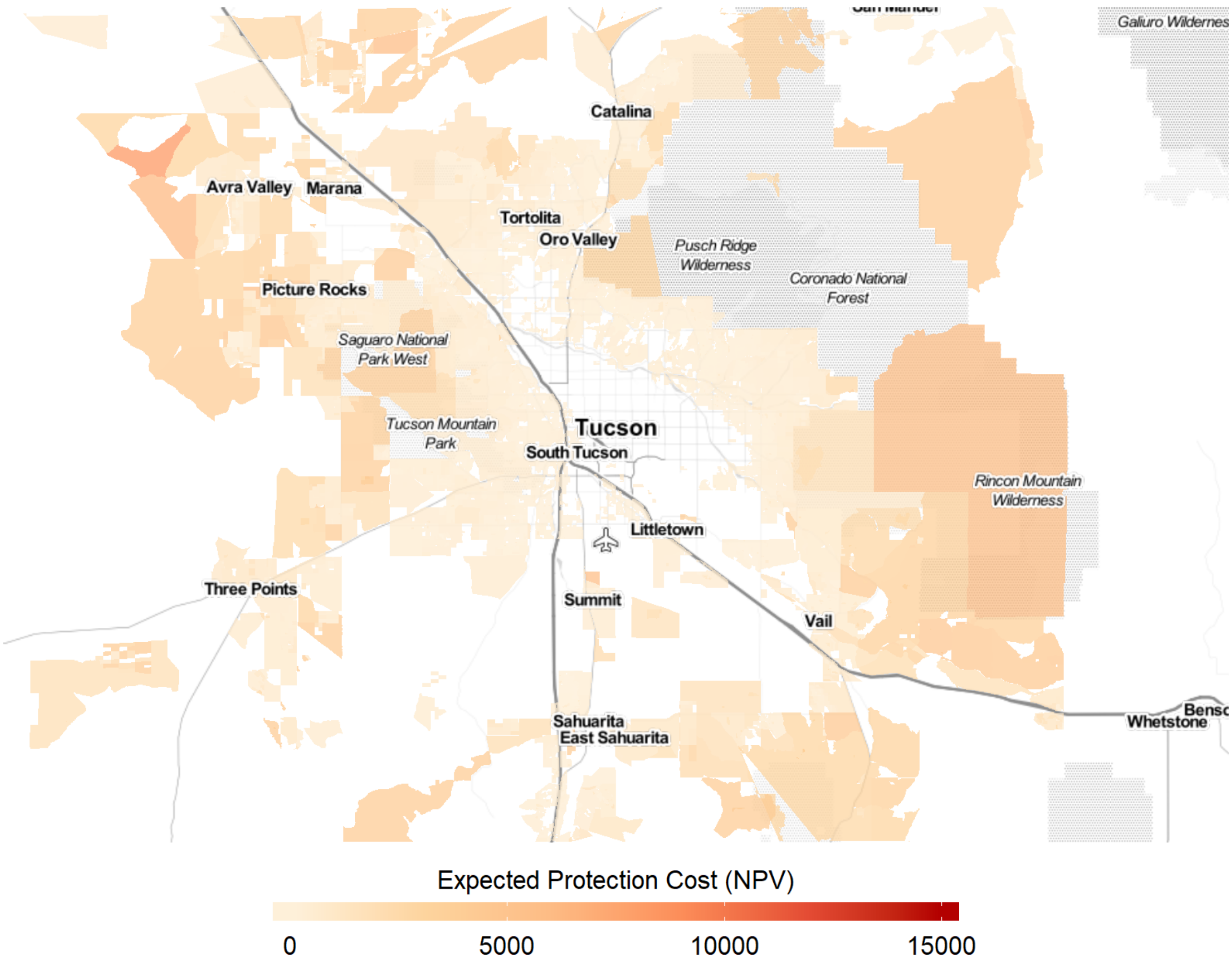
Each panel shows a single national forest area in green. The X's represent individual wildfires, colored according to the distance to the nearest home. Black dots indicate private homes. Clockwise from upper left, the forests are Bitterrroot National Forest (Montana), Kootenai National Forest (Montana), Salmon-Challis National Forest (Idaho), and Plumas National Forest (California).

Appendix Figure 4A: Local variation in Expected Cost, Additional Examples



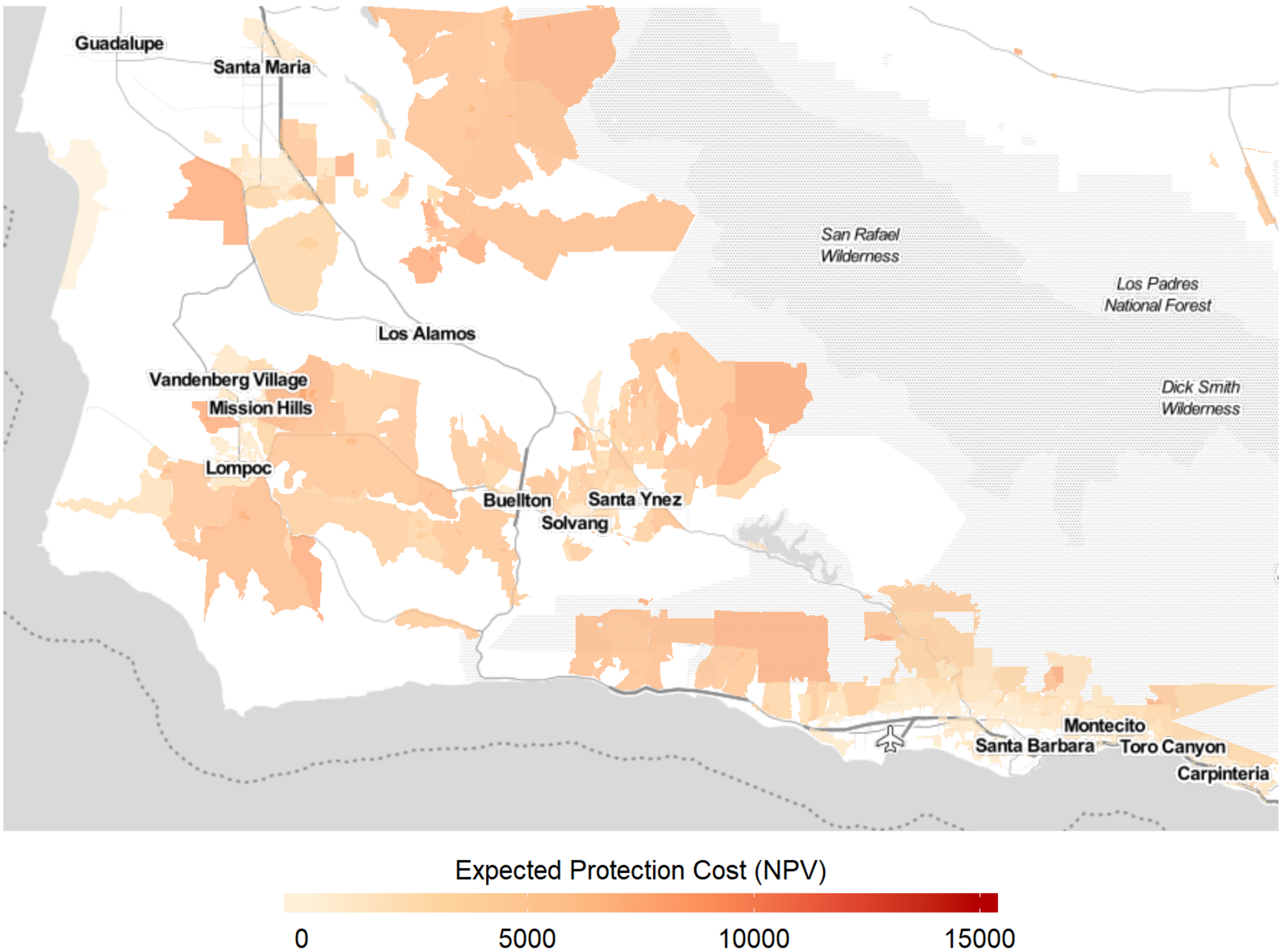
This map shows expected protection costs averaged by Census block for the Boise, Idaho area. The color scale indicates the average expected NPV of a home's protection cost, and is top-coded at $15,000.

Appendix Figure 4B: Local variation in Expected Cost, Additional Examples



This map shows expected protection costs averaged by Census block for the Tucson, Arizona area. The color scale indicates the average expected NPV of a home's protection cost, and is top-coded at $15,000.

Appendix Figure 4C: Local variation in Expected Cost, Additional Examples



This map shows expected protection costs averaged by Census block for the Santa Barbara, California area. The color scale indicates the average expected NPV of a home's protection cost, and is top-coded at $15,000.

1.1 Effect of Home Density on Fire Costs: Robustness checks

Appendix Table 1 shows additional robustness checks for the effects of the number of nearby homes on fire costs. Columns (1) through (5) show the same checks that we show in Table 2 for the effect of the nearest home on fire costs. Our results are robust to these various tests. The estimated effects of the other fire characteristics are also very similar to those in Table 2, as expected. Column (6) shows an additional specification that measures the stock of nearby homes by total transaction value, instead of number of homes. Results are similar.

Appendix Table 1: The Effect of Number or Value of Homes, Robustness Checks

	Number					Value
	(1)	(2)	(3)	(4)	(5)	(6)
Quintile Bins						
1	0.87***	0.87***	0.65**	0.89***	0.93***	0.86***
	(0.28)	(0.30)	(0.31)	(0.31)	(0.28)	(0.31)
2	1.41***	1.39***	1.34***	1.37***	1.29***	1.34***
	(0.34)	(0.36)	(0.38)	(0.36)	(0.38)	(0.38)
3	1.59***	1.53***	1.19***	1.29***	1.38***	1.55***
	(0.41)	(0.44)	(0.32)	(0.44)	(0.42)	(0.44)
4	1.71***	1.68***	1.50***	1.50***	1.92***	1.71***
	(0.34)	(0.36)	(0.32)	(0.39)	(0.38)	(0.36)
5	1.76***	1.79***	1.31***	1.62***	1.69**	1.93***
	(0.40)	(0.40)	(0.46)	(0.49)	(0.64)	(0.40)
Additional Controls						
Terrain Slope		0.008***	0.006	0.009**	0.007	0.008***
		(0.003)	(0.004)	(0.004)	(0.005)	(0.003)
Wind (mph)		0.026	0.008	0.015	0.008	0.026
		(0.019)	(0.035)	(0.028)	(0.039)	(0.019)
Vapor Pressure Differential		0.021*	0.013	0.010	0.023	0.021*
		(0.011)	(0.016)	(0.014)	(0.021)	(0.011)
South/southwest-facing		0.252*	0.340	0.312	0.115	0.251*
		(0.136)	(0.224)	(0.197)	(0.275)	(0.135)
Fuel Model FE		X	X	X		X
National Forest FE	X	X	X	X	X	X
Month-of-Year by State FE	X	X		X	X	X
Year by State FE	X	X		X	X	X
Week-of-Sample by State FE			X			
Lightning fires only				X		
Timber Fuels only					X	
N	2,069	2,069	1,365	1,437	1,018	2,069

Columns (1) through (5) reproduces Table 2 from the main text, using bins of the number of homes within 30 kilometers as the variables of interest. The bins are equal observation bins for fires with at least 1 nearby home (see Table 3 for bin ranges). The omitted category is fires with zero nearby homes. Column (6) shows an alternative specification that measures the stock of homes within 30 km by total transaction value. Again, bins are equal observation bins for fires with at least 1 nearby home, and the excluded category is fires with zero nearby homes.

Appendix Table 2 shows the effects of the number of nearby homes on fire costs using alternative radii around the ignition point to count the number of homes. Each table row shows coefficients for five equal-observation bins corresponding to the distribution of number of homes, conditional on any homes within the radius. The omitted category in each regression is fires with zero homes within the radius. For all three radii, there is a clear pattern of quick increases across the first two bins, and then roughly constant costs at higher numbers of homes. Note that direct comparisons of these coefficients across bins are difficult, since the comparison group of fires with zero threatened homes is systematically different across columns (e.g., in the 40 km column, all fires with zero homes are very remote by construction). Several other effects also presumably occur simultaneously as we widen the radius: since further-away homes have less effect on costs, these measures attenuate somewhat; however, because calculating density over a wider area reduces noise in our assessment of the number of threatened homes, there is another factor making these measurements more precise. Finally, note that the actual bin endpoints vary across models. The choice of radius is ultimately a somewhat arbitrary decision. Importantly, however, the obvious non-linear pattern of costs by number of homes exists for any radius.

Appendix Table 2: Costs by Number of Homes: Alternative Radii

	20 km		30 km		40 km	
Bin	Number of homes	Log Cost	Number of homes	Log Cost	Number of homes	Log Cost
0	0	0.00	0	0.00	0	0.00
1	1	0.65**	1	0.84***	1	0.86***
		(0.26)		(0.28)		(0.14)
2	36	1.11***	115	1.35***	300	1.74***
		(0.29)		(0.33)		(0.31)
3	185	1.20***	626	1.55***	1,336	1.67***
		(0.34)		(0.39)		(0.24)
4	859	1.05***	2,499	1.68***	4,983	1.85***
		(0.29)		(0.33)		(0.24)
5	3,257	1.32***	8,524	1.69***	15,529	1.66***
		(0.31)		(0.38)		(0.34)
Fires with Homes		1,806		2,082		2,233
Fires Without Homes		528		252		101

This table reproduces Table 3 from the main text using alternative radii. Each table row shows coefficients for five equal-observation bins corresponding to the distribution of number of homes, conditional on any homes within the radius. The omitted category in each regression is fires with zero homes within the radius. For all three radii, there is a clear pattern of quick increases across the first two bins, and then roughly constant costs at higher numbers of homes. Note that direct comparisons of these coefficients across bins are difficult, since the comparison group of fires with zero threatened homes is systematically different across columns (e.g., in the 40 km column, all fires with zero homes are very remote by construction). Several other effects also presumably occur simultaneously as we widen the radius: since further-away homes have less effect on costs, these measures attenuate somewhat; however, because calculating density over a wider area reduces noise in our assessment of the number of threatened homes, there is another factor making these measurements more precise. Finally, note that the actual bin endpoints vary across models. The choice of radius is ultimately a somewhat arbitrary decision. Importantly, however, the obvious non-linear pattern of costs by number of homes exists for any radius.

1.2 Effect of Homes on the Number of Fires

To evaluate whether the addition of new homes causes a larger number of fires (in addition to larger expenses on each fire that occurs), we take advantage of panel variation in home construction near each of the national forests in our dataset. We construct a year-by-national forest panel including 67 national forests and 20 years of fire experience. Because new homes are most likely to affect the number of ignitions in places with relatively low levels of development, we exclude national forests that had more than 150,000 homes within 30 kilometers of the national forest boundary in 1995 (this excludes the 20% of most densely-populated national forests).

We implement a variety of panel regression specifications. Our preferred statistical approach is a Poisson regression, since the number of fires in each national forest-year is a count variable with many zeros and a small number of other values.[20] The key identification challenge in this setting is to separate the effect of new home construction from other time-varying determinants of fire probability. Because homes are durable, the number of homes near each national forest increases monotonically across the sample. We adopt a variety of time trends and year fixed effects specifications to control as flexibly as possible for potential secular trends in the number of forests in each national forest caused by factors like climate change or annual drought cycles. Our results in this section should be interpreted with caution, since they rest on the somewhat strong assumption that, conditional on these controls, the trend in new home construction near each national forest is uncorrelated with other trends in fire occurrence.

Appendix Table 3 shows the results. All of these regressions include national forest fixed effects which remove the effect of time-invariant determinants of fire risk, such as local topography. Across specifications, new home development has a small positive effect on the number of fires each year. In Column (1), the estimated coefficient in the Poisson regression is 0.028. This implies that adding 1,000 new homes increases the annual number of fires in this national forest by 2.8%.[21] The average number of fires in each national forest-year is 1.7, so this implies that an additional 1,000 homes lead to 0.05 additional fires per year. Columns (2)–(5) include alternative polynomial time trends and find similar results. Column (6) instead includes year fixed effects, which allows for arbitrary annual trends at the West-wide level. Column (7) shows the same fixed effects specification in an OLS regression, for comparison to the Poisson results.

[20] We address the typical limitation of classic count regression, the restriction that the mean equal the variance for the estimated effects, by using a cluster-robust variance estimator which eliminates this problem.

[21] Expected changes in counts are calculated as $exp^{\beta} - 1$, where β is the Poisson regression coefficient.

Appendix Table 3: The Effect of Homes on the Number of Fires

	(1) Poisson	(2) Poisson	(3) Poisson	(4) Poisson	(5) Poisson	(6) Poisson	(7) OLS
Thousands of Homes	0.028***	0.035***	0.029***	0.037***	0.033***	0.030***	0.021*
	(0.005)	(0.007)	(0.008)	(0.008)	(0.007)	(0.008)	(0.011)
National Forest FE	X	X	X	X	X	X	X
Linear Time Trend		X					
Quadratic Time Trend			X				
Regional Linear Trends				X			
Regional Quadratic Trends					X		
Year Fixed Effects						X	X
N	1,060	1,060	1,060	1,060	1,060	1,060	1,060

This table reports the results of seven separate regressions. In each regression the dependent variable is the number of fires larger than 300 acres in each national forest-year. Columns (1)-(6) show results for several Poisson regression specifications, and Column (7) shows an OLS specification for comparison. The variable of interest is the number homes within 30 kilometers of the national forest boundary, in thousands. The table reports regression coefficients and standard errors, which are calculated using a cluster robust variance estimator at the national forest level. For the Poisson specifications, the coefficients can be converted to expected percentage changes in the number of large fires using calculation $e^{\beta} - 1$. See text for details. The mean number of fires in each national forest-year is 1.7. "Regional Linear Trends" and "Regional Quadratic Trends" indicate that the regression includes separate polynomial time trends for each of the five forest service regions included in the sample area.

2 Construction of the Dataset

2.1 Fires data

The initial dataset includes 2,613 fires larger than 300 acres during 1995–2014. We restrict to fires managed by the US Forest Service. For every fire, we calculate the distance from the ignition point to the nearest home, using the homes dataset described below.

2.2 Homes data

2.2.1 Sample restrictions

The initial dataset includes nearly 17,701,699 million single-family homes in the 11 western states.[22] We restrict the sample to include 8,117,482 homes at risk from wildland fires due to the presence of wildland vegetation, based on geographic classifications in Radeloff et al. (2005). The vegetation categories we include are high density interface, high density intermix, medium density interface, medium density intermix, low density interface, low density intermix, very low density vegetated, and uninhabited vegetated. We exclude homes in areas without wildland vegetation, including high density no vegetation, medium density no vegetation, low density no vegetation, very low density no vegetation, and uninhabited no vegetation. We further restrict the sample to the 6,452,290 homes that are less than 40km from a national forest boundary.

2.2.2 Assigning fire costs to homes

We calculate the expenditures on each fire that were due to the presence of homes, using the statistical model in Section 5.1. First, we construct an estimate of the residual fire cost due to the presence of any single-family homes near the point of ignition to avoid capturing firefighting costs that would have been occurred regardless of nearby habitation. We use the estimate of the proportional increase in firefighting costs due to the nearest home given by our results in section 5.1 minus one divided by that estimate. Mathematically, if r_d is the ratio of costs estimated for a fire with nearest parcel at distance d over the cost for a fire with $d > 40$km (the distance above which firefighting costs no longer decrease), then our estimate of the residual fire cost is $C = \text{Fire cost} \times \frac{r_d - 1}{r_d}$, which is by construction between zero and the total cost of

[22] Arizona, California, Colorado, Idaho, Montana, New Mexico, Nevada, Oregon, Utah, Washington, Wyoming.

the fire. For example way, if the ignition point of a fire is between 12 and 16km from the nearest home, then the estimated coefficient from Equation 1 is 5.3, meaning that the relative cost is 5.3 times larger than a fire with the nearest home more than 40 km away. To obtain the residual fire cost C we multiply the total cost times the ratio $\frac{5.3-1}{5.3} \approx 0.81$ to get the residual cost. Using this method, we assign around 72% of total firefighting costs to the single-family homes in our dataset.

We then divide those residual fire costs, or home protection expenditures, over all homes in our sample that are within 40 km of the ignition point. In assigning these costs to homes, we place more weight on homes located near the ignition point. To do so we make use of two approaches, an inverse-distance weighting (IDW) algorithm and an empirical estimate based on the results in section 5.1. For the IDW algorithm, houses within 40km are assigned a weight of $\frac{1}{d}$, weights are normalized to one within each fire, and home protection expenditures by fire are divided using the normalized weights. Our second and preferred approach is identical except that the weights assigned to each fire-parcel combination are the estimated coefficients from Equation 1 for distance between the igition point and the parcel location, normalized to sum to one for each fire.

2.3 Additional Data

We assign both a categorical and a continuous measure of wildfire hazard potential to nearly every parcel[23] in our dataset using data from Dillon (2015).

Data on the area and number of housing units in each Census Block come from 2010 Census TIGER/Line data files. To smooth idiosyncratic variation in housing density due to small block areas, we aggregate this information to the Census tract level and then calculate the number of housing units per square meter. There are 15,599 Census tracts in the 11 western states. We merge this housing density data to the homes dataset using Census geographies included in the homes dataset. We successfully merge 99.7% of homes to tract-level housing density. For the 0.3% of homes that we fail to match at the tract level, we use county-level average housing density.

[23] 156 parcels lie outside the boundaries given by the spatial data in Dillon (2015).

3 Model Appendix

3.1 Optimal Level of Defensive Expenditures

Conditional on a disaster occurring, the optimal level of defensive expenditures minimizes the sum of total expected damages and defensive expenditures:

$$f^*(N) = \arg\min_f \, [f + NH(f)]$$

Taking first order conditions and solving yields that $-H'(f^*) = \frac{1}{N}$. By Assumption 1 from the main text, $-H'(f)$ is decreasing in f. By implication, f^* is increasing in N.

3.2 Per-Household Disaster Costs are Decreasing in Population

Each household's proportional share of expected disaster costs is $\pi_i = \phi_i[\frac{f^*(N_i)}{N_i} + H_i(f^*(N_i))]$. Note that $f^*(N_i)$ is chosen optimally to minimize total disaster costs in the location conditional on a disaster, which are equal to $\frac{N_i}{\phi_i} * \pi_i$. Thus, $f^*(N_i)$ also minimizes per-household disaster costs. The envelope theorem implies that the total derivative of $\pi_i(N, f^*(N))$ with respect to N is equal to the partial derivative of $\pi_i(N, f^*(N))$ with respect to N, which is $-\frac{f^*(N)}{N^2}$. This is negative.