NBER WORKING PAPER SERIES

HOW ARE INSURANCE MARKETS ADAPTING TO CLIMATE CHANGE? RISK CLASSIFICATION AND PRICING IN THE MARKET FOR HOMEOWNERS INSURANCE

Judson Boomhower
Meredith Fowlie
Jacob Gellman
Andrew Plantinga

Working Paper 32625
http://www.nber.org/papers/w32625

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2024, Revised February 2025

This paper was previously circulated under the title "How Are Insurance Markets Adapting to Climate Change? Risk Selection and Regulation in the Market for Homeowners Insurance". We thank seminar participants at the NBER Summer Institute, NBER Insurance Working Group, the 2023 ASSA meetings, the 2024 AERE Summer Conference, TWEEDS, Arizona State University, University of Chicago, Johns Hopkins University, Harvard University, MIT, Indiana University, Resources for the Future, the New York Federal Reserve, the San Francisco Federal Reserve, Stanford University, University of Alaska Anchorage, UC Berkeley, UC Davis, UC Santa Barbara, University of Southern California, University of Wisconsin, University of Illinois, University of Calgary, Georgia State University, and Oregon State University. We thank Josh Aarons, Kaleb Javier, and Kendra Marcoux for outstanding research assistance. We thank Benjamin Collier, Chris Costello, Mark Duggan, Penny Liao, Philip Mulder, and Joel Sobel for helpful discussions. This work received generous financial support from the University of California Office of the President Laboratory Fees Program (LFR-20-652467). Boomhower gratefully acknowledges funding via an Andrew Carnegie Fellowship from the Carnegie Corporation of New York. CoreLogic, Inc. is a source of data for this analysis. CoreLogic was not involved in the preparation of any aspect of the report. Any conclusions/interpretations of the CoreLogic data are those of the authors of the report, and not CoreLogic. CoreLogic is not responsible for any use of, nor any decisions based on or in reliance on, the CoreLogic data used in the report. CoreLogic does not represent or warrant that the CoreLogic data is complete or free from error and does not assume, and expressly disclaims, any liability to any person or entity for any loss or damage caused or resulting in whole or in part by any reliance on the CoreLogic data including, but not limited to, any loss or damage caused or resulting in whole or in part by errors or omissions in the CoreLogic data, whether such errors or omissions result from negligence, accident, or other cause. CoreLogic® is the registered trademark of CoreLogic, Inc. and/or its subsidiaries or affiliates. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2024 by Judson Boomhower, Meredith Fowlie, Jacob Gellman, and Andrew Plantinga. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

How Are Insurance Markets Adapting to Climate Change? Risk Classification and Pricing in the Market for Homeowners Insurance
Judson Boomhower, Meredith Fowlie, Jacob Gellman, and Andrew Plantinga
NBER Working Paper No. 32625
June 2024, Revised February 2025
JEL No. D82, G22, Q54

ABSTRACT

Property insurance is critical for reducing household exposure to severe weather risks and aiding in recovery when disaster strikes. As climate risks escalate, insurers are investing in more sophisticated methods to classify and price complex disaster risks. We use proprietary data on parcel-level wildfire risk, together with insurers' regulatory filings, to investigate how wildfire risk insurance is being priced and provided in a large market for homeowners insurance. We document striking variation in insurers' risk classification and pricing strategies. Firms that rely on coarser measures of wildfire risk are exposed to potentially severe adverse selection as a result of their information disadvantage relative to insurers with richer models. Consistent with this fact, firms relying on simpler pricing methods charge relatively high prices in high-risk market segments – or choose not to serve these areas at all. A theoretical model of a market for natural hazard insurance that incorporates both asymmetric risk classification and market regulation helps rationalize the empirical patterns we document. Our results highlight how the winner's curse can increase prices and limit participation in insurance markets for large, hard-to-model risks.

Judson Boomhower
Department of Economics
University of California at San Diego
9500 Gilman Drive #0508
La Jolla, CA 92093
and NBER
jboomhower@ucsd.edu

Meredith Fowlie
Department of Agricultural &
Resource Economics
University of California, Berkeley
301 Giannini Hall
Berkeley, CA 94720-3310
and NBER
fowlie@berkeley.edu

Jacob Gellman
Department of Economics
University of Alaska Anchorage
3211 Providence Drive
Rasmuson Hall 205K
Anchorage, AK 99508
jgellman@alaska.edu

Andrew Plantinga
Bren School of Environmental Science
and Management
University of California
Rm 3424
Santa Barbara, CA 93106
plantinga@bren.ucsb.edu

How Are Insurance Markets Adapting to Climate Change? Risk Classification and Pricing in the Market for Homeowners Insurance

Judson Boomhower Meredith Fowlie
Jacob Gellman Andrew J. Plantinga*

January 2025

Property insurance is critical for reducing household exposure to severe weather risks and aiding in recovery when disaster strikes. As climate risks escalate, insurers are investing in more sophisticated methods to classify and price complex disaster risks. We use proprietary data on parcel-level wildfire risk, together with insurers' regulatory filings, to investigate how wildfire risk insurance is being priced and provided in a large market for homeowners insurance. We document striking variation in insurers' risk classification and pricing strategies. Firms that rely on coarser measures of wildfire risk are exposed to potentially severe adverse selection as a result of their information disadvantage relative to insurers with richer models. Consistent with this fact, firms relying on simpler pricing methods charge relatively high prices in high-risk market segments – or choose not to serve these areas at all. A theoretical model of a market for natural hazard insurance that incorporates both asymmetric risk classification and market regulation helps rationalize the empirical patterns we document. Our results highlight how the winner's curse can increase prices and limit participation in insurance markets for large, hard-to-model risks.

Keywords: Insurance, adverse selection, natural disasters, climate change, wildfire
JEL Codes: D82, G22, Q54

---

*(Boomhower) Department of Economics, University of California San Diego and NBER; jboomhower@ucsd.edu. (Fowlie) Department of Agricultural and Resource Economics, University of California Berkeley and NBER. (Gellman) Department of Economics, University of Alaska Anchorage. (Plantinga) Bren School of Environmental Science and Management, University of California Santa Barbara. This paper was previously circulated under the title "How Are Insurance Markets Adapting to Climate Change? Risk Selection and Regulation in the Market for Homeowners Insurance". We thank seminar participants at the NBER Summer Institute, NBER Insurance Working Group, the 2023 ASSA meetings, the 2024 AERE Summer Conference, TWEEDS, Arizona State University, University of Chicago, Johns Hopkins University, Harvard University, MIT, Indiana University, Resources for the Future, the New York Federal Reserve, the San Francisco Federal Reserve, Stanford University, University of Alaska Anchorage, UC Berkeley, UC Davis, UC Santa Barbara, University of Southern California, University of Wisconsin, University of Illinois, University of Calgary, Georgia State University, and Oregon State University. We thank Josh Aarons, Kaleb Javier, and Kendra Marcoux for outstanding research assistance. We thank Benjamin Collier, Chris Costello, Mark Duggan, Penny Liao, Philip Mulder, and Joel Sobel for helpful discussions. This work received generous financial support from the University of California Office of the President Laboratory Fees Program (LFR-20-652467). Boomhower gratefully acknowledges funding via an Andrew Carnegie Fellowship from the Carnegie Corporation of New York. CoreLogic, Inc. is a source of data for this analysis. CoreLogic was not involved in the preparation of any aspect of the report. Any conclusions/interpretations of the CoreLogic data are those of the authors of the report, and not CoreLogic. CoreLogic is not responsible for any use of, nor any decisions based on or in reliance on, the CoreLogic data used in the report. CoreLogic does not represent or warrant that the CoreLogic data is complete or free from error and does not assume, and expressly disclaims, any liability to any person or entity for any loss or damage caused or resulting in whole or in part by any reliance on the CoreLogic data including, but not limited to, any loss or damage caused or resulting in whole or in part by errors or omissions in the CoreLogic data, whether such errors or omissions result from negligence, accident, or other cause. CoreLogic® is the registered trademark of CoreLogic, Inc. and/or its subsidiaries or affiliates.

1 Introduction

As the climate changes, natural disasters are increasing in frequency and intensity (Summers et al. 2022). In the United States, losses from natural disasters now exceed $120 billion annually.[1] Property insurance markets are critical to help households and firms adapt to these risks, providing financial resilience when disaster strikes (Shi and Moser 2021). However, there are mounting concerns that property insurance markets are not equipped to manage escalating climate risk (U.S. Department of the Treasury 2023).

Across the United States, property insurance premiums have been increasing faster than inflation (Keys and Mulder 2024). Insurers have also been limiting underwriting or exiting some market segments altogether. The public conversation about these trends has focused on two important explanations. The first is cost-based. More frequent and more severe losses imply higher expected costs for insurers. Moreover, the correlated nature of disaster claims further raises insurers' costs through the need for surplus capital or reinsurance to protect the firm's solvency (Jaffee and Russell 1997). The second is regulation, which has been argued to constrain firms' ability to set premiums at levels that keep pace with increasing climate risk (Oh, Sen, and Tenekedjieva 2024).

This paper explores a third factor at the heart of insurers' response to climate risk. Quantifying and classifying catastrophic weather risk is a complex exercise. Unlike more predictable risks such as health conditions or auto accidents, damages from extreme weather events are infrequent, spatially correlated, and often catastrophic. As climate-driven loss risk grows, insurers must weigh the costs of adopting more advanced approaches to modeling these risks against the benefits of more accurate, granular pricing. Better risk information helps firms assess loss potential in risky areas. In addition, higher fidelity risk premiums can provide better incentives for self-protection when premiums align with true risk. Paradoxically, however, innovation in proprietary pricing models also has the potential to reduce participation in competitive insurance markets. If an insurer finds it is offering a customer a lower price than its competitors, this could indicate that competitors have better information about climate risk exposure. These information asymmetries can expose firms to potentially severe adverse selection, analogous to the winner's curse in the literature on common-value auctions (Engelbrecht-Wiggans, Milgrom, and Weber 1983; Hendricks and Porter 1988; Milgrom and Weber 1982).

To understand proprietary investments in risk modeling and the economic implications of associated adverse selection, we focus on a particular form of climate risk. Catastrophic wildfires are the fastest-growing source of climate-related damages in the United States.[2] Over the past two decades, wildfires in the United States have quadrupled in size and tripled in frequency (Iglesias, Balch, and Travis 2022). Despite large economic implications, wildfire risk remains under-explored in the

1. NOAA National Centers for Environmental Information (NCEI). (2024). U.S. Billion-Dollar Weather and Climate Disasters. https://www.ncei.noaa.gov/access/billions/.

2. Swiss Reinsurance. (2023). Continued high losses from natural catastrophes in 2022. https://www.swissre.com/institute/research/sigma-research/sigma-2023-01/5-charts-losses-natural-catastrophes.html.

economics literature addressing insurance market pricing, underwriting, and competition.

Our empirical setting is California, home to an estimated 4.6 million properties with high wildfire risk exposure.[3] Since 2017, the state has seen a significant increase in average annual wildfire losses as compared to previous decades.[4] The wildfire seasons of 2017 and 2018 were particularly devastating, raising concerns about the insurability of catastrophic wildfire risk (Cignarale et al. 2019). We leverage a feature of California's insurance regulatory regime that subjects all insurance premium changes to prior regulatory approval. This setting allows us to access detailed, insurer-specific information about risk classification filed with the California regulator. We use these public rate cases, together with the proprietary information that firms use to assess wildfire risk exposure, to reconstruct parcel-level wildfire risk classification and pricing strategies for six large property insurers in this market.

We begin with a descriptive analysis of how property insurers in California have been adapting and responding to escalating wildfire risk. In the years following the catastrophic 2017 and 2018 wildfire seasons, premiums have risen, the rate of policy cancellations in high-hazard areas has escalated, and participation in the California FAIR plan, a quasi-private insurer of last resort, has increased rapidly. Over this same period, many insurers have developed more sophisticated wildfire risk classification strategies. The pace of innovation, and the adoption of more granular risk information, has been uneven across firms. Asymmetries in wildfire risk information across insurers create the potential for adverse selection because firms with more granular risk information can identify lower-risk properties and offer them lower insurance premiums.

To assess the potential cost implications of wildfire risk information asymmetries, we develop a model of cost-based insurance pricing. Canonical work on insurance markets (e.g., Jaffee and Russell 1997; Kreps 1990; Kunreuther 1996; Stone 1973) provides a normative prescription for "fair and adequate" pricing in markets for catastrophic risk insurance that are characterized by perfect competition and symmetric information.[5] Building on these theoretical foundations, we estimate the cost of providing wildfire risk insurance in terms of expected losses, operating costs, and a loading factor that reflects the costs of protecting insurer solvency through capital surplus or reinsurance. We show how asymmetries in risk classification information can increase costs for insurers operating at a risk information disadvantage.

For each insurer in the dataset, we identify the relationship between offered insurance premiums and assessed wildfire risk, conditioning on other insurance premium determinants. We compare these empirical risk pricing gradients against average cost benchmarks implied by the insurance cost model. Pricing patterns in the data are consistent with strategies that incorporate winner's curse adjustments. For the insurer with the most granular risk classification strategy, risk segment aver-

3. This number is projected to increase to 5.5 million, or 7.6 percent of all properties, by 2050 (First Street Foundation 2022).

4. The average annual loss in California from 2009 to 2018 was almost $1 billion, compared to $0.40 billion from 1999 to 2008, $0.19 billion from 1989 to 1998, and $0.03 billion from 1979 to 1988 (Buechi et al. 2021).

5. For a summary of these issues, an interested reader could refer to Kunreuther and Michel-Kerjan (2011).

age insurance prices increase commensurately with wildfire risk-related costs. In contrast, among insurers that use relatively coarse risk classification, average prices within higher-risk segments exceed segment-average levels of assessed wildfire risk exposure.

To definitively quantify the cost implications of adverse selection in this market, we would need to know which properties are insured by each firm. We cannot access this proprietary information. However, we are uniquely positioned to assess the size of the winner's curse in a pricing game where firms differ in the sophistication of their risk pricing strategies. In a stylized oligopoly model of price competition, we show that the expected losses of households who purchase from firms using coarser risk information are substantially higher than those purchasing from the firm with the most granular pricing. These results hold when accounting for consumer switching costs and alternative market structures.

Motivated by these findings, we develop an equilibrium model to analyze interactions between information asymmetries and regulatory constraints that empirically appear to bind in this property insurance market. In the model, the value of adopting more sophisticated information and insurance pricing are determined endogenously. The model helps to rationalize the incomplete adoption of more sophisticated risk measurement. Resulting information asymmetries expose less-informed firms to adverse selection. The model further predicts that the more-informed firm uses a more granular risk classification strategy to capture the lower-risk customers within a risk segment, while the relatively less-informed firms must raise their prices to avoid selling unprofitable policies to high-risk customers. The overall effect of this asymmetric risk information is an increase in the average price of insurance.

Finally, we use the model to explore the likely implications of existing regulations and some recent policy reforms in California. The model elucidates how information asymmetries interact with regulations that limit insurers' ability to raise prices. Binding regulatory constraints induce relatively uninformed firms to exit high-risk segments, reducing insurance availability and increasing market concentration in these segments. We show how policies that improve access to better risk information for all insurers can improve both the affordability and the availability of insurance.

Our paper contributes to a growing literature on natural disaster insurance. Early work by Kunreuther (1996) and Jaffee and Russell (1997) examined the causes of private market failures in catastrophe insurance and the conditions that must be established to make these markets viable. Related work has investigated the response of insurance markets to catastrophic events (Born and Viscusi 2006; Klein and Kleindorfer 1999), the challenges of insuring catastrophic risks (Kousky and Cooke 2009; Marcoux and Wagner 2024), and the pricing of catastrophe insurance (Gourevitch, Kousky, and Liao 2023). Our work is most closely related to Keys and Mulder (2024), who study the relationship between insurance premiums and disaster risk across zip codes in the United States. We examine this relationship from a supply-side perspective, using firm-level price schedules and parcel-level measures of wildfire risk. The granular property-level data allow us to estimate highly detailed risk price gradients and to observe differences in pricing across insurers. We demonstrate the importance of accounting for insurers' risk classification practices when interpreting the empirical relationships between insurance prices and assessed wildfire risk.

Our analysis also builds on a classic literature that investigates costly risk classification in competitive insurance markets. Seminal work in this area has explored insurers' endogenous investments in risk classification (see, for example, Borenstein 1989), efficiency implications (e.g., Abraham 1985; Crocker and Snow 1986), and concerns around how risk classification practices redistribute risk (Chamberlin 1985). New innovations in machine learning, catastrophe modeling, satellite observation, and "big data" are transforming the practice of risk classification in insurance markets (Einav, Finkelstein, and Mahoney 2021; Jin and Vasserman 2021). These more sophisticated approaches to risk classification and insurance pricing raise new questions about efficiency and fairness (Eling, Gemmo, and Guxha 2024). We document how uneven adoption of this information can reduce the availability and affordability of insurance.

A closely related line of inquiry explores the relationship between risk classification and adverse selection. Much of this literature has considered selection market contexts in which more granular risk classification can reduce information asymmetries and mitigate adverse selection problems (see, for example, Einav and Finkelstein 2011; Handel, Hendel, and Whinston 2015; Hoy 1982). More recently, economists have considered contexts in which risk classification based on proprietary information held by an innovative insurer *creates* adverse selection against insurers that do not have this information. Einav, Jenkins, and Levin (2012) show how proprietary health information can confer strategic "cream skimming" benefits, leaving relatively less-informed insurers with an adversely selected pool of customers. Cather (2018) demonstrates how innovations in auto insurance risk classification can lead to adverse selection against insurers that are slow to adopt these innovations. We show how asymmetries in risk classification practices can lead to higher premiums and increased concentration in a market where the purchase of insurance is effectively mandatory and prices are subject to binding regulatory constraints.

Lastly, we contribute to the literature on insurance market regulation. Past work has explored supply-side implications of capital requirements and dynamic pricing regulations across a range of insurance market contexts (see, for example, Aizawa and Ko 2023; Ge 2022; Koijen and Yogo 2015). Our paper is closely related to work investigating how private property insurers respond to economic regulation (Born and Klimaszewski-Blettner 2013; Oh, Sen, and Tenekedjieva 2024; Taylor, Turland, and Weill 2023). We leverage parcel-level data to show how economic regulations can interact with asymmetries in risk information in economically significant ways.

The remainder of this paper is organized as follows. Section 2 provides background on wildfire risk and homeowners insurance regulation and summarizes recent trends in insurance market outcomes. Section 3 introduces the new dataset of insurance rate filings and proprietary wildfire risk measures, exploring key descriptive facts that emerge from the data. Section 4 introduces a theoretical cost-based benchmark for the fair price of wildfire risk, which guides the empirical analysis. Section 5 estimates the empirical relationship between offered premiums and assessed wildfire risk. Section 6 presents additional evidence on the roles of adverse selection and regulation. Section 7 introduces an equilibrium model of a regulated insurance market where access to more sophisticated risk information is costly. Section 8 concludes.

2 Institutional Background and Publicly-Available Descriptive Data

This section describes the operation of homeowners insurance markets in the United States, with a particular focus on California. Sections 2.1 through 2.3 review the basics of homeowners insurance contracts, trends in wildfire risk and risk pricing, and insurance market regulation. Section 2.4 uses publicly available data to summarize aggregate outcomes in the homeowners insurance market.

2.1 Homeowners insurance in the United States

We study the market for homeowners (HO) multi-peril insurance. Under standard HO insurance contracts, wildfire losses are bundled with other perils. HO multi-peril premiums exceed $125 billion annually in the United States, with California representing 9.1 percent of the market.[6] Multi-peril HO policy terms typically last for one year. Holders receive annual renewal statements that include information about any changes to rates or policy terms. Contracts are automatically renewed unless they are canceled by insurers or homeowners. Most mortgage lenders require multi-peril insurance coverage as a mortgage precondition. It is estimated that 88 percent of U.S. homeowners purchase and hold HO insurance.[7]

2.2 Wildfire risk

Wildfire risk is escalating across the western United States. In California, annual increases in projected wildfire damages exceed 1 percent in some of the highest-hazard zip codes (Dixon, Tsang, and Fitts 2018). Multiple factors explain this trend, including climate change, population growth in the wildland-urban interface, and a history of aggressive fire suppression practices that have increased fuel loads.

As wildfire damages increase and become more salient, more resources are being allocated to the modeling and classification of wildfire risk. Insurance companies determine customer-level prices through a process of risk classification – i.e. the assignment of individuals to groups with similar expected damages. Traditionally, insurers have assessed wildfire risk very coarsely, for example at the zip code level. More recently, insurers have begun to incorporate third-party models and tools.

6. NAIC. (2023). 2022 Market Share Reports For Property/Casualty Groups and Companies by State and Countrywide. https://content.naic.org/sites/default/files/publication-msr-pb-property-casualty.pdf.

7. If a homeowner neglects to purchase and hold HO insurance, their lender will purchase "force-placed insurance" which is typically more expensive and only covers the bank's interest in the house. See: Insurance Information Institute. (2023). Homeowners Perception of Weather Risks 2023Q2 Consumer Survey. https://www.iii.org/sites/default/files/docs/pdf/2023_q2_ho_perception_of_weather_risks.pdf.

Catastrophe (or "CAT") models can be used to simulate probability distributions of dollar-denominated insured losses for an individual property or an insurer's entire book of business. A number of firms have built proprietary CAT models for wildfire which insurers can use to classify and price wildfire risk. CAT modeling has significantly improved insurers' ability to assess wildfire risk. However, model-based risk metrics are inherently uncertain due to the complexity of wildfire risk.

2.3 Insurance market regulation

State regulators exercise considerable authority over insurers' entry, exit, underwriting, pricing, and claims settlement choices. These regulations are rationalized, in part, by potential market failures in an unregulated market. If customers have limited attention or lack the sophistication to understand and compare complex contracts, pricing terms, or insurer solvency, there may be potential for information-related market failures or agency problems. Unequal bargaining power between homeowners and insurance companies could also result in HO premiums that are higher than necessary.[8] Although setting premiums above what it costs to provide insurance does not generate deadweight loss in markets where the purchase of insurance is effectively mandatory, this practice negatively affects homeowners. A pursuit of "fair" pricing motivates regulatory oversight of firms' underwriting practices across the United States.

In California, insurance regulation is the purview of the California Department of Insurance (CDI). CDI's stated objective is to promote solvency, affordability, and availability of insurance. Under the provisions of Proposition 103, CDI is required to review and approve rates for most property and casualty lines of insurance to ensure that they are "fair" (i.e., not excessive, inadequate, or unfairly discriminatory).[9] Requests to increase premiums across a book of business by more than 6.9 percent are more likely to be subject to costly public rate hearings. Historically, California insurers have been somewhat restricted in their use of CAT modeling. Under the current regulatory regime, insurers can use CAT models to classify wildfire risk and price differentially on the basis of assessed risk exposure. However, insurers have not been permitted to use CAT modeling to justify an overall rate of increase in earned premiums. Instead, they must appeal to the historical record of their past catastrophe claims when requesting a rate increase.[10]

2.4 Publicly-available administrative data on market outcomes

We use zip code-level data collected by the California Department of Insurance (CDI) to summarize trends in homeowners insurance premiums, admitted market participation, and policy cancellations

8. Insurance contracts are considered "contracts of adhesion" in which the consumer must either accept the terms of the policy or reject the terms and accept similar terms from another company.

9. Rate regulations apply to "admitted market" insurers, which comprise 98 percent of the California insurance market (Dixon, Tsang, and Fitts 2018). "Surplus lines" are not subject to the same regulations.

10. See California Code of Regulations Title 10, § 2644.5 - Catastrophe Adjustment: "The catastrophic losses for any one accident year in the recorded period are replaced by a loading based on a multi-year, long term average of catastrophe claims. The number of years over which the average shall be calculated shall be at least 20 years for homeowners multiple peril fire." Recently proposed reforms seek to update these regulations in ways that facilitate increased use of CAT modeling, subject to regulatory oversight and underwriting requirements.

over time. Beginning in 2018, insurers writing more than $10 million in premiums were required to report zip code-level information about the assessed wildfire risk exposure of the properties they insure. This includes information about the distribution of insured parcels across wildfire risk categories. We use these data to coarsely classify zip codes into wildfire risk quantiles, a process which is detailed in Appendix A. These data reveal several trends.

2.4.1 Insurance premiums are rising

Figure 1 summarizes zip code-level average increases in HO insurance premiums over time, in 2020 dollars. Across all wildfire risk quantiles, real premiums increased noticeably after the destructive 2017 and 2018 wildfire seasons. Statewide, premiums rose 23.3 percent from 2017 to 2022.

2.4.2 Insurance availability is declining

The second panel of Figure 1 shows that the size of the admitted market had been increasing in all but the highest-hazard zip codes over the period 2009 to 2016. Policy counts fell across all categories during the 2017 and 2018 fire seasons; these reductions were particularly significant in the highest-hazard zip codes. In more recent years the decrease in admitted market policies in the highest-hazard zip codes has accelerated. The third graphic in Figure 1 tracks participation in the California FAIR Plan, which provides basic backstop coverage for properties that cannot find coverage in the admitted market.[11] Growth in FAIR Plan participation has been increasing since 2018, particularly in the highest-hazard areas.

2.4.3 Rates of consumer switching are low

The bottom panels of Figure 1 summarize trends in insurance policy cancellations, i.e. non-renewals. These data are only available for 2015 through 2021. Prior to 2019, insurer-initiated cancellations averaged around 2 to 3 percent. Since 2019, insurers have been canceling policies at higher rates in high-hazard areas.[12] The right-hand panel shows customer-initiated policy cancellations, which have historically been around 8 percent. These switching rates are low given the significant variation in premiums we document below.[13]

2.4.4 Economic regulation appears binding

Figure 2 summarizes 636 requested rate increases filed with CDI from 2008 to 2023 for owner-occupied homeowners' insurance (HO-3) policies. In California, the regulatory filings that insurers

11. The FAIR Plan is instituted at a state level but is backed by admitted market insurers. Insurers divide the profits and expenses associated with the FAIR plan according to their statewide market share.

12. This increase coincides with the introduction of Senate Bill 824, which imposed a one-year moratorium on insurance companies canceling insurance policies in or adjacent to wildfire perimeters.

13. One model of consumer behavior that would be consistent with these trends would be if most households consider HO insurance choices only when purchasing a home. Rates of consumer switching in California align with median homeowner tenure, which ranges from 11 to 13 years. See: Anderson, Dana. (2022). The Typical U.S. Home Changes Hands Every 13.2 Years. Redfin. https://www.redfin.com/news/2021-homeowner-tenure.

submit to request rate changes are all in the public domain.[14] These rate increase requests are visibly bunched at 6.9 percent, the threshold beyond which insurers are more likely to face costly public rate hearings. This suggests that pricing regulations have been a limiting factor in California. More recently, firms have been requesting much larger increases and facing the prospect of public hearings.[15]

2.4.5 Annual insurer profits are variable

We also collect information from the National Association of Insurance Commissioners (NAIC), which tracks industry profits by state and by insurance line. Appendix Figure 1 summarizes insurer profits in homeowners insurance from 1985 to 2021. This figure helps to illustrate the year-to-year variability in insurer profits in states that are affected by high-severity, low-frequency weather events. Property insurers operating in these areas must build up surplus capital during uneventful years so as to be able to cover losses incurred during a catastrophe. California's catastrophic wildfires in 2017 and 2018 erased several years of modest insurer profit.

3 New data and descriptive facts on wildfire risk, risk classification, and insurance pricing

This section introduces the new dataset of insurer-level, structure-specific insurance price schedules that we created for this paper. We present a number of new stylized facts about wildfire risk classification and wildfire pricing that emerge from this data. All datasets are described in more detail in Appendix A.

We use proprietary wildfire risk analytics and property characteristics data from CoreLogic, Inc., a leading provider of property information and risk analytics. We obtained these data for a sample of 100,000 single-family houses from 400 California zip codes with meaningful wildfire risk exposure. Details on the sampling procedure are in Appendix A. There are three components of these data:

- **Property characteristics:** Parcel characteristics include estimated reconstruction costs, year of construction, fire department quality, roof type, and other variables commonly used to price insurance.
- **Categorical wildfire risk scores (WRS):** These scores, which range from 5 to 100, are designed to help insurers classify wildfire risk. A home's score reflects a number of location-specific factors that can be measured by remote sensing, such as slope, fuel loads, and a location's proximity to ember-generating, high-risk vegetated areas.

14. Individual filings are publicly available from: CDI. (2024). Web Access to Rate and Form Filings (WARFF). http://www.insurance.ca.gov/0250-insurers/0800-rate-filings/0050-viewing-room.

15. Balber, Carmen. (2024). Consumer Watchdog Calls for Public Hearing on State Farm's Unprecedented Request for 4-Year, $5.2 Billion Policyholder Bailout. Consumer Watchdog. https://consumerwatchdog.org/insurance/consumer-watchdog-calls-for-public-hearing-on-state-farms-unprecedented-request-for-4-year-5-2-billion-policyholder-bailout.

- **Parcel-level loss predictions:** CoreLogic's 2021 wildfire CAT model is used to simulate probabilistic estimates of annual losses from wildfire. This simulation-based model considers both landscape attributes and structure characteristics, such as construction material and year built. We obtain measures of yearly losses for each home in the data based on thousands of model simulations. The reported statistics for each home include the average annual loss (averaged across all model simulations for a given year), the standard deviation of modeled annual loss realizations, and aggregate exceedance probability losses over return periods of 50, 100, 250, and 500 years.

Parcel-specific average annual losses (AALs) serve as our preferred measure of assessed wildfire risk. We use AAL values to proxy for the best available estimates, in 2021, of location-specific wildfire risk exposure. Assessing the accuracy of AAL estimates is challenging because wildfires are infrequent. However, in recent work, Wylie et al. (2024) find that CoreLogic AAL estimates compare reasonably well against recorded long-run historical wildfire losses. In Appendix E.1 we assess the robustness of our findings to alternative wildfire damage probabilities which we derive from publicly available gridded U.S. Forest Service data.

We use these proprietary data, together with information extracted from regulatory rate filings, to document four facts that motivate the empirical analysis.

3.1 Assessed wildfire risk varies significantly across homes

The top panel of Table 2 summarizes the variation in structure characteristics and wildfire AAL across the homes in the dataset. The average home is 2,135 square feet, was built in 1976, and would cost about $600,000 to rebuild. The mean and standard deviation of wildfire AAL values in 2021 was $303 and $596, respectively. Notably, these AAL values vary significantly within zip codes. Regressing AAL on zip code fixed effects explains only 46 percent of the total variation in the data. The AAL distribution is right-skewed; almost 5 percent of the homes in the data are associated with AAL values that exceed $1,300. To contextualize the size of these AALs, the average premium for a Californian HO policy with greater than $500,000 of coverage was approximately $1,960 in 2021.[16]

3.2 Risk classification strategies vary significantly across insurers

Due to modeling limitations and operational constraints, estimating property-specific wildfire premiums is challenging. Insurers instead pool individuals with similar risk profiles into risk classes. This practice of risk classification is a general feature of insurance markets (Abraham 1985; Chamberlin 1985; Crocker and Snow 1986). Insurers in California's admitted market must submit detailed documentation of the data sources, formulas, modeling, and risk factors they use in their

16. NAIC. (2023). Dwelling Fire, Homeowners Owner-Occupied, and Homeowners Tenant and Condominium/Cooperative Unit Owner's Insurance Report: Data for 2021. https://content.naic.org/sites/default/files/publication-hmr-zu-homeowners-report.pdf.

risk classification and insurance pricing.

We use these public documents to identify the risk classification strategies used by the ten largest California homeowners insurance groups and the FAIR plan. There has been rapid innovation in this area in recent years. Our review found that prior to 2013 no large admitted market insurer used third-party wildfire risk rating to determine either premium or eligibility. By 2019, approximately 68 percent of the market was insured by a firm using some form of wildfire risk rating, though companies' risk rating strategies vary widely.

Table 1 describes the risk classification strategies used by the ten largest insurers in 2021, the period of our study. Column (1) reports statewide market shares, while column (2) focuses on market share in the 20 percent of zip codes with the highest average wildfire risk. The three insurers with the largest market presence in California – State Farm, Farmers, and the CSAA Insurance Group – are also the firms with the largest market presence in high-risk areas. Other private insurers tend to be underrepresented in high-hazard zip codes compared to their statewide presence, while the FAIR plan is substantially overrepresented.

Column (3) provides a rough summary measure of the level of granularity in wildfire risk classification. To build this measure, we count the number of risk-rating variables that each insurer uses to assess the likelihood of wildfire damages in the location of a given home.[17] Some insurers classify risk on the basis of zip code-level territory factors. Other insurers use categorical wildfire risk scores. Lastly, some insurers use more granular measures generated using probabilistic CAT models. Notably, the insurers with the largest market shares in high-risk zip codes use the most granular risk classification systems. State Farm, for example, divides the state into 1 kilometer grids and applies separate rating factors for wildfire risk in each grid.

The significant variation in risk classification strategies across firms begs the question: Why do all firms not use the most granular risk classification strategies? In theory, firms will only adopt risk classification if the competitive benefits justify the costs (see, for example, Borenstein 1989). Licensing a state-of-the-art wildfire model costs millions of dollars per year.[18] In addition, insurers must hire skilled professionals to manage, maintain, and deploy more sophisticated risk analytics. This includes training local agents and negotiatiating the extensive regulatory approvals required.[19] Adopting the most granular risk classification strategies may only be cost-effective for those insurers with a significant market presence in high-risk areas. We evaluate this hypothesis more formally in Section 7.

17. More information on the measurement of risk-rating variables is available in Appendix A.

18. Jergler, Don. (2021). Grim California Wildfire Outlook Has Insurers Forking Over Big Bucks for Modeling. The Insurance Journal. https://www.insurancejournal.com/magazines/mag-features/2021/07/05/621088.htm.

19. In California, prior approval regulations require each insurer to present evidence that their proposed risk classification systems and associated premiums are actuarially justified. These rules also limit firms' ability to mimic the approved premiums charged by other insurers.

3.3 More sophisticated classification methods better capture wildfire risk

The final column of Table 1 shows that differences in risk classification lead to meaningful differences in the precision with which insurers can price wildfire risk. The reported values are the R^2 from a regression of wildfire AAL on dummy variables for the rating classes implied by each firm's wildfire classification method. These R^2 values range from 0.16 for AAA of Southern California to 0.82 for State Farm.

This measure of the degree to which risk classes predict individual risk relates to the concept of *homogeneity* in the prior literature (see, for example, Abraham 1985). In addition to implications for perceived fairness, the homogeneity of risk classes, i.e. the extent to which customers assigned to the same class share the same risk profile, can affect the incentive value of insurance prices with respect to individual behavior. When risk classes are homogeneous, premiums that reflect the average assessed insurance premiums for each risk class convey information about individual-level risk that can usefully incentivize behaviors to reduce risk exposure.

The box plots in Figure 3 illustrate the extent to which assessed wildfire risk, measured using parcel-level AAL values, varies within and across risk classes defined using categorical CoreLogic wildfire risk scores. The vertical box plots summarize individual expected annual losses per $1,000 of insurance coverage for homes with each risk score. Each box shows the interquartile range, the 10th percentile, and the 90th percentile. On average, higher assessed wildfire losses are associated with higher risk category values. However, individual expected losses vary significantly within risk score classes. The variance of AAL values within a risk segment increases with the risk score level.

A hypothetical insurer that classifies and prices wildfire risk on the basis of categorical risk scores will charge the same wildfire premium to all parcels assigned to the same risk class. This pricing strategy will result in charging premiums that exceed (fall below) assessed risk exposure for households with below (above) class-average risk exposure. Because the purchase of insurance is effectively mandatory, we do not expect that these discrepancies will impact household demand for insurance. In other words, coarse risk classification should not directly result in deadweight loss by distorting the purchase of insurance. However, a disconnect between insurance price and individual risk may affect homeowners' perceived returns to risk-reducing behaviors. Investigating these trade-offs in the context of wildfire is an important area for future research.

3.4 Offered insurance prices vary across insurers

Six large California insurers use CoreLogic wildfire risk scores or geographic territory factors to classify and price insurance. These insurers collectively represent 45 percent of California's homeowners insurance market.[20] We combine CoreLogic data with information collected from insurer rate filings

20. We are unable to calculate parcel-specific prices for the other large insurers because they use wildfire analytics from other data providers, such as Verisk's FireLine scores.

to reconstruct the prices that these six insurance groups would charge each of the 100,000 homes in the dataset.[21] HO insurance prices depend on parcel-level characteristics we can observe, such as the age of the home, construction materials, roof type, and distance to vegetation. Insurance prices also depend on *occupant* characteristics and coverage choices that we cannot observe. To calibrate insurers' pricing formulas, we assume constant reference values for these occupant characteristics.[22] Appendix A describes this approach in more detail.

The middle panel of Table 2 summarizes these calibrated insurance prices. Whereas the insurance prices offered to homes in the data are quite similar on average, the prices offered to any particular home differ substantially across insurers. The five firms that operate statewide all charge average prices close to \$2,400 per year based on coverage at reconstruction cost. The sixth firm, AAA Southern California, operates only in the southern part of the state and thus we only observe prices for about half of the homes in the data for that firm. The final panel of Table 2 shows significant dispersion in prices across insurers for a given home. The average standard deviation in premiums offered to a home in the dataset is \$578. The average range between the minimum and maximum price offered is almost \$1,500.

3.5 Richer pricing methods embed a wildfire information advantage

This section presents a regression-based variance decomposition exercise that achieves two objectives. First, it provides a check on the dataset construction by confirming that the wildfire component of insurance price for each firm does not vary within the customer segments used by that firm for pricing. Second, it begins to illustrate the wildfire information advantage for more granular pricing methods.

For each property i and insurer j in the dataset, we regress insurance premiums p_{ijk} on a set of controls and a set of wildfire risk segment indicators indexed by k:

$$p_{ijk} = g_j(R_i) + X_i\theta_{1j} + \sum_k \varsigma_{jk} D_k + h(l_i; \varphi_j) + \varepsilon_{ijk}. \tag{1}$$

Covariates include a polynomial function $g_j(R_i)$, where R_i is demeaned reconstruction cost, and controls X_i which include 5-year bins for the age of home (with single bins for homes built before 1950 or after 2015), indicators for roof type, and categorical variables for public protection class.

21. This generated dataset is unique in that it includes complete price menus for each firm, versus the subset of transaction-based prices paid by customers who choose to buy from a given firm. Not all of the prices we construct will be offered to customers in the market. As we discuss below, over the time period we study, some insurers had stopped offering insurance to new customers in some market segments. Thus, some of the prices we calibrate would only be available to existing customers. Some firms also engaged in non-renewals of a subset of existing customers in some segments.

22. For example, we assume that a homeowner purchases coverage equal to the reconstruction cost of the home (which is generally advised by insurers), chooses a \$1,000 deductible, has not had a recent claim, and bundles their homeowners and automobile insurance policies. About 78 percent of consumers bundle their HO and auto policies. See: J.D. Power. (2015). 2015 U.S. Household Insurance Study. https://www.jdpower.com/business/press-releases/2015-us-household-insurance-study.

We parameterize a flexible function of parcel-specific wildfire AAL values $h(l_i; \varphi_j)$ using a step-wise function of l_i with bin widths of \$25. The variable φ_j is an insurer-specific vector of coefficients that summarize the relationship between residual variation in insurance premium and AAL. In other words, $h(l_i; \varphi_j)$ is a binned regression specification with \$25 bins of wildfire AAL.

We estimate Equation 1 using increasingly granular sets of risk segment indicators D_k. When the D_k indicators are coarser than the risk classification used by an insurer, we should find a positive relationship between the residual variation in premiums and residual variation in AAL. This relationship should be completely absorbed once we condition on the risk segments used by the firm to price wildfire risk because all parcels assigned to the same risk class are offered the same insurance price.

Figure 4 illustrates results from this variance decomposition. The top panel reports results for Allstate and Nationwide. Both insurance groups use CoreLogic's categorical risk scores to classify wildfire risk and price insurance. Using various sets of fixed effects, a positive relationship between premium and risk remains; it is only when the regression includes wildfire category indicators that the relationship is absorbed completely. The middle panel shows results for USAA and Liberty Mutual, both of which group homes into wildfire rating territories based on zip codes. As expected, zip code fixed effects soak up all of the relationship between premium and risk. The lower panel displays results for the last two insurers. AAA Southern California uses the fewest risk classes of any insurer in the data, with only 19 distinct categories derived from zip codes; this insurer displays limited correlation between premium and risk, even when omitting zip code fixed effects. In contrast, State Farm assesses wildfire risk at a 1-square km resolution using CAT model predictions, generating 12,000 distinct risk classes. Figure 4 shows a strong and positive relationship between premiums and AAL that persists even when the regression includes wildfire category scores and zip code fixed effects. Controlling for individual grid cell dummies eliminates the residual relationship between prices and wildfire risk.

4 Average Cost Benchmarks for Wildfire Risk Pricing

We have documented striking variation in the risk information that insurers use to classify and price wildfire risk. To assess the implications of these risk information asymmetries for insurer costs and insurance pricing, we specify a model of insurance costs that incorporates not only expected damages but also administrative costs, fixed costs, and risk load. Building on earlier models of economic regulation and average-cost pricing (e.g., Viscusi, Harrington, and Sappington 2018), we adopt the perspective of a regulator who is guided by principles of fair and adequate insurance pricing. From this regulator's perspective, we investigate how prices should rise to account for adverse selection. Our goal in this section is to construct benchmarks for cost-based insurance pricing that we can compare to the empirical risk pricing gradients in Section 5. We defer the richer equilibrium model to Section 7, which considers how a firm with superior information would strategically leverage its risk information advantage in a competitive equilibrium.

4.1 Cost-based insurance pricing under symmetric information

Consider an insurance market in which firms offer a single insurance contract that covers multi-peril property damages. In the spirit of Akerlof (1970), insurers compete on prices but do not compete on contract features. Households indexed by i must buy insurance for disaster losses from insurers indexed by j. Households vary in terms of how costly they are to insure. The variable l_i denotes expected disaster claims for property i during the annual contract period. This variable is distributed in the population of properties with mean $\bar{l}$ and variance σ^2.

When households choose an insurance provider, we assume they select the contract to maximize indirect utility u_{ij}. The price charged by insurer j to insure property i is p_{ij}. Indirect utility is denoted $u_{ij} = \delta_j - p_{ij}$, where δ_j represents the average brand preferences for insurer j plus any switching costs incurred by households who already hold insurance.[23] The value associated with choosing a firm other than j is given by $\bar{u}_{ij} = \max_{\text{-}j}(\delta_{\text{-}j} - p_{i,\text{-}j})$, where -$j$ indexes insurers other than insurer j. Demand for firm j is thus:

$$Q_j = \sum_i 1[\delta_j - p_{ij} \geq \bar{u}_{ij}]. \tag{2}$$

Insurer j's book of business, Ω_j, is comprised of the group of customers who purchase insurance from insurer j.

When considering the costs to insure customer i, the expected value of insurance payouts provides a useful point of departure. However, setting insurance prices at actuarially fair levels will not in general be economically sustainable because insurers must also cover fixed and variable operating expenses, plus the costs of holding sufficient capital reserves to pay out claims with an acceptably low probability of default. The amount of surplus capital or reinsurance that a firm needs to hold will depend on the risk characteristics of its book of business Ω_j.[24] The cost of holding the required capital surplus or reinsurance is often referred to as catastrophe risk "load" (Kunreuther and Michel-Kerjan 2011; Stone 1973). Let $\phi(\Omega_j)$ denote risk load, and define the firm's total annual cost function as:

$$c_j = \sum_i (l_i + a_i) 1[\delta_j - p_{ij} \geq \bar{u}_{ij}] + \phi(\Omega_j) + F_j. \tag{3}$$

The variable a_i denotes administrative costs plus the expected costs for non-catastrophe losses such as burst pipes or liability claims, whereas F_j represents any fixed costs, such as the costs to license more granular risk analytics.

A cost-based insurance price ρ_i reflects the expected marginal cost to insurer j from adding customer i to its portfolio, plus a cost component f_i that allows the firm to prudently recovery fixed

23. When two firms yield identical utility, we assume the household randomizes which to buy.

24. As an example, solvency regulations may require firms to hold sufficient reserves or reinsurance to cover their full losses in a 1-in-200 loss year, that is, a 99.5th percentile realization of total losses.

costs:

$$\rho_i(\Omega_j) = l_i + a_i + \phi'_{ij}(\Omega_j) + f_i. \tag{4}$$

In practice, insurers group customers into risk classes within which they apply constant pricing factors, as described in Section 3.2. We first consider a scenario in which all insurers use the same information to classify and price wildfire risk. Assume that all insurers observe the same discrete signal $s(l_i) \in \{s_0, s_1, ..., s_K\}$ which assigns a home to a risk class or segment indexed by k. Insurers know the average risk level for each segment, $\bar{l}_k$, but not the individual-specific l_i values. The expected cost associated with parcel i in segment k is given by:

$$\rho_k(\Omega_j) = \bar{l}_k + \bar{a}_k + \bar{\phi}'{}_{jk}(\Omega_j) + \bar{f}_k, \tag{5}$$

where $\bar{a}_k$, $\bar{\phi}'{}_{jk}(\Omega_j)$, and $\bar{f}_k$ are segment-level averages over parcels i in segment k. Equation 5 yields an adequate insurance premium because it covers insurers' expected costs. It is also "fair" insofar as members of each class are charged in accordance with expected costs.

In the analysis that follows, we will estimate the empirical relationship between insurance premiums and assessed wildfire risk. To motivate this *ceteris paribus* comparison, consider the difference in insurance costs between a home assigned to the zero risk class ($\bar{l}_0 = 0$) and an otherwise identical home ($a_i = a$; $f_i = f$) assigned to a higher risk segment ($\bar{l}_k > 0$). The slope of this relationship is:

$$\frac{\bar{l}_k + a + \bar{\phi}'_{jk}(\Omega_j) + f - (\bar{l}_0 + a + f)}{\bar{l}_k - \bar{l}_0} = 1 + \frac{\bar{\phi}'{}_{jk}(\Omega_j)}{\bar{l}_k} \equiv \beta_{jk}. \tag{6}$$

The β_{jk} parameter measures the increase in the segment-average insurance costs associated with a unit increase in the segment-average expected loss. We will refer to this subsequently as the risk price *gradient*. This cost-based risk price gradient has a slope of 1 plus the average marginal surplus term $\frac{\bar{\phi}'{}_{jk}(\Omega_j)}{\bar{l}_k}$.

4.2 Marginal surplus calculations

To calibrate the risk price gradient empirically, we need to estimate insurers' marginal surplus costs. Following Kreps (1990), we formulate this cost component as the product of the insurer's costs of capital, a "distribution statistic" z, and the change in the standard deviation of the firm's loss distribution after adding parcel i to the risk portfolio:

$$\phi'_{ij}(\Omega_j) = \underbrace{\frac{y}{1+y}}_{\text{capital cost}} \times \underbrace{z}_{\text{“distribution statistic”}} \times \underbrace{\frac{(2S_jC_{ij} + \sigma_i)\sigma_i}{S_j + S'_{ij}}}_{\text{change in s.d. of firm's losses}}. \tag{7}$$

The y parameter denotes the market cost of capital and z is a distribution statistic that indicates the number of standard deviations above a mean loss year that the firm can survive. The change in the standard deviation of the firm's annual total losses after adding parcel i to the portfolio is expressed in terms of the standard deviation of losses from the existing book of business S_j, the standard deviation of losses from the combined book of business S'_{ij} with the addition of parcel i, the standard deviation of annual losses associated with parcel i, σ_i, and the correlation of losses between the new parcel and the existing book of business, C_{ij} (Kreps 1990).

We approximate the risk load associated with covering successively larger portfolios of HO policies in high wildfire hazard areas of California using CoreLogic's parcel-specific AAL distributions. Appendix C describes how we build a statewide pseudosample to explore the relationship between risk load and market share. We calibrate measures of marginal surplus under a range of assumptions about the correlation in risks across homes and the risk profile of an insurer's country-wide book of business.

Appendix Table 5 shows how the calibrated marginal catastrophe load costs per dollar of assessed AAL increase with market share. Intuitively, firms with a higher concentration of customers in high-hazard areas have higher surplus requirements due to spatially correlated risks. For a representative 30 percent market share in high-hazard zip codes, we estimate that an additional dollar of California wildfire risk exposure increases surplus requirements by about \$0.18.[25] The largest HO insurer in California claims less than a 20 percent share in higher-hazard zip codes, as reported in Table 1. If we conservatively assume a market share of 30 percent, this implies that an additional dollar of wildfire risk exposure should increase insurance costs by no more than \$1.18.

4.3 Cost-based insurance pricing under asymmetric information

Thus far, we have assumed that all insurers access the same wildfire risk information. In practice, insurers bring different information to wildfire risk classification and pricing. We now investigate the implications of asymmetric information for average cost pricing.

Consider a scenario in which an insurer j acquires superior wildfire risk information at a cost of F^c, which figures into the firm's fixed costs F_j. If we assume the information is sufficiently rich to support parcel-level risk assessment, cost-based pricing for this insurer is given by Equation 4. If we further assume that the fixed costs of licensing and using the information are recovered symmetrically across customers, the property-level fixed cost component f now includes this information cost. Other insurers in the market, denoted -j, do not incur the costs of adopting superior information and continue to rely on relatively coarse risk classification strategies. Maintaining the assumption that customers purchase from the insurer offering the lowest price, the probability that a relatively uninformed firm "wins" a customer will now be correlated with expected disaster claims. In other words, firms at an information disadvantage will face a "winner's curse":

$$\omega_{\text{-}jk} = E[l_i | s_{\text{-}jk}, \delta_{\text{-}j} - p_{i,\text{-}j} > \delta_j - p_{ij}] - E[l_i | s_{\text{-}jk}] > 0, \tag{8}$$

25. Kunreuther and Michel-Kerjan (2011) calibrate loading factors for insurance pricing in areas at high risk for hurricanes that include both risk load and administrative costs. They calibrate a "loading factor" of \$0.50, which serves as a useful point of comparison. We should expect that risk loads for hurricane risk should be significantly higher than wildfire risks given the magnitude of damages and high degree of spatial correlation in these risks.

where p_{ij} denotes the price offered by the relatively more-informed firm in segment k and $s_{\text{-}jk}$ indexes the risk segments used by the less-informed firm. To ensure that revenues cover expected costs, firms operating at an information disadvantage should adjust their prices to account for this winner's curse.

4.4 An applied example

To elucidate the potential for an economically significant winner's curse in this market, we consider a stylized scenario that we calibrate using risk classification strategies akin to those we observe and the data on assessed wildfire risk. This example temporarily ignores several features of the market, including strategic behavior by the firm pricing on better information. We extend our framework to accommodate strategic behavior in the empirical analysis of Section 5 and the equilibrium model of Section 7.

Suppose one insurer, denoted Firm A, classifies wildfire risk using CoreLogic's categorical risk scores. If all insurers in the market use this same risk classification system, average cost pricing would involve setting premiums equal to the segment-specific average AALs plus the calibrated risk load factor. This wildfire risk pricing schedule is illustrated graphically by the circular markers in Figure 3.

Next, consider an alternative scenario in which Firm A competes against a firm, denoted Firm B, that segments based on granular one-kilometer grid cells (as in the case of State Farm). A property owner purchases insurance from either Firm A or Firm B. For expositional ease, we ignore perils other than wildfire and we abstract away from risk load costs.[26]

Under the assumption of average-cost pricing (which we impose here for expositional ease but relax in the equilibrium model in Section 7), Firm B charges premiums equal to the expected average wildfire AAL within each of its highly granular risk categories. If Firm A were to naively set premiums equal to the average assessed wildfire risk within each of its coarser risk classes (i.e. the circular markers in Figure 3), these prices would fall below expected costs because Firm A now serves an adversely selected sample of customers. The square markers in Figure 3 denote the average AAL values for the customers that purchase insurance from Firm A in this scenario. The average expected loss in each risk category (i.e. the winner's curse) is estimated by the distance between the conditional (square marker) and unconditional (circular marker) average AAL values. Because the variance of AAL values is larger in higher wildfire risk categories, the winner's curse is larger in higher wildfire risk categories.

To ensure that earned premiums cover expected costs, Firm A should charge prices that reflect the

26. We assume that insurance is a homogeneous good and that there is a homogeneous switching cost. In this stylized example, we use a switching cost of $500 and hold reconstruction cost constant for all homes at the sample average of $600,000. Initial firm-customer matches are randomly determined in an assumed "period 0" prior to introduction of the improved risk classification model which randomly and equally allocates customers across the two insurers. We drop a small number of homes in one-kilometer grid cells that contain only a single home in our data, which makes this exercise conservative, i.e. reduces implied selection.

average costs of the properties it insures. These costs will exceed risk-class average AAL values. Note that, were Firm A to increase its prices to follow the square markers in Figure 3, the risk profile of its customer pool would shift toward higher-cost customers. In auxiliary calculations in Appendix Section B, we simulate break-even prices for Firm A in this stylized example. In the higher-risk classes, these break-even prices are an order of magnitude higher than segment-average expected losses. Thus, the break-even risk pricing gradient is steeper than the gradient that ignores adverse selection.

5 Empirical Estimates of the Risk Price Gradient

In this section, we analyze the empirical relationship between insurance prices and assessed wildfire risk exposure using data from California's HO insurance market. We use data on wildfire risk and firm-specific annual premiums and risk classification strategies to estimate empirical analogs of the risk price gradient in Equation 6. We estimate separate empirical risk price gradients for each insurer and compare these against corresponding average cost benchmarks.

5.1 Econometric specifications and identification

For each insurer in the data, we estimate an empirical analog of Equation 6, the risk price gradient. For each insurer, we estimate the following two equations:

$$l_{ik} = \psi_j + \sum_{k=1}^{K_j} \lambda_{jk} D^*_{jk} + X_i\theta_{2j} + g_j(R_i) + \zeta_{ijk}; \tag{9}$$

$$p_{ijk} = \mu_j + \sum_{k=1}^{K_j} \gamma_{jk} D^*_{jk} + X_i\theta_{3j} + g_j(R_i) + \xi_{ijk}, \tag{10}$$

where l_{ik} denotes the parcel-level AAL and p_{ijk} denotes the insurance premium charged by insurer j for parcel i in segment k. The D^*_{jk} are indicator variables for the k segments that firm j uses to price wildfire risk.[27] The variables $g_j(R_i)$ are flexible polynomial functions of demeaned reconstruction costs.[28] The vector X_i includes additional controls for observable property characteristics.[29] The λ_{jk} parameters in Equation 9 estimate the average wildfire AAL in pricing segment k used by firm j, holding constant reconstruction cost and other factors. The γ_{jk} coefficients in Equation 10 recover the average insurance price charged by firm j in segment k, holding constant reconstruction cost and other factors. We choose the lowest-risk wildfire segment as the omitted D^*_{jk} category.

27. Since each insurer uses a different segmentation strategy, k is itself a function $k(j)$ of j. We omit this additional subscript to simplify notation.

28. The $g_j(\cdot)$ controls adjust for premium differences driven by differing reconstruction costs R_i. Our main specifications use a seventh-degree polynomial to allow for non-linearity in R_i. The results are similar if we control linearly for R_i or simply divide by R_i to normalize premiums and wildfire AALs by dollar of reconstruction cost.

29. In our main specification, X_i includes age of home in five-year bins, public protection class, and roof type.

Thus, the parameters ψ_j and μ_j estimate the average AAL and annual premiums, respectively, in the lowest wildfire risk segment.

For each insurer, we estimate a large number of γ_{jk} and λ_{jk} coefficients (one for each risk class). We use these estimates to assess the relationship between segment-average premiums and corresponding segment-average assessed risk, i.e., the risk price gradient β_{jk}. A linear approximation of this relationship is estimated by regressing the estimated γ_{jk} on the estimated λ_{jk}. We weight this bivariate regression according to the number of homes in each risk segment and calculate standard errors for β_{jk} by bootstrapping all steps of the estimation procedure using a zip code-level block bootstrap.

This empirical strategy can consistently estimate β_{jk} provided that our conditioning strategy effectively captures the variation in non-catastrophe losses and other cost components that generate variation in insurance premiums. Even after controlling flexibly for reconstruction costs and other factors we can observe (e.g. roof type, home age), the residual variation in assessed wildfire risk exposure could still be correlated with insurance cost drivers such as local crime rates that are priced at the zip code level. To address the potential for omitted variable bias, we estimate a second set of specifications that include a full set of zip code fixed effects. The advantage of this approach is that, with the possible exception of State Farm, this strategy will absorb variation in all cost drivers assessed by the firm that could generate variation in zip code average premiums and induce spurious correlation between premiums and wildfire risk.[30]

The approach described up to this point can be applied directly to the two firms, Allstate and Nationwide, that classify wildfire risk coarsely based on categorical risk scores. Because we observe many parcels within each of the small number of risk segments used by these insurers, we can estimate average AAL for homes in each segment with precision. For insurers using more granular approaches, an additional correction is required to address potential measurement error. When firms use many rating segments, there is greater sampling error in the λ_{jk} coefficients, since our estimate of the average AAL for homes in each of the many risk classes is based on fewer observations. The increased noise in the estimated λ_{jk} may lead to measurement error-induced attenuation bias in the bivariate regression of γ_{jk} on λ_{jk}. To mitigate this problem, we use the CDI measures of wildfire risk introduced in Section 2.4 as instruments for segment-average AAL values when estimating the risk price gradients.[31] CDI zip-code-by-year measures of wildfire risk are noisy, but strongly correlated with AAL values and presumably uncorrelated with the error process. Appendix D discusses this two-stage least squares (2SLS) strategy in detail.

30. Inspection of the detailed pricing formulas in the rate filings confirms that with the exception of State Farm, wildfire risk is the only spatially-varying peril that is priced at the sub-zip code level. Other perils, such as crime or water pipe failures, are priced at the zip code level or higher.

31. Wooldridge (2010, Section 5.3) reviews how a two-stage least squares approach reduces attenuation bias.

5.2 Empirically-estimated risk price gradients

Figure 5 summarizes estimates of γ_{jk} and λ_{jk} from Equations 9 and 10 for Allstate and Nationwide. The vertical axis in each plot shows the λ_{jk} coefficient estimates with confidence intervals. The horizontal axis in each plot shows the γ_{jk} estimates with confidence intervals. Each marker corresponds to a different risk class or segment used by the insurer to price wildfire risk.

Panel (a) reports γ_{jk} and λ_{jk} estimates for Allstate without including zip code fixed effects. Panel (d) reports analogous results for Nationwide's prices and risk segments. In each case, the average annual premium in the lowest wildfire risk segments (which have negligible wildfire AALs) is in the range of \$1,500 to \$1,750. This gives an estimate of average per-customer non-wildfire claims costs and variable administrative costs (i.e. the a cost component) as well as profit margins and the fixed cost component (f). These average premia for the lowest wildfire-risk homes define the anchor point for the cost-based risk pricing benchmarks derived in Section 4. The grey wedges span the upper and lower bounds of the cost-based gradients introduced above. Recall that these benchmarks pertain to scenarios in which all firms use the same risk classification strategies and thus offer the same wildfire premiums.

The β_{jk} estimate is reported in each panel, along with bootstrapped standard errors. The Allstate and Nationwide slope parameters in panels (a) and (d) are both significantly steeper than the upper bound on the cost-based symmetric information benchmark. This steep relationship is qualitatively consistent with our theoretical predictions. Given how assessed AAL values are distributed across and within these categorical risk classes, price adjustments that account for adverse selection will generate steeper price gradients. Notably, empirical relationships we observe appear concave in the highest-risk segments. This shape could be due to regulations that constrain insurer pricing, a topic which we return to in Section 6.2.

Panels (b) and (e) in Figure 5 report the coefficient estimates from the more saturated specifications that include zip code fixed effects. In these figures, the estimated γ_{jk} parameters in very low-risk segments are close to zero because the zip code fixed effects absorb the average costs of insuring homes in these segments. Comparing the risk price gradients from these more saturated regressions against the cost-based benchmarks, we observe similar qualitative patterns as above.

For comparison purposes, panels (c) and (f) report results from estimating the same regression equations reported in panels (b) and (e) using State Farm's prices as the dependent variable in Equation 10. In other words, we assess the prices that State Farm charges for homes in the risk segments defined by its rival firms. The linear approximation to State Farm's gradient parameters is noticeably less steep. With the exception of the highest wildfire risk bins, which contain a small fraction of properties, this risk price gradient closely tracks the cost-based benchmark.

5.3 2SLS estimates of insurer-specific risk pricing gradients across firms

Table 3 reports the results of the 2SLS estimation described in Section 5.1 for a larger group of firms.[32] The top panel reports 2SLS estimates of the linear relationship between γ_{jk} and λ_{jk} for each insurer. These 2SLS β_{jk} estimates for Allstate and Nationwide are very similar to the OLS-based β_{jk} estimates reported in Figure 5. This result suggests minimal attenuation bias for these two firms given the large number of parcels we observe in each of the risk classes used by those two firms. Bringing in estimates for other firms that price more granularly, the estimated price gradients for USAA and Liberty Mutual are 2.17 and 2.46, respectively. Both are significantly steeper than the corresponding cost-based benchmarks. The State Farm price gradient estimate is very close to the cost-based gradient benchmark; we fail to reject a slope of 1.18.

The bottom panel replicates this same estimation exercise to consider State Farm's pricing within the risk segments defined by each of its competitors. For each insurer other than State Farm, the table reports 2SLS estimates of the relationship between average wildfire AAL and average State Farm price in each risk segment in the other firm's classification system. In contrast with the steep gradients in other firms' prices, State Farm prices rise commensurately with assessed average wildfire risk across the range of classification systems used by other firms.

6 Further Evidence on Adverse Selection and Regulation

We have documented significant differences in the risk classification and pricing strategies used by major property insurers in the California market. The risk pricing gradient associated with the most sophisticated firm, State Farm, tracks the cost-based benchmark quite closely. In contrast, among firms using relatively less granular risk classification strategies, estimated risk pricing gradients are significantly steeper. These findings are qualitatively consistent with our model of cost-based pricing under asymmetric information introduced in Section 4.3. In what follows, we provide more quantitative evidence on the potential cost implications of adverse selection. We also provide evidence that winner's curse adjustments made by firms at an information disadvantage interact with binding price regulations.

6.1 Adverse selection

If we could observe how properties in the dataset were allocated across insurers, we could directly address the degree of adverse selection in this market. Because this information is proprietary and unavailable, we use the data we do have to indirectly assess the potential for adverse selection.

32. AAA Southern California is excluded from Table 3 because the first stage regression for that firm does not pass conventional rule of thumb tests for weak instruments (e.g. Staiger and Stock 1997), possibly due to the small number of rating segments used by AAA Southern California. Appendix Section D reports OLS and IV estimates for all firms.

6.1.1 Contextualizing the empirical risk price gradient

The winner's curse implies that firms may "shade up" prices in high-risk segments to account for adversely-selected parcels. To assess whether empirically estimated risk price gradients associated with information-disadvantaged firms could possibly be rationalized by adverse selection, we leverage information we have about the distribution of AAL values within risk classes or segments. We compare insurers' premiums across different segments with corresponding conditional quantiles of the distribution of AAL values.

Figure 6 contrasts the empirically-estimated risk price gradients against quantile regression estimates that summarize the distribution of expected wildfire insurance costs in each risk segment bin, holding constant the controls in the main regressions. The colored markers denote the average wildfire prices charged by Allstate, Nationwide, and State Farm within each WRS category. The solid black line shows the mean wildfire AAL, multiplied by 1.18 to reflect the assumed risk load. The dashed lines show the 90th and 95th conditional quantiles of 1.18 $\times$ wildfire AAL in each segment.[33]

For wildfire score segments below 50 (the level CoreLogic defines as "low"), all three firms charge prices that closely track the wildfire risk segment mean AAL values plus the calibrated risk load values. For higher risk classes, State Farm continues to price close to the segment average values of assessed wildfire risk, whereas Allstate and Nationwide charge much higher prices. Allstate's prices approximately track the 90th percentile of customer AAL, while Nationwide's price schedule lies between the 90th and 95th values. Thus, the risk pricing gradients we observe for Nationwide and Allstate could possibly be rationalized by high levels of adverse selection in this market.

6.1.2 Modeling adverse selection generated by observed risk pricing

We can quantify the impact of adverse selection on an insurer's costs, even absent detailed data on insurers' risk exposure, if we make some assumptions about the nature of competition in this market. We consider a stylized Bertrand setting that builds on the demand model introduced in Equation 2. We begin with a homogeneous product setting where the δ_j parameters are assumed to be zero so that the insurer offering the lowest price p_{ij} "wins" parcel i. Within this simple framework, given the offered insurance premiums we observe in this market, we can simulate the allocation of parcels across firms. We also explore how this allocation is affected by varying the δ_j "switching cost" parameter.

This exercise is conceptually related to the calculation of a cost-based pricing benchmark under asymmetric risk classification in Section 4.3 and Figure 3. Those calculations model insurer costs when both firms naively set prices equal to the mean of expected losses for all customers in each rating segment. Here, we instead study outcomes at the prices that insurers have chosen to charge

33. We cannot extend this analysis to the firms that price at the zip code level because zip code fixed effects absorb all the insurance price variation.

in the observed market equilibrium. These prices reflect any strategic adjustments that firms have made to protect against selection or exploit information advantages. Thus, this exercise is informative about the degree of selection that exists in this market in equilibrium. At the same time, we note that this simple modeling exercise does not capture all features of the market. Our dataset does not include all insurers active in California, so that some of the parcels "won" by an insurer in this dataset may have faced lower prices from another insurer. Our parameterization of δ_j also simplifies a suite of potentially relevant factors including varied attention and brand preferences into a single reduced-form "switching cost."[34]

We focus primarily on hypothetical duopolistic competition between State Farm and one other firm in the dataset. For each of the relatively less-informed firms in the data, define an indicator variable $1[Win]_{ij}$ that equals 1 if firm j offers a lower insurance premium to parcel i, as compared to State Farm. We estimate the following equation:

$$\log(l_i) = \sum_{k=1}^{K_j} \tilde{\lambda}_{jk} D^*_{jk} + X_{ij}\theta_{4j} + \omega_j 1[Win]_{ij} + e_{ij}, \tag{11}$$

where l_i is the AAL of parcel i, D^*_{jk} are indicators for firm j's risk segments, and X_{ij} includes all other variables used by the less-informed firm. The coefficient ω_j estimates the log difference in wildfire AAL for properties won by firm j, relative to properties in the same risk segment won by its relatively more-informed rival.

Table 4 summarizes the adverse selection that manifests in this static duopoly setting. The households to whom Allstate offers a lower price in duopoly competition with State Farm have AAL values that are approximately 50 percent higher, on average, than the properties that appear identical in Allstate's rating system but are won by State Farm. Although the magnitude of this winner's curse varies across insurers, all estimated values are economically and statistically significant.[35]

The results in Table 4 will overstate the extent of adverse selection if choice frictions captured in δ_j limit switching behavior. Prior work has documented meaningful search and switching costs. In automobile insurance markets, for example, Honka (2014) estimates search costs in the range of \$35 to \$170 and average switching costs of \$40. In a health insurance context, Heiss et al. (2021) estimate switching costs in the range of \$300 to \$600. We extend the static duopoly framework to incorporate choice frictions and switching costs δ, where the j subscript is suppressed because switching costs are assumed equal across firms. As in Section 4, we assume that insurers initially compete in a symmetric information setting (e.g., the pre-2010 period when insurers invested little in wildfire pricing), such that parcels are randomly allocated across insurers. In a second stage, differences in risk classification create a more granular risk pricing strategy for the more-informed

34. As an aside, we note that the regression method proposed in Equation 11 could exactly quantify the degree of selection faced by an insurer if transaction-level data on the book of customers served by a given firm were to be available, for example through a research partnership with an insurer.

35. The bottom panel reports the fraction won by the relatively uninformed firms in these Bertrand duopoly games. The market is roughly evenly split between the two firms, with State Farm winning lower-risk homes on average.

insurer. We model each customer's choice of duopolists in the second stage given a range of potential δ values. Table 5 shows how the estimated winner's curse from Equation 11 varies with these assumed switching costs. Intuitively, the degree of adverse selection is declining with switching costs.

Appendix B.2 presents analogous regressions where we model simultaneous competition between all six firms in the dataset, instead of the duopoly framework. The results similarly point to economically and statistically significant adverse selection for insurers other than State Farm.

6.1.3 Alternative explanations

Our empirical results are consistent with insurers which use less granular risk classification strategies setting wildfire premiums above risk segment average risk to account for adverse selection. Stylized Bertrand simulations that allocate consumers to firms support the rationality of such a strategy in this market. However, there could be other factors that also contribute to the divergent price gradients that we observe across the six firms in the constructed dataset. Differential premium markups in high-risk market segments could in principle relate to differences in risk load across insurers, differences in risk assessment, or differences in ambiguity aversion.

Appendix E presents additional analysis germane to these alternative explanations. First, we replicate the empirical analysis using alternative and publicly available measures of wildfire risk. We observe similar qualitative patterns, both with respect to State Farm's information advantage and the associated winner's curse implications for firms using coarser risk classification. These alternative results imply that our conclusions are not the result of having used a particular wildfire model to generate the AALs in the analysis. We additionally investigate the possibility that differences in ambiguity and/or risk loading can explain steeper risk pricing gradients among these firms. Similarities along key dimensions we can observe (e.g., loss ratios and risk diversification) suggest it is unlikely that differences in rate schedules are explained by differences in risk load costs.

6.2 Regulation

Public discussions about increasing prices and declining availability in California's property insurance market have raised important questions about the role of regulation, and the extent to which price regulations are limiting firms' ability to set premiums at a level that fully reflects assessed risk. Although the wildfire risk price gradients we estimate shed important light on how offered premiums vary with assessed wildfire risk, holding constant other cost components, they are not informative about the extent to which these premium *levels* cover all insurance costs because we cannot observe expected non-wildfire claims for this multi-peril product.

We observe bunching in requested rate increases, as depicted in Figure 2, which suggests that constraints on average premium increases are binding in this market. For State Farm, the estimated risk pricing gradient tracks our benchmark cost gradient across almost all segments, as shown in

Figure 6. Given the potential for positive selection in the case of State Farm, this suggests that State Farm is able to price wildfire risk at a level that is commensurate with – or exceeds – risk-related costs in all but the very highest-risk parcels. More generally, we observe some degree of concavity in all of the empirical risk pricing gradients. One possible explanation is that these gradients are compressed by regulation in the highest-risk segments. This compression could also have implications for insurance availability.

We have shown how a rational firm should adjust prices upward to account for adverse selection when competing with a rival that prices risk using superior information. If price regulations limit firms' ability to shade up premiums in high-hazard areas, relatively uninformed firms may elect to stop writing new policies in high-risk market segments. Empirically, this is what we observe among the subset of insurers in the data for whom we can reconstruct eligibility rules. Using these rules, we can estimate the fraction of homes in the dataset that would have been eligible for a new insurance policy in 2021. Figure 7 shows how these eligibility rates vary with categorical wildfire risk scores. At this time, Allstate and Nationwide were not offering new policies to most parcels with risk scores above 30, whereas State Farm has historically maintained much higher acceptance rates in higher-risk categories. Taken together, these results reinforce the importance of the interaction between rate regulation and imperfect information about wildfire risk. These interactions are an important focus of the equilibrium model developed in the next section.

7 An Equilibrium Model of Asymmetric Information and Binding Regulation

We develop an equilibrium model of an insurance market to explore further implications of the empirical findings for pricing and availability. The model illustrates how a more informed firm can "cream skim" low-risk customers within a given segment, leaving higher-cost customers for less-informed firms. In addition, we show how asymmetric information interacts with regulatory price constraints, inducing less-informed firms to exit the market. Policies that reduce information asymmetries can reduce insurance premiums while also increasing availability of insurance.

We model a single risk segment, such as a geographic area (e.g., a zip code) or a set of properties assigned to a particular risk score. We assume that risk, measured by AAL, is uniformly distributed across properties in the segment. Similar to the model in Section 4, firms can make costly investments in risk information, and consumers face costs of switching insurers. Initially, we assume that firms' prices are unconstrained by regulation. We derive the market equilibrium and the expected value of information associated with the adoption of risk information. We then extend the model by incorporating regulatory constraints. In the latter case, we consider the affordability and availability implications of policy reform.

7.1 Setup

Firms offer insurance to a group of property owners within a wildfire risk segment. Structure values are assumed to be identical, but wildfire risk varies among properties. Property risk, defined as the expected loss l, is uniformly distributed according to $U(0, l^*)$, with mean risk $\bar{l}$ and variance σ^2. A firm charges p for a homeowners policy and makes expected profits $\pi = p - l$. Because the relevant variation in costs for the model comes from differences in risk, we normalize a, ϕ', and f in Equation 4 to zero. We assume that firms can charge different prices to different consumers, but that a given firm charges the same price to all consumers with the same modeled risk.[36]

Consumers buy one unit of insurance to maximize utility $u = I(0, \delta) - p$. The total number of consumers within the risk segment is normalized to 1. Once a consumer purchases insurance coverage from a particular firm, they face costs of switching to other firms: $I(0, \delta)$ is an indicator function equal to $-\delta$ if the consumer switches insurers, 0 otherwise. Specifically, after an initial period in which consumers are matched with an insurer, they need to be offered a price more than δ below another firm's price to induce them to switch.

Similar to the Bertrand model introduced in Section 6.1, we specify a two-stage model. In an initial stage, all firms have the same information about how risk is distributed within a market segment, but firms do not know the specific risk of any individual property. In time t, a risk modeling technology (e.g., a CAT model) becomes available at cost F^c that provides perfect information about the risk of properties. We characterize the decision by a firm to adopt the technology in terms of the additional profits it can earn, i.e., the value of information.

7.2 Initial conditions

In the initial period, firms and consumers enter the de novo market. Firms compete in prices knowing only the risk distribution $U(0, l^*)$. In equilibrium, firms make zero expected profits, which occurs at the price $p_0 = \bar{l}$. Once consumers have selected an insurer, they have no incentive to switch insurers because there is a single price for insurance.

To approximate the market structure we observe, we assume that one firm captures a relatively large share α of the market and the remaining "fringe" of competing insurers share the rest of the market.[37] We assume that losses of the dominant firm's customers are distributed according to $\alpha U(0, l^*)$, implying that losses of the fringe's customers are distributed according to $(1 - \alpha)U(0, l^*)$.

36. This assumption is consistent with California's insurance regulations. Insurers must propose rate structures that amount to an algorithm mapping risk of a given property to a price for a homeowners policy.

37. The dominant firm may have "brand recognition" due to past advertising investments. Brand recognition does not give the firm a pricing advantage, but consumers may be more likely to select them when indifferent on the price dimension.

7.3 Market equilibrium with asymmetric information

In time t, a more sophisticated information technology becomes available. Here, we assume that the technology provides perfect information about risk. In Appendix F, we allow for the possibility that the technology provides imperfect information about a property's risk. The dominant firm decides whether to adopt this technology, weighing the cost of adoption and expected equilibrium profits.

7.3.1 Pricing by the dominant firm at the initial equilibrium

If the dominant firm adopts the technology, it can segment customers by risk and charge them different prices. At the initial equilibrium, the dominant firm will either set a price at $p^D = \bar{l} + \delta$ to earn positive expected profits on its existing customers or a price at $p^D = \bar{l} - \delta$ to capture all the customers.[38] The latter strategy is more profitable for low-risk customers provided δ is not too large. In particular, at $l = 0$, profits from selling to all customers, $\pi^D = \frac{1}{l^*}(\bar{l} - \delta)$, are larger than those from selling only to its original customers, $\pi^D = \frac{\alpha}{l^*}(\bar{l} + \delta)$, provided that $\bar{l} > \delta\frac{1+\alpha}{1-\alpha}$, an assumption we adopt hereafter.[39] It will be convenient below to define $\tau = \frac{\bar{l}}{\delta}$; thus, the assumption is restated as $\tau > \frac{1+\alpha}{1-\alpha}$. As risk rises, the difference in profits shrinks until at $\tilde{l} = \bar{l} - \delta\frac{1+\alpha}{1-\alpha}$, the dominant firm earns equal profits with the two pricing strategies. As shown by the red line in panel (a) of Figure 8, below $\tilde{l}$ the dominant firm prices at $p^D = \bar{l} - \delta$ and captures both types of consumers. Above $\tilde{l}$, it prices at $p^D = \bar{l} + \delta$ and sells only to its existing consumers. However, the strategy $p^D = \bar{l} + \delta$ is only profitable up to $l_1 = \bar{l} + \delta$ and so the dominant firm elects not to sell to any consumers with risk above l_1.

The dominant firm's pricing strategy yields profits given by:

$$\pi^D = \frac{1}{2}\left\{\bar{l}^2 - 2\delta\bar{l} + \frac{\delta^2}{(1-\alpha)^2}[1 + 2\alpha - 3\alpha^2]\right\}. \tag{12}$$

Given the dominant firm's pricing, the competitive fringe earns positive expected profits on the interval $[\tilde{l}, \bar{l}]$ and negative profits on the interval $(\bar{l}, l^*]$. Because the fringe firms cannot distinguish risks, they cannot avoid losing money on high-risk customers. In the aggregate, the fringe's profits are given by:

$$\pi^F = \frac{1}{2}\left\{\delta^2\frac{1+3\alpha}{1-\alpha} - \bar{l}^2\right\}. \tag{13}$$

Given the assumption $\tau > \frac{1+\alpha}{1-\alpha}$, the fringe's profits are negative. Thus, once the dominant firm has adopted the technology, the initial equilibrium cannot be supported.

38. We focus on pure strategy equilibria, which accords with the rate filings by insurers in California that map house characteristics to a single price.

39. We expect prices for insurance policies to be large relative to switching costs, suggesting $p_0 = \bar{l} \gg \delta$. As well, for market share values in Table 1, the term $\frac{1+\alpha}{1-\alpha}$ is at most 1.44.

7.3.2 Equilibrium and the value of information

We define market equilibrium in a risk segment as the set of prices that yield zero profits for the competitive fringe and maximum profits for the dominant firm conditional on the fringe's price. Formally, market equilibrium is given by:

Definition I: Market equilibrium is the set of prices p^F and $\mathbf{p}^D(p^F)$ such that (1) the competitive fringe earns zero profits at p^F and (2) $\mathbf{p}^D(p^F)$ is the best response by the dominant firm to the price p^F.

Because the competitive fringe cannot distinguish customers by risk, it charges a single price p^F. However, given its information advantage, the dominant firm can charge different prices to different customers and decline coverage to a subset of customers. The fringe firms have no such ability to limit coverage and so accept all customers willing to buy policies at the price p^F.

For the uniform risk distribution, the following proposition defines the market equilibrium:

Proposition I: If (a) $l \sim U(0, l^*)$ and (b) $\tau = \frac{\bar{l}}{\delta} > \frac{1+\alpha}{1-\alpha}$, then market equilibrium is given by: (1) $p^F = l^* - \delta\sqrt{\frac{1+3\alpha}{1-\alpha}}$ and (2) $p^D = p^F - \delta$ for $l \in [0, \tilde{l})$, $p^D = p^F + \delta$ for $l \in [\tilde{l}, l_1)$, and $p^D > p^F + \delta$ for $l \in [l_1, l^*]$, where $\tilde{l} = p^F - \delta\frac{1+\alpha}{1-\alpha}$ and $l_1 = p^F + \delta$.
Proof: See Appendix F.

Condition (b) ensures that the dominant firm can set a low enough price on the interval $[0, \tilde{l})$ to take all the low-risk customers. This requires that δ and/or α not be too large.[40] On the interval $[l_1, l^*]$, the dominant firm ensures it gets no customers by charging any price above $p^F + \delta$.[41]

The market equilibrium is represented in panel (b) of Figure 8. It can be shown that $p^F > \bar{l}$, indicating that the competitive fringe must raise its price in order to break even in a risk segment with a better-informed dominant firm. The dominant firm exploits its information advantage by serving low-risk customers and declining coverage to high-risk customers.[42] It earns positive profits given by:

$$\pi^D = \frac{1}{2}\left\{\tilde{l}(2p^F - 2\delta - \tilde{l}) + \alpha(p^F + \delta - \tilde{l})^2\right\} > 0. \tag{14}$$

The competitive fringe, on the other hand, sells money-losing policies to the high-risk portion of the segment and only breaks even with profitable policies sold to medium-risk customers. Its profits

40. The competitive fringe can only earn non-negative profits at a price $p^F > \bar{l}$. Together with condition (b), this implies $\tilde{l} > 0$ and that a two-tiered pricing strategy is optimal for the dominant firm.

41. In practice, an insurer would simply decline coverage to particular customers.

42. Although the lowest-risk consumers pay the lowest prices, the dominant firm faces no competition for these consumers as long as its price is at or below $p^F - \delta$. In practice, a firm like State Farm is likely to face competition from other large, well-informed firms (e.g., Farmers) operating in a given market segment. Although we do not formally consider this extension, we would expect competition among well-informed firms to drive down prices closer to expected cost (l in this model), yielding a positive relationship between price and risk at low prices. This is consistent with Radner (2003), who finds in a model with sticky adjustment of consumers among firms ("viscous demand") that duopoly equilibria tend to be more competitive than the monopoly outcome.

are given by:

$$\pi^F = \frac{1}{2}\left\{(1-\alpha)(p^F - \tilde{l})^2 - (l^* - p^F)^2 + \alpha\delta^2\right\} = 0. \tag{15}$$

Since the dominant firm makes zero profits when it does not adopt the technology, the value of information (VOI) is given by the equilibrium profits in Equation 14. Although the VOI is always positive, it is only profitable for the dominant firm to adopt the technology when $VOI > F^c$. Since VOI is a function of market share, we consider how it varies with α. In Appendix F, we show that for a given value of τ, there is always a range of market shares at which $\frac{\partial VOI}{\partial \alpha} > 0$. A positive relationship between VOI and α is consistent with the data in Table 1 showing that insurers with larger market shares use more granular risk classification methods. Thus, the model can offer an explanation for the incomplete adoption of sophisticated risk rating technologies.

The equilibrium model of asymmetric information generates several predictions that are consistent with empirical findings and analysis presented throughout the paper: (1) The high-information firm uses a more granular pricing strategy than low-information firms; (2) The high-information firm uses its superior information to offer lower prices to low-risk customers within a risk segment; (3) Low-information firms set higher prices to avoid selling money-losing policies to high-risk consumers;[43] (4) The incentive to adopt a more granular risk classification strategy is increasing in the market share of the dominant firm under certain conditions.

7.3.3 Average prices in a market with asymmetric information

The adoption of the information technology by the dominant firm affects equilibrium insurance prices. A key question is whether the average price paid by consumers increases as a result of asymmetric information. The average price of insurance following adoption is given by:

$$\bar{p} = (p^F - \delta)\frac{\tilde{l}}{l^*} + \alpha(p^F + \delta)\frac{l_1 - \tilde{l}}{l^*} + (1-\alpha)p^F\frac{l_1 - \tilde{l}}{l^*} + p^F\frac{l^* - l_1}{l^*}. \tag{16}$$

We show in the following proposition that $\bar{p}$ is greater than the average price under the initial equilibrium in Section 7.2:

Proposition II: The average price of insurance, $\bar{p}$, under the market equilibrium with technology adoption (Proposition I) is greater than the average price under the original market equilibrium ($p_0 = \bar{l}$).
Proof: See Appendix F.

In this market in which demand for insurance is perfectly inelastic, an increase in firm profits must be accompanied by an increase in average price and, hence, a wealth transfer from consumers to firms.

43. In Appendix F, we derive the market equilibrium when the risk information acquired by the dominant firm is imperfect. We show that the "low/high" pricing strategy by the dominant firm (see Figure 8) remains optimal for a large set of parameters and that the fringe price gets bid up above the mean $\bar{l}$.

7.4 Equilibrium with price regulation

In California, the extent to which an insurer can raise its average price – evaluated across a book of business – has been somewhat constrained by historical loss experience. In this section, we explore the implications of price regulation for a market with asymmetric information. Of particular interest is how regulation affects insurance availability and how public information provision affects prices and availability. We approximate price regulation by imposing an upper bound on the average prices charged by the dominant firm, noting that in practice insurers can choose to incur the significant costs of appealing to exceed the regulatory threshold. We continue to examine a single risk segment and assume a uniform risk distribution.

7.4.1 The effect of price regulation on market equilibrium

We consider formally the effect of a price constraint on the dominant firm. The average price of the dominant firm is given by:

$$\bar{p}^D = \eta(p^F - \delta) + (1 - \eta)(p^F + \delta), \tag{17}$$

where $\eta = \frac{\tilde{l}}{\tilde{l}+\alpha(p^F+\delta-\tilde{l})}$. Suppose that the firm's average price cannot exceed $\bar{p}^R$ under the regulation. Then the maximization problem for the dominant firm is:

$$\max_{\tilde{l}} \pi^D \quad s.t. \quad \bar{p}^D \leq \bar{p}^R, \tag{18}$$

where π^D and $\bar{p}^D$ are given in Equations 14 and 17, respectively. The maximization is over $\tilde{l}$, the risk level at which the dominant firm switches from selling to the whole market at $p^F - \delta$ and selling to the share α of the market at $p^F + \delta$. When the constraint is binding, $\bar{p}^D = \bar{p}^R$, and we can use Equation 17 to derive the chosen value of $\tilde{l}$ as:

$$\tilde{l} = \frac{\bar{p}^R \alpha(p^F + \delta) - \alpha(p^F + \delta)^2}{p^F(1 - \alpha) - \delta(1 + \alpha) - \bar{p}^R(1 - \alpha)}. \tag{19}$$

The regulation forces the dominant firm to depart from the unconstrained pricing rule under which the low and high prices earn the same profits at $\tilde{l}$.

The fringe's profits change when the dominant firm adjusts $\tilde{l}$. Thus, to define the market equilibrium, we need a new value of p^F that makes the fringe's profits equal to zero. In general, this value is given by[44]:

$$p^F = \frac{l^* - \tilde{l}(1 - \alpha)}{\alpha} - \frac{1}{\alpha}\sqrt{(1 - \alpha)(l^* - \tilde{l})^2 + \alpha^2\delta^2}. \tag{20}$$

For a given value of $\bar{p}^R$, the constrained market equilibrium is given by the values $\{p^{F*}, \tilde{l}^*\}$ that satisfy Equations 19 and 20. The equilibrium can be illustrated with isoclines in $\{p^F, \tilde{l}\}$ space

44. In Equation 20, p^F is the solution to a quadratic equation. We can rule out one of the solutions because it implies a value of p^F that exceeds l^*.

corresponding to the zero profit and average price conditions, as depicted in Figure 9. We show in Appendix F that at the unconstrained equilibrium[45] the isoclines are upward sloping and that the relative magnitude of the average price and zero profit isocline slopes is indeterminate. Numerical analysis shows that except for small values of τ, requiring large values of δ relative to $\bar{l}$, the zero profit isocline has a steeper slope than the average price isocline, as shown in Figure 9. We adopt this as the empirically relevant case.

7.4.2 Regulation and insurance availability

Due to the randomness of catastrophic wildfire events, at any given time, some firms will accumulate loss experience that can be used to rationalize a price increase, while others will not. This unevenness has implications for insurance availability in a market with asymmetric information. As shown in Section 7.3.2, once the dominant firm adopts and uses superior risk information, the fringe firms must raise their prices to remain profitable. If they cannot raise prices sufficiently due to the regulatory constraint, their profits will be negative and they will exit the risk segment.

Fringe firms may also be forced to exit when the dominant firm raises its average price. Suppose that the dominant firm experiences losses that relax the constraint on $\bar{p}^R$. Relaxing the constraint shifts up the $\bar{p}^D$ isocline, as shown in Figure 9. When the constraint on the dominant firm's average price is relaxed ($\bar{p}_1^R > \bar{p}_0^R$), the equilibrium values of p^F and $\tilde{l}$ increase. Profits for the fringe are increasing in the fringe's price p^F (see Equation 15), implying that a regulatory constraint on the fringe's price can be at most weakly binding, because otherwise fringe profits are negative and fringe firms will exit the risk segment. Thus, as long as fringe firms cannot raise their own prices, relaxing the regulatory constraint on the dominant firm will have the effect of lowering availability as fringe firms decline coverage to their current customers. After the exit of the fringe, the dominant firm would want to pick up some, but not all, of the fringe's customers if its price remains constrained at $\bar{p}_1^R$.

7.4.3 Improving access to risk information

As the aforementioned analysis makes clear, regulations that limit firms' ability to increase premiums can trade off affordability for insurance availability. In contrast, the government could improve both affordability and availability by helping insurers access and use superior risk information in their pricing and underwriting decisions. We consider how the market equilibrium changes when fringe firms gain access to more granular risk information. In panels (c) and (d) of Figure 8, we allow fringe firms to distinguish properties with risk distributed as $U(0, \bar{l})$ or $U(\bar{l}, l^*)$, compared to the original case in which they know only that risk is distributed $U(0, l^*)$. If $\tilde{l} > \bar{l}$ under the original equilibrium, the overall average price unambiguously declines. Prices do not change on the interval

45. The unconstrained equilibrium is the values $\{p^{F*}, \tilde{l}^*\}$ such that $\tilde{l}^*$ is freely chosen in Equation 18. Associated with this value of $\tilde{l}^*$ is an average price $\bar{p}^D$ according to Equation 17.

$[\bar{l}, l^*]$ and are lower on the interval $[0, \bar{l})$.[46] In this scenario, a policy intervention that improves access to better information can improve affordability and availability insofar as average premiums fall and the fringe firms can remain in segments of the market that would otherwise be associated with negative profits. Under the alternative scenario wherein $\tilde{l} < \bar{l}$, the effect on the overall average price is unclear. However, Appendix F provides numerical results showing that average price falls for a large range of parameter values. Although the cost of information provision would need to be considered, our results suggest that this may be a more effective way to address affordability and availability objectives than price constraints.

8 Conclusion

In the face of escalating climate risk, well-functioning property insurance markets can provide households and businesses with crucial protection from economic losses as well as incentives to reduce risk exposure. However, there are signs that property insurance markets are struggling to adapt to climate change pressures. This study investigates some of the reasons that climate change risk presents significant challenges for property insurance markets, with a focus on wildfire risk and homeowners insurance in California.

Our analysis combines proprietary parcel-level wildfire risk analytics with information in insurers' public rate filings to analyze how insurers are pricing wildfire risk in California. We document significant variation in the information insurers use to price wildfire risk. We also find economically significant differences between the price schedules we observe and cost-based benchmarks that account for expected losses, operating costs, and a loading factor that reflects the costs of ensuring insurer solvency. We show that the empirical evidence is consistent with a form of adverse selection in which less-informed firms face a winner's curse in high-wildfire risk market segments.

Motivated by the empirical evidence on asymmetric information and economic regulation, we develop an equilibrium model of the property insurance market that incorporates both information asymmetries and binding economic regulation. In the model, insurers can access detailed risk information through the costly adoption of sophisticated modeling tools. We show that if the costs of adopting and using more sophisticated risk information are sufficiently high, there are conditions under which only the firms with the largest market shares will adopt the information. This is consistent with what we observe in the California market. The model also predicts that the high-information firm will use its superior information to win the lower-risk customers within a risk segment, and that low-information firms will set high prices to mitigate effects of adverse selection. If regulation prevents upward price adjustments, insurers wary of the winner's curse may exit high-risk market segments to limit their exposure. Thus, in markets characterized by asymmetric information, regulations limiting premium increases can have unintended consequences. Policies

46. On the interval $[0, \bar{l})$, the fringe price is $p^F = \bar{l} - \delta\sqrt{\frac{1+3\alpha}{1-\alpha}}$. Since the original fringe price is greater than $\bar{l}$, we know that all prices are lower on $[0, \bar{l})$, implying a lower overall average price.

that improve market-wide understanding of wildfire risk could improve affordability without sacrificing availability.

Our findings are relevant to current policy discussions about property insurance market reform in California and elsewhere. As of 2025, several leading insurers, including State Farm and Allstate, have begun to limit the writing of new policies and tighten underwriting standards for existing customers. Major insurers have been requesting rate increases in excess of 20 to 30 percent. Wildfire risk exposure is just one of many factors driving these developments. Other factors include increases in non-catastrophe liability claims and construction cost inflation. Our results highlight the underappreciated importance of wildfire risk information, and more specifically, the winner's curse, as a barrier to participation in insurance markets for large, hard-to-model risks. Further investigation of the potential for adverse selection, the implications for insurance pricing and underwriting, and the policy changes that could be warranted will be critical to informing insurance market policies in an era of changing climate.

References

Abraham, Kenneth S. 1985. "Efficiency and Fairness in Insurance Risk Classification." *Virginia Law Review,* 403–451.

Aizawa, Naoki, and Ami Ko. 2023. *Dynamic Pricing Regulation and Welfare in Insurance Markets.* National Bureau of Economic Research.

Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84 (3): 488–500.

Baylis, Patrick W, and Judson Boomhower. 2022. *Mandated vs. Voluntary Adaptation to Natural Disasters: The Case of U.S. Wildfires.* National Bureau of Economic Research.

Bohrnstedt, George W., and Arthur S. Goldberger. 1969. "On the Exact Covariance of Products of Random Variables." *Journal of the American Statistical Association* 64 (328): 1439–1442.

Borenstein, Severin. 1989. "The Economics of Costly Risk Sorting in Competitive Insurance Markets." *International Review of Law and Economics* 9 (1): 25–39.

Born, P., and W.K. Viscusi. 2006. "The Catastrophic Effects of Natural Disasters on Insurance Markets." *Journal of Risk Uncertainty* 33:55—72.

Born, Patricia H., and Barbara Klimaszewski-Blettner. 2013. "Should I Stay or Should I Go? The Impact of Natural Disasters and Regulation on U.S. Property Insurers' Supply Decisions." *The Journal of Risk and Insurance* 80 (1): 1–36.

Buechi, Hanna, Paige Weber, Sarah Heard, Dick Cameron, and Andrew J Plantinga. 2021. "Long-term trends in wildfire damages in California." *International Journal of Wildland Fire* 30 (10): 757–762.

Cather, David A. 2018. "Cream Skimming: Innovations in Insurance Risk Classification and Adverse Selection." *Risk Management and Insurance Review* 21 (2): 335–366.

Chamberlin, John R. 1985. "Assessing the Fairness of Insurance Classifications." *Research in Law and Economics* 7:65–87.

Cignarale, T., Joel Laucher, Ken Allen, and L. Landsman-Smith. 2019. *The Availability and Affordability of Coverage for Wildfire Loss in Residential Property Insurance in the Wildland-Urban Interface and Other High-Risk Areas of California: CDI Summary and Proposed Solutions.* California Department of Insurance.

Crocker, Keith J., and Arthur Snow. 1986. "The Efficiency Effects of Categorical Discrimination in the Insurance Industry." *Journal of Political Economy* 94 (2): 321–344.

Dietz, Simon, and Falk Niehörster. 2021. "Pricing ambiguity in catastrophe risk insurance." *The Geneva Risk and Insurance Review* 46 (2): 112–132.

Dixon, Lloyd, Flavia Tsang, and Gary Fitts. 2018. *The Impact of Changing Wildfire Risk on California's Residential Insurance Market: A Report for California's Fourth Climate Change Assessment.* California Natural Resources Agency.

Einav, Liran, and Amy Finkelstein. 2011. "Selection in Insurance Markets: Theory and Empirics in Pictures." *Journal of Economic Perspectives* 25 (1).

Einav, Liran, Amy Finkelstein, and Neale Mahoney. 2021. "The IO of selection markets." In *Handbook of Industrial Organization,* 5:389–426. 1. Elsevier.

Einav, Liran, Mark Jenkins, and Jonathan Levin. 2012. "Contract Pricing in Consumer Credit Markets." *Econometrica* 80 (4): 1387–1432.

Eling, M., I. Gemmo, and D. Guxha. 2024. "Big Data, Risk Classification, and Privacy in Insurance Markets." *Geneva Risk Insurance Review* 49:75–126.

Engelbrecht-Wiggans, Richard, Paul R Milgrom, and Robert J Weber. 1983. "Competitive bidding and proprietary information." *Journal of Mathematical Economics* 11 (2): 161–169.

First Street Foundation. 2022. *5th National Risk Assessment Fueling the Flames.* First Street Foundation.

Ge, Shan. 2022. "How Do Financial Constraints Affect Product Pricing? Evidence from Weather and Life Insurance Premiums." *Journal of Finance* 77 (1): 449–503.

Goodman, Leo A. 1960. "On the Exact Variance of Products." *Journal of the American Statistical Association* 55 (292): 708–713.

Gourevitch, J.D., C. Kousky, and Y. Liao. 2023. "Unpriced climate risk and the potential consequences of overvaluation in U.S. housing markets." *Nature Climate Change.*

Handel, Ben, Igal Hendel, and Michael D Whinston. 2015. "Equilibria in Health Exchanges: Adverse Selection Versus Reclassification Risk." *Econometrica* 83 (4): 1261–1313.

Heiss, Florian, Daniel McFadden, Joachim Winter, Amelie Wuppermann, and Bo Zhou. 2021. "Inattention and switching costs as sources of inertia in Medicare part D." *American Economic Review* 111 (9): 2737–2781.

Hendricks, Kenneth, and Robert H Porter. 1988. "An empirical study of an auction with asymmetric information." *American Economic Review,* 865–883.

Honka, Elisabeth. 2014. "Quantifying search and switching costs in the U.S. auto insurance industry." *The RAND Journal of Economics* 45 (4): 847–884.

Hoy, Michael. 1982. "Categorizing Risks in the Insurance Industry." *The Quarterly Journal of Economics* 97 (2): 321–336.

Iglesias, Virginia, Jennifer K. Balch, and William R. Travis. 2022. "U.S. Fires Became Larger, More Frequent, and More Widespread in the 2000s." *Science Advances* 8 (11): eabc0020.

Jaffee, Dwight M, and Thomas Russell. 1997. "Catastrophe Insurance, Capital Markets, and Uninsurable Risks." *Journal of Risk and Insurance,* 205–230.

Jin, Yizhou, and Shoshana Vasserman. 2021. *Buying Data from Consumers: The Impact of Monitoring Programs in U.S. Auto Insurance.* National Bureau of Economic Research.

Keys, Benjamin J, and Philip Mulder. 2024. *Property Insurance and Disaster Risk: New Evidence from Mortgage Escrow Data.* National Bureau of Economic Research.

Klein, Robert W., and Paul R. Kleindorfer. 1999. *The Supply of Catastrophe Insurance Under Regulatory Constraints.* Center for Financial Institutions Working Paper 99-25. Wharton School Center for Financial Institutions, University of Pennsylvania.

Koijen, Ralph S. J., and Motohiro Yogo. 2015. "The Cost of Financial Frictions for Life Insurers." *American Economic Review* 105 (1): 445–75.

Kousky, Carolyn, and Roger M Cooke. 2009. *The Unholy Trinity: Fat Tails, Tail Dependence, and Micro-correlations.* Resources for the Future 09-36.

Kreps, Rodney. 1990. "Reinsurer Risk Loads From Marginal Surplus Requirements." In *Proceedings of the Casualty Actuarial Society,* 77:196–203. Citeseer.

Kunreuther, Meszaros, Hogarth, and Spranca. 1995. "Ambiguity and underwriter decision processes." *Journal of Economic Behavior & Organization* 26 (3): 337–352.

Kunreuther, Michel-Kerjan, Doherty, Grace, Klein, and Pauly. 2009. *At War with the Weather: Managing Large-Scale Risks in a New Era of Catastrophes.* The MIT Press.

Kunreuther, Howard. 1996. "Mitigating Disaster Losses Through Insurance." *Journal of Risk and Uncertainty* 12:171–187.

Kunreuther, Howard C, and Erwann O Michel-Kerjan. 2011. *At War with the Weather: Managing Large-Scale Risks in a New Era of Catastrophes.* MIT Press.

Marcoux, Kendra, and Katherine R. H. Wagner. 2024. *Fifty Years of U.S. Natural Disaster Insurance Policy.* Technical report 10431. CESifo Working Paper.

Milgrom, Paul R, and Robert J Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica: Journal of the Econometric Society,* 1089–1122.

Moore, Frances. 2024. *Learning, Catastrophic Risk, and Ambiguity in the Climate Change Era.* National Bureau of Economic Research.

Mumpower, Jeryl L. 1991. "Risk, Ambiguity, Insurance, and the Winner's Curse." *Risk Analysis* 11 (3): 519–522.

Oh, Sangmin, Ishita Sen, and Ana-Maria Tenekedjieva. 2024. *Pricing of Climate Risk Insurance: Regulation and Cross-Subsidies.* Federal Reserve Board Finance and Economics Discussion Series.

Radner, Roy. 2003. "Viscous Demand." *Journal of Economic Theory* 112 (2): 189–231.

Scott, Joe H., Gregory K. Dillon, Michael N. Callahan, Melissa R. Jaffe, Kevin C. Vogler, Julia H. Olszewski, Eva C. Karau, Mitchell T. Lazarz, Karen C. Short, Karin L. Riley, Mark A. Finney, and Isaac C. Grenfell. 2024. *Wildfire Risk to Communities: Spatial Datasets of Landscape-Wide Wildfire Risk Components for the United States, Second Edition.* Forest Service Research Data Archive.

Shi, Linda, and Susanne Moser. 2021. "Transformative Climate Adaptation in the United States: Trends and Prospects." *Science* 372 (6549): eabc8054.

Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65 (3): 557–586.

Stone, James M. 1973. "A Theory of Capacity and the Insurance of Catastrophe Risks (Part I)." *Journal of Risk and Insurance,* 231–243.

Summers, J. K., A. Lamper, C. McMillion, and L. C. Harwell. 2022. "Observed Changes in the Frequency, Intensity, and Spatial Patterns of Nine Natural Hazards in the United States from 2000 to 2019." *Sustainability* 14 (7).

Taylor, Reid, Madeline Turland, and Joakim A. Weill. 2023. *Climate Change and the Regulation of a Crashing Insurance Market.* Working paper.

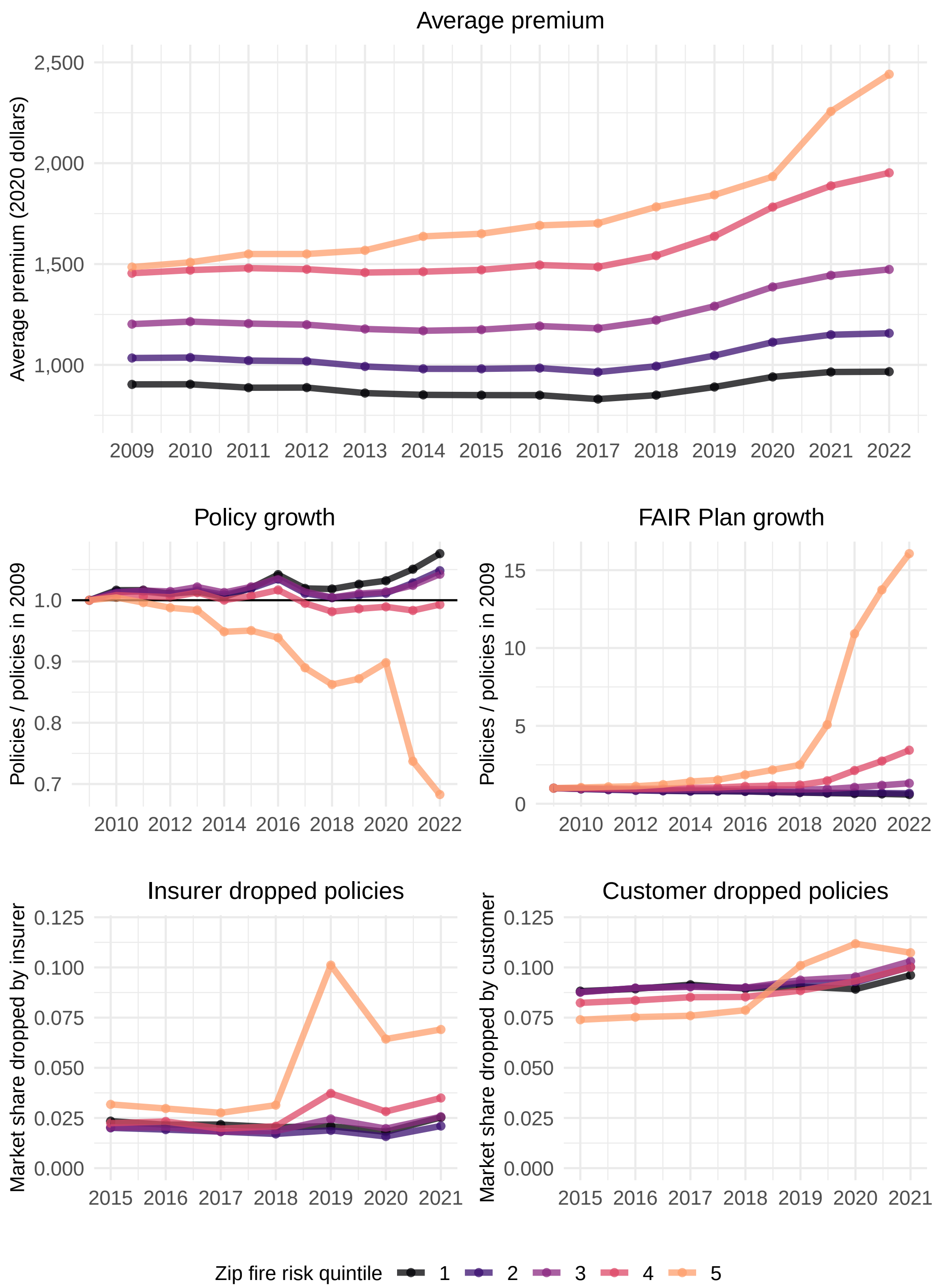
U.S. Department of the Treasury. 2023. *Insurance Supervision and Regulation of Climate-Related Risks.* Federal Insurance Office, U.S. Department of the Treasury.

Viscusi, W Kip, Joseph E Harrington Jr, and David EM Sappington. 2018. *Economics of Regulation and Antitrust.* MIT press.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* MIT press.

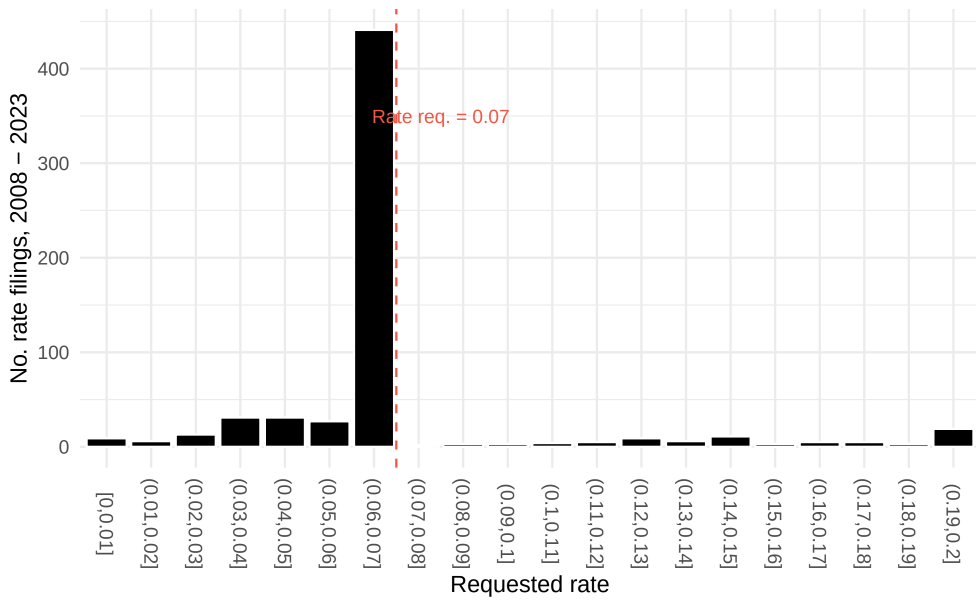
Wylie, David, Natee Amornsiripanitch, John Heilbron, and Kevin Zhao. 2024. *Who Bears Climate-Related Physical Risk?* Federal Reserve Bank of Philadelphia.

Figure 1: California homeowners insurance by year



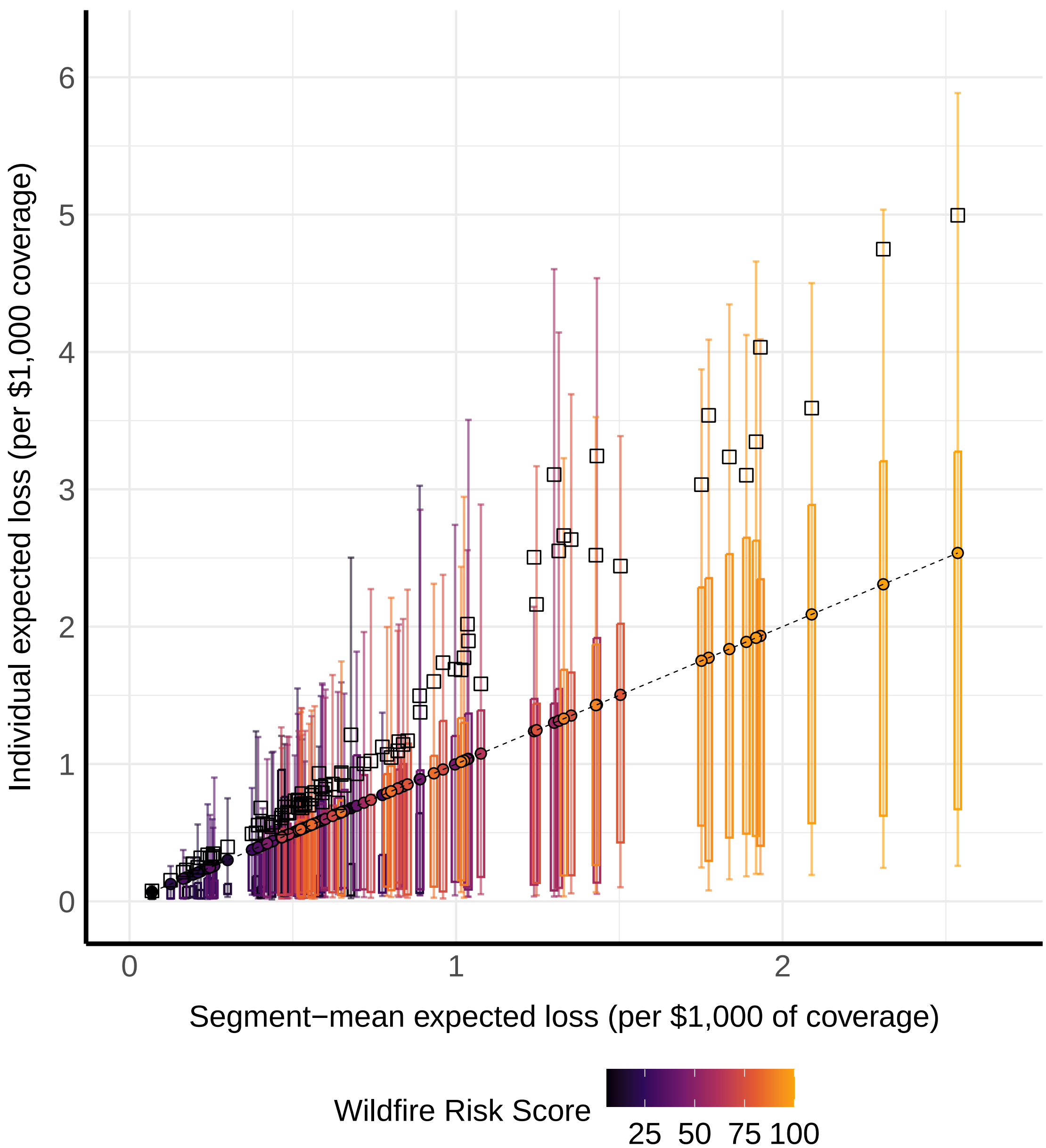
Notes: Figure summarizes zip code-level outcomes over time. Zip codes are classified into wildfire risk quintiles on the basis of zip code average risk. Risk scores are derived from California Department of Insurance Wildfire Risk Information Reporting, which lists proportions of insurance policies in various hazard categories (Negligible = 0, Low = 1, Moderate = 2, High = 3, Very High = 4). Annual zip code-level premiums, policy growth, and dropped policies are reported by the California Department of Insurance. Premium is in 2020 dollars.

Figure 2: Rate filing behavior



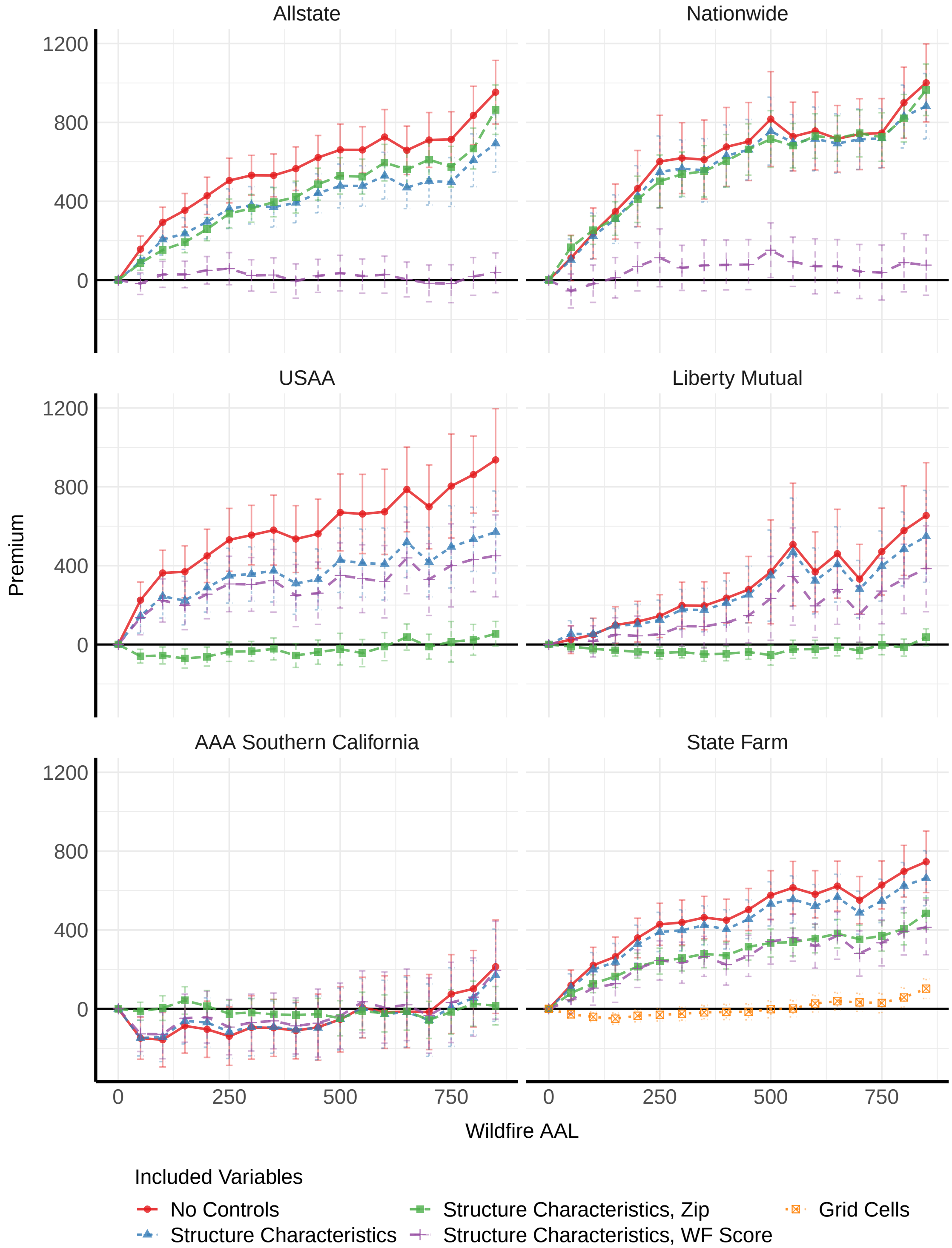
Notes: Figure displays the distribution of all 636 requested rate increases to the California Department of Insurance from 2008 to 2023 for owner-occupied homeowners' insurance (HO-3) policies for all insurers in California. Rate filings are available via the Web Access to Rate and Form Filings (WARFF) system.

Figure 3: Illustrative per-customer costs under symmetric and asymmetric risk classification



Notes: Box plots show the distribution of individual expected losses within each wildfire risk score. Whiskers show the 10th and 90th percentiles; notches show the interquartile range; circular markers are means. Horizontal axis plots the mean of expected losses in each wildfire risk score. The square markers show Firm A's expected per-customer costs in the illustrative asymmetric classification example in Section 4.4.

Figure 4: Validating firms' reported pricing variables

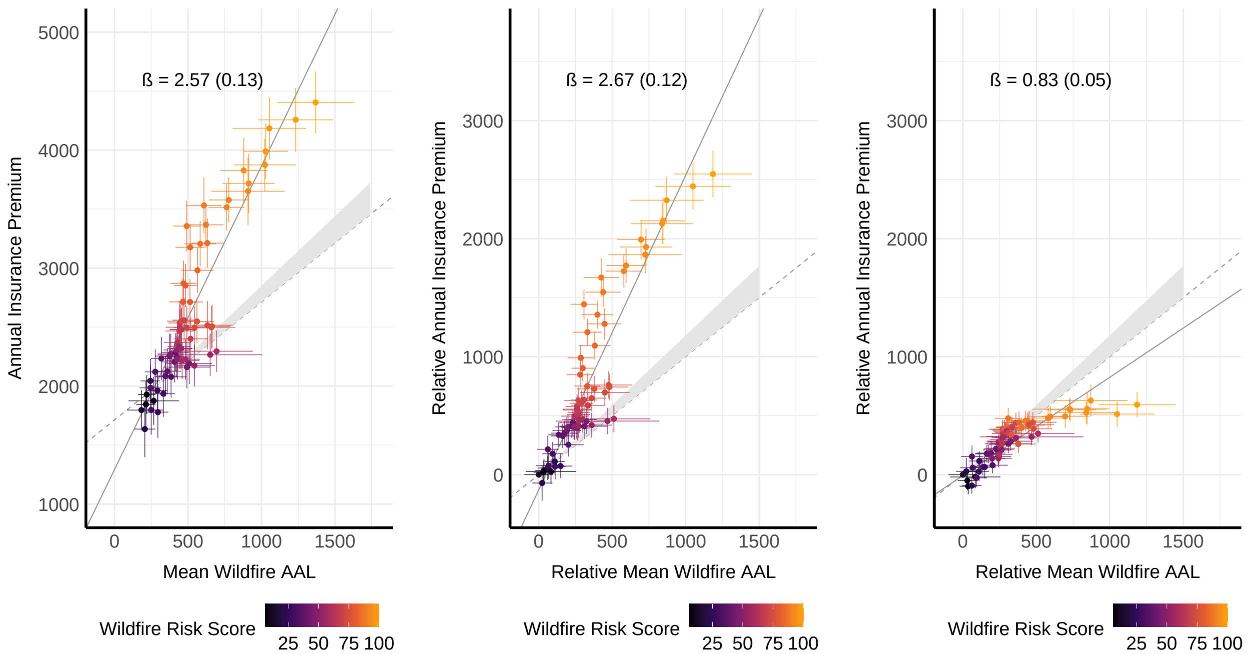


Notes: Each panel shows coefficients from multiple separate regressions of annual premium on a binned specification of wildfire AAL. All regressions (including "no controls") include a seventh-degree polynomial in demeaned reconstruction cost. Standard errors in all cases are clustered by zip code. The red lines summarize the average relationship between insurance premiums and AAL, holding constant reconstruction cost. The blue lines add controls for structure characteristics (age of home, public protection class, and an indicator for Class A roof). The green lines control for structure characteristics and zip code dummies. The purple lines control for structure characteristics and dummies for CoreLogic wildfire risk scores.

Figure 5: Wildfire price gradients for Allstate and Nationwide

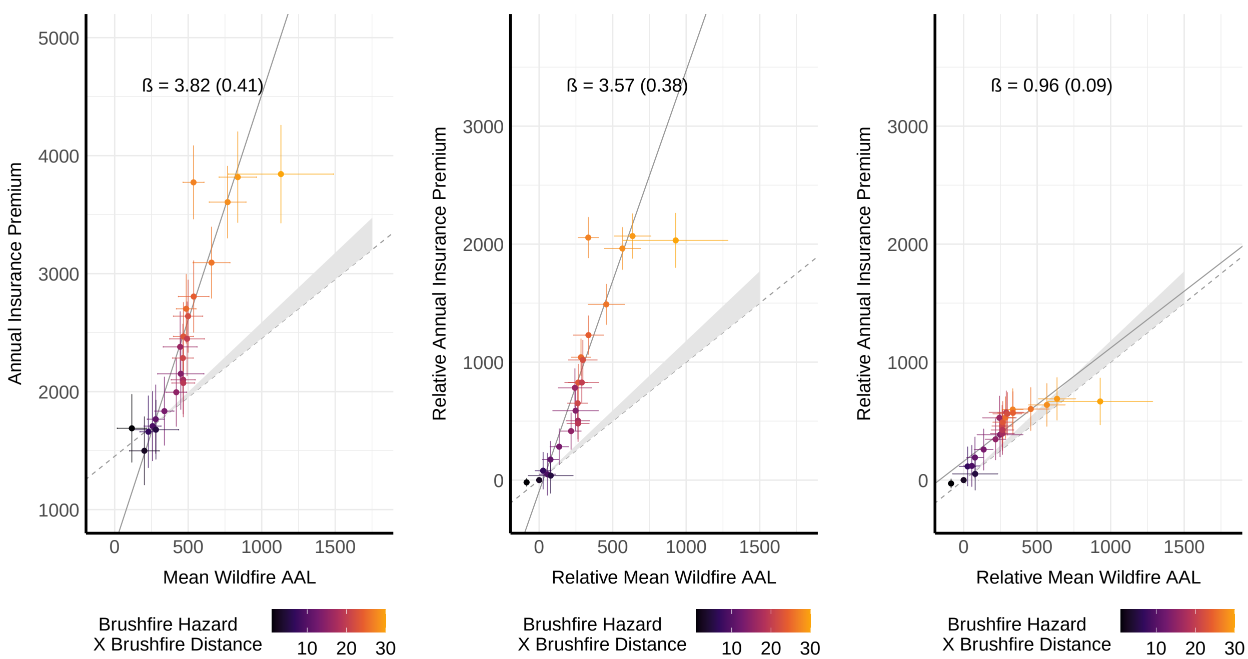
Panel A. Allstate

(a) Allstate, no FE | (b) Allstate, with FE | (c) State Farm, with FE



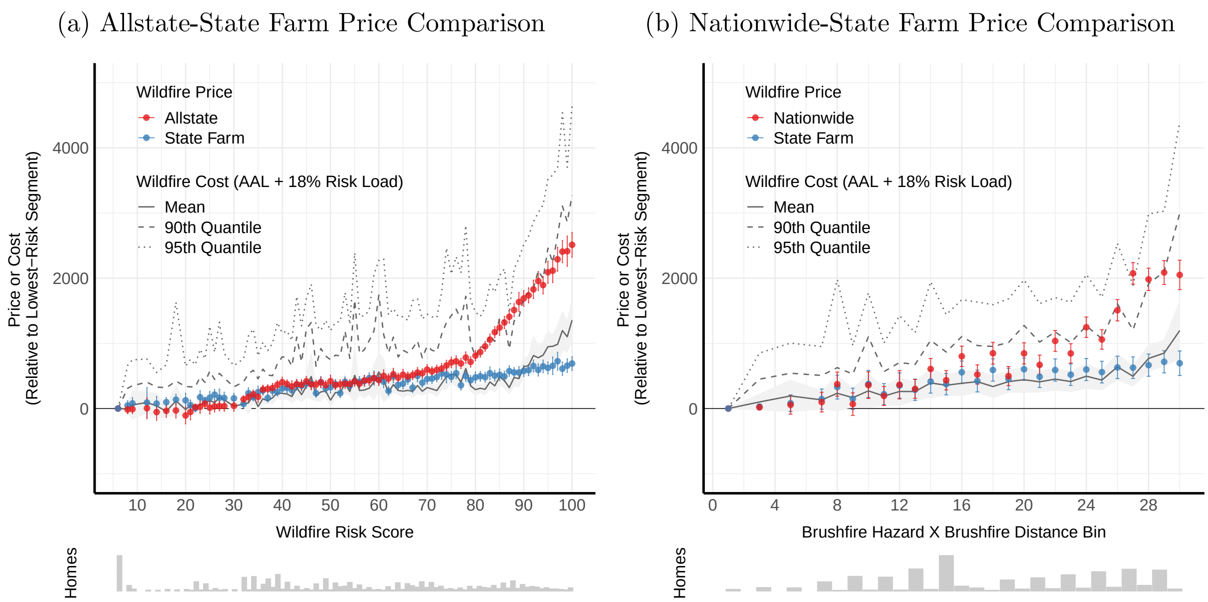
Panel B. Nationwide

(d) Nationwide, no FE | (e) Nationwide, with FE | (f) State Farm, with FE



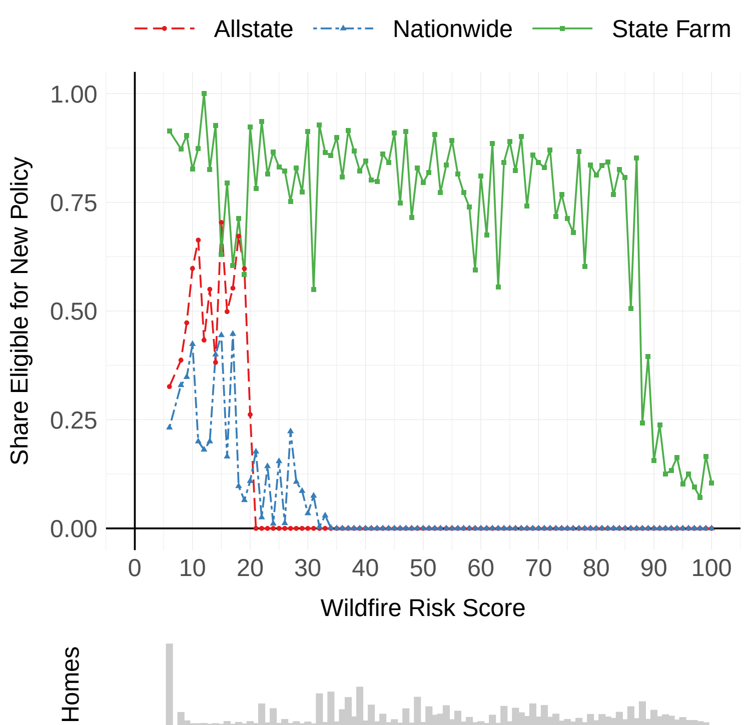
Notes: Each panel reports separate regressions of annual premium and wildfire AAL on wildfire score dummies, with controls for reconstruction cost and other factors as described in the text. The regressions used to create the figures in the second and third columns also include zip code fixed effects. Vertical bars represent confidence intervals for premiums; horizontal bars represent confidence intervals for mean AAL. Standard errors are clustered by zip code to allow for arbitrary within-zip code shocks to residuals. The gray shaded regions show slopes between 1 and 1.18, the symmetric classification cost-based benchmark.

Figure 6: Price and Cost by Risk Segment



Notes: Figures compare price and cost across risk segment bins, holding constant the controls in the main regressions of the paper. The colored markers denote the average wildfire prices charged by Allstate, Nationwide, and State Farm within each WRS category. The solid black line shows the mean wildfire AAL, multiplied by 1.18 to reflect assumed risk load. The dashed lines show the 90th and 95th conditional quantiles of 1.18 × wildfire AAL in each segment. The gray histograms show the share of homes in each risk segment.

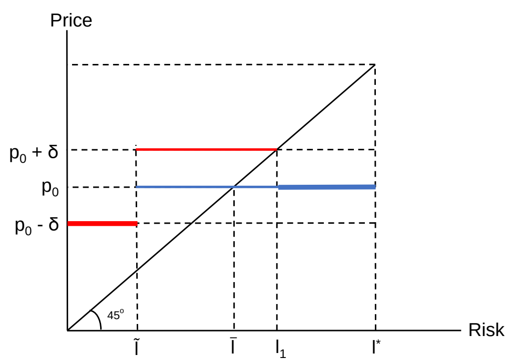
Figure 7: New policy eligibility versus parcel wildfire scores



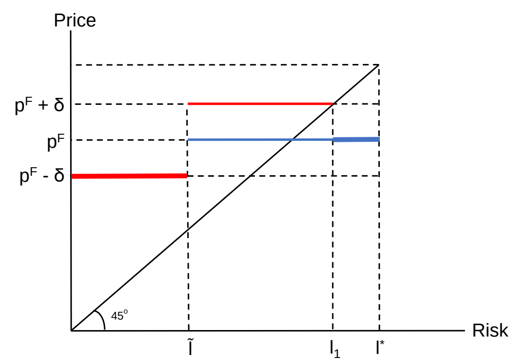
Notes: Figure reports the fraction of homes in each risk score bin that would have been eligible for a new homeowners insurance policy in 2021 from each firm.

Figure 8: Equilibrium pricing by dominant firm and competitive fringe

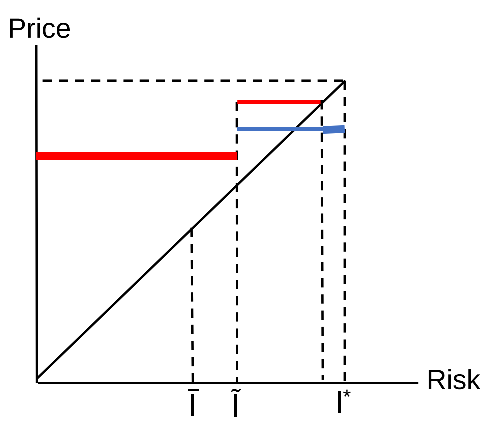
(a) Dominant firm best response at initial p_0



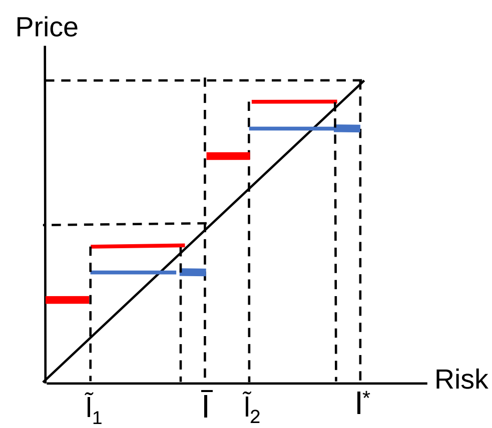
(b) Equilibrium when fringe profit is zero



(c) Baseline, limited information for fringe



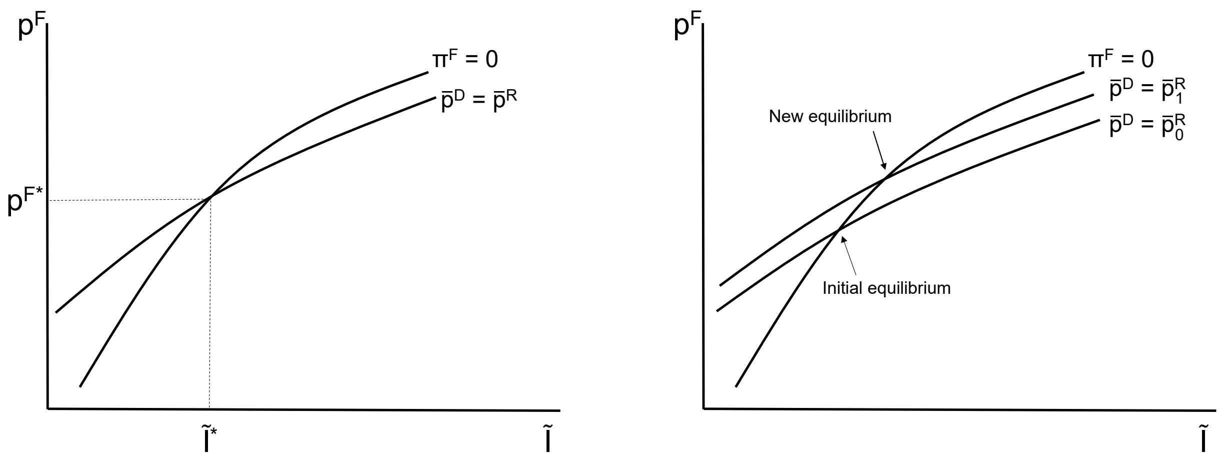
(d) Information provision



Notes: Panel (a) shows the dominant firm's best response at the initial equilibrium price p_0. The price charged by the dominant firm at each risk level l is shown in red, where the width of the line indicates the firm's market share (the widest line corresponds to a share of one). The competitive fringe's price and market shares are shown in blue. Panel (b) shows the market equilibrium prices at which the profits of the competitive fringe are zero. Panels (c) and (d) show how equilibrium prices change when new information is provided to the competitive fringe firms. Panel (c) is the baseline case in which fringe firms know only that properties are distributed according to $U(0, l^*)$ and for the case $\tilde{l} > \bar{l}$. In panel (d), the firms can distinguish whether properties are distributed according to $U(0, \bar{l})$ or $U(\bar{l}, l^*)$.

Figure 9: Market equilibrium with regulatory price constraints

(a) Isoclines for profit and price conditions (b) Equilibrium with relaxed constraint



Notes: Panel (a) shows the isoclines corresponding to the zero profit and average price conditions in Equations 19 and 20. The intersection of the isoclines defines the equilibrium for $\bar{p}^D = \bar{p}^R$. Panel (b) shows how the equilibrium changes when the average price constraint on the dominant firm is relaxed ($\bar{p}_1^R > \bar{p}_0^R$).

Table 1: Insurer market shares and granularity of wildfire rating

Insurer	Market Share (Percent)		Wildfire Hazard Variables	Hypothetical Best Fit
	Statewide	High-Risk Zip Codes		
State Farm	18.0	18.4	434,252	0.82
Farmers	15.5	14.7	2,149	nd
CSAA	7.6	8.6	26,055	0.67
Mercury	7.1	0.9	2,248	nd
Auto Club Enterprises	6.9	0.2	22	0.13
Liberty Mutual	6.5	3.4	1,698	0.47
Allstate	5.8	3.3	107	0.18
USAA	5.3	5.9	845	0.43
Travelers	3.2	4.8	1,568	nd
Nationwide	2.5	2.5	45	0.16
FAIR Plan	2.5	20.4	726	nd
All Others	19.2	16.9		

Notes: Market shares are based on California Department of Insurance Community Service Statement exposures data for 2020 at the insurer group level for the HO insurance line plus FAIR Plan. High-hazard zip codes are those falling into the highest quintile of average wildfire risk as reported in the CDI Wildfire Risk Information Report for 2021. Wildfire hazard variables count the number of factors an insurer uses to capture the likelihood of wildfire occurrence; more information on wildfire hazard variables and references to specific rate filings are available in Appendix A. Hypothetical best fit is the R^2 from a regression of catastrophe model wildfire risk (average annual loss) on rating variable indicator variables using the 100,000 homes in the dataset. Firms with no data ("nd") use other proprietary variables such as Verisk FireLine scores, which were unavailable for analysis.

Table 2: Summarizing Structure Characteristics, Wildfire Risk, and Insurance Prices

	Mean	SD	p5	p95	N
Structure Characteristics					
Reconstruction Cost	594,671	290,183	270,724	1,191,194	95,352
Year Built	1976	21	1937	2006	95,352
Square Feet	2,135	954	953	3,965	88,503
Wildfire Risk Score	52	27	6	92	95,352
Wildfire Average Annual Loss	303	596	9	1,298	95,352
Insurance Price Schedules					
Allstate	2,362	1,443	924	5,181	95,352
Nationwide	2,466	1,735	866	5,661	95,351
Liberty Mutual	2,389	1,229	1,176	4,763	94,290
State Farm	2,309	1,201	1,007	4,735	95,351
USAA	2,470	1,704	998	5,724	95,297
AAA Southern CA	2,756	1,914	964	8,290	55,641
Within-Home Dispersion of Offered Prices					
Average of Offered Prices	2,419	1,337	1,070	5,219	95,352
Lowest Offered Price	1,805	927	808	3,675	95,352
Std. Dev. of Offered Prices	578	546	148	1,716	95,352
Range of Offered Prices	1,488	1,447	384	4,470	95,352

Notes: Units for reconstruction cost, average annual loss, and insurance price are dollars. Within-home dispersion of offered prices describes the variation in insurance prices for each individual home across the six insurers in the data.

Table 3: IV estimates of price gradients by firm

	(1)	(2)	(3)	(4)	(5)
	Segment Definition				
	Allstate Scores	Nationwide Scores	USAA Territories	Liberty Mutual Territories	State Farm Grid Cells
Own Price					
Segment-Mean AAL	2.87***	3.42***	2.17***	2.46***	1.1
	(0.26)	(0.43)	(0.40)	(0.41)	(0.14)
State Farm Price					
Segment-Mean AAL	0.95	0.95	1.01	0.91	
	(0.10)	(0.13)	(0.19)	(0.15)	
F-statistic	255.6	61.1	62.1	88.8	1694.2
Zip Code FE	Yes	Yes	No	No	No
Number of Segments	94	30	323	393	11946

Notes: Table reports estimates of β_{jk} from a two-step estimation procedure that first calculates regression-adjusted mean prices and wildfire AAL by segment following Equations 9 and 10, and then regresses these segment-level prices against segment-level mean wildfire risk. The first-step estimation of segment means includes zip code fixed effects for Allstate and Nationwide, where there is cross-cutting variation in zip codes and wildfire segments. The second-step regression of mean price on mean wildfire AAL is estimated by two-stage least squares to remove measurement error in segment-mean wildfire AAL due to sampling variation. See text for details. Standard errors are calculated by bootstrapping the full estimation procedure 500 times. Stars indicate statistical significance against a null of $1 \leq \beta_{jk} \leq 1.18$, the benchmark cost gradient.

Table 4: The winner's curse in stylized Bertrand duopoly

	log(Wildfire AAL)				
	Allstate	Nationwide	USAA	Liberty Mutual	AAA SoCal
	(1)	(2)	(3)	(4)	(5)
1[Win]	0.52***	0.33***	0.44***	0.39***	0.56***
	(0.08)	(0.08)	(0.03)	(0.04)	(0.09)
Structure Characteristics	✓	✓	✓	✓	✓
Allstate Segments FE	✓				
Nationwide Segments FE		✓			
USAA Segments FE			✓		
Liberty Segments FE				✓	
AAA SoCal Segments FE					✓
Observations	95,295	95,295	95,295	93,976	55,606
R^2	0.36	0.32	0.74	0.77	0.28
Mean AAL ($)	302.21	302.21	302.21	299.09	235.61
Fraction Won	0.53	0.55	0.50	0.44	0.50

Notes: 1[Win] is an indicator for the firm's price being less than or equal to State Farm's price for that customer. Dependent variable is the log of wildfire AAL. Structure characteristics include reconstruction cost, age of home, class A roof indicator, and public protection class. Standard errors are clustered by zip code.

Table 5: Winner's curse in Bertrand duopoly, alternative switching costs

δ	Allstate vs. State Farm	Nationwide vs. State Farm	USAA vs. State Farm	Liberty Mutual vs. State Farm	AAA SoCal vs. State Farm
0	0.52 (0.08)	0.33 (0.08)	0.44 (0.03)	0.39 (0.04)	0.56 (0.09)
50	0.50 (0.08)	0.33 (0.08)	0.40 (0.03)	0.38 (0.04)	0.56 (0.09)
100	0.47 (0.08)	0.35 (0.07)	0.37 (0.03)	0.36 (0.03)	0.55 (0.08)
500	0.21 (0.04)	0.22 (0.04)	0.14 (0.02)	0.12 (0.01)	0.27 (0.04)
∞	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)	0.00 (0.01)

Notes: Table summarizes estimates from twenty-five separate OLS regressions following the specification in Table 4. The table reports the coefficient and standard error for 1[Win]. In each regression, each home is initially randomly assigned to one of the duopolists in a "period zero." That home's perceived price for the other firm is incremented by the indicated switching cost. Customers then choose the firm with the lower perceived cost.

Online Appendix to: How Are Insurance Markets Adapting to Climate Change? Risk Classification and Pricing in the Market for Homeowners Insurance

Judson Boomhower, Meredith Fowlie,
Jacob Gellman, Andrew Plantinga

A Data Documentation

We compile several sources of data. The first data source, taken from the National Association of Insurance Commissioners (NAIC), provides state-level insurance company market shares and earned premium for the United States. The second data source was obtained from the California Department of Insurance (CDI) and contains zip code-level information on annual insurance premiums, coverage, dropped policies, and company market share. Next, we use property-level information from CoreLogic on house characteristics, categorical wildfire risk scores, and probabilistic wildfire risk distributions. Lastly, we use public insurer rate requests made to CDI to develop price and eligibility schedules for the property-level dataset.

A.1 National Association of Insurance Commissioners statewide data

Since 1974, NAIC has produced annual reports on the profitability of insurance lines by state.[47] Specifically, these data report profit on insurance transactions, which accounts for underwriting profit, investment gains, and federal taxes. We gather these reports for years representing 1985 through 2021. The data are used to produce Appendix Figure 1, which shows the profitability of HO insurance lines by state.

A.2 California Department of Insurance zip code data

We obtained three datasets from CDI in a Public Records Act request.[48] These data report zip code-level information on insurer market share, premium, number of policies, total coverage, deductible, insurer-initiated non-renewals, and customer-initiated non-renewals, for the years 2009 through 2020. In addition, we used a publicly available CDI dataset, Wildfire Risk Information Reporting, which breaks down the distribution of wildfire risk within California zip codes. Although these datasets do not capture sub-zip code variation in premium and exposure, they provide broader assessments of company behavior, market concentration, competition, and industry trends.

The first dataset is the Community Service Statement (CSS) data, which reports company by year by zip code information on total premium and number of policies, segmented by policy type. All insurance companies licensed to operate in California's admitted market respond to this data collection survey. We use these data to impute zip code- or company-specific premium and market share. These data are used, for instance, in Table 1 and Figure 1.

The second dataset is the Personal Property Experience (PPE) data, which gives zip code-level information on number of policies, total coverage, and deductibles, separated by policy type. It is not company-specific. Insurers that wrote more than $5 million in premium for either dwelling fire or homeowners insurance report this information. While most CDI data are reported for all years from 2009 to 2020, the coverage and deductible data in PPE were linearly interpolated for the years 2010, 2012, 2014, and 2016.

The third dataset is the Residential Property Experience Data (RPE). Beginning in 2015, insurers with combined total written premiums of $5 million or more for dwelling fire or homeowners lines of business were required to respond to an annual RPE data call. These data report, at the zip code level, the number of new residential policies written, the number renewed, the number of non-

47. National Association of Insurance Commissioners (NAIC). (2021). Report on Profitability by Line and by State. https://naic.soutronglobal.net/Portal/Public/en-US/RecordView/Index/7008.

48. The request numbers were PRA-2022-00204 and PRA-2023-00342.

renewed policies, and the number of cancelled policies. These data are used to show the cancellation panels of Figure 1.

The last zip code-level dataset from CDI is the Wildfire Risk Information Report. Starting in 2018, all admitted insurers with at least $10 million or more in written California premium in dwelling fire or homeowners lines of business have submitted reports to CDI on wildfire risk exposure. The dataset reports, at a zip code level, fire- or wildfire-incurred losses, as well as the distribution of insured parcels across wildfire risk categories. We use these data to coarsely classify zip codes into quantiles of risk for Table 1 and Figure 1. In addition, they are used for the two-stage least squares (2SLS) analysis of Section 5.

A.3 Wildfire risk and home characteristics

We obtained proprietary data from CoreLogic, LLC on parcel-level house characteristics and wildfire risk. The parcel-level sample consists of 100,000 single family homes in California. These data include standard assessor's information such as the home address, geolocated coordinates, reconstruction cost, and the year of construction. They also provide relevant information related to wildfire risk, such as the construction material, the presence of fire resistive siding or roofing, distance to high hazard vegetation, distance to a responding fire station, and a public protection classification that rates community fire protection services.

A key feature of these data is a set of deterministic categorical wildfire risk scores (WRS) which are used by many insurers in the pricing and underwriting process. The main WRS is a rating that ranges from 5 to 100. This measure is based on factors such as slope, aspect, fuel, past burns, and distance to vegetation. Other risk scores are included, such as a brushfire risk rating and a set of crime indices. None of these factors are derived from probabilistic models but are commonly used by insurers in decision-making.

Separately, the data report a set of probabilistic catastrophe loss measures which are derived from simulations. For each property we observe probabilistic measures of the annual average loss (AAL), the standard deviation of losses, and aggregate exceedance probability (AEP) losses over return periods of 50, 100, 250, and 500 years. The AAL is the average yearly loss in dollars, which is roughly the probability of destruction times the reconstruction cost of the home. The AEP describes the probability distribution of the sum of losses over various return periods; for example, for a 250 year return period we might observe that a house has a $1/250 = 0.4\%$ chance of $\$k$ in total losses, where k is reported for each parcel.

To develop the sample of 100,000 homes, we drew on zip code-level data from CoreLogic. These data report the total number of single family homes in a zip code falling into categorical wildfire risk scores of 1 to 50, 51 to 60, 61 to 80, and 81 to 100. We used this dataset to identify 400 California zip codes with meaningful variation in wildfire risk scores; then, we received a sample of 250 houses per zip code with the aforementioned parcel-level characteristics, giving us 100,000 properties for the analysis. The final estimation sample drops 4,648 homes where the reported reconstruction costs exceed $2 million or data are missing for quality of local fire department, distance to high risk vegetation, or distance to fire station. This leaves 95,352 homes in the main estimation sample.

A.4 Insurer rate filings

To determine insurance pricing and eligibility we developed data from public insurance rate filings. Because California is a prior approval state, any company in the admitted market must submit rate increase requests for approval by CDI. As part of the rate request, insurers must provide complete copies of their rate manuals which they use to set premiums and eligibility for customers. All rate filings are publicly available through CDI's website.[49]

We reviewed rate filings for dozens of large insurance companies. Rate requests can range from several hundred pages to more than 10,000 pages. A typical rate request takes between six months and two years from submission to approval and involves several rounds of correspondence and objection letters from state rate specialists. There is no limit on how often an insurer may file a new request, although an insurer cannot submit a new request for an insurance line while another is pending.

Using the property and risk data, we develop the full insurance pricing and eligibility schedule of the 100,000 home sample for six large insurers listed in Appendix Table 1. As a result of this exercise we are able to observe the price a company would offer to any house, even if the house is ineligible for a policy. For each company, we use the most recent rate filing as of 2021; correspondingly, the premiums are priced in 2021 dollars.

Several parcel-level characteristics affect the insurance premium or eligibility, such as the age of the home, zip code, geocoordinates, construction, roof, proprietary wildfire risk score, public protection class of the community, or distance to vegetation. To derive full prices, we must make several assumptions about the policy. First, we assume that a homeowner purchases coverage equal to the reconstruction cost of the home, which is generally advised by insurers. In addition, we assume a $1,000 deductible. For all other coverages, such as liability or loss of use, we assume default coverage for each insurer, which is summarized in Appendix Table 1. We do not assume any additional coverages, such as for scheduled items like furs and jewelry. Lastly, we assume the most standard homeowners policy, i.e. not a deluxe or premium plan.

Appendix Table 1: Standard coverage levels for rate filings

Company filing	Coverage A - Dwelling	Coverage B - Other Structures	Coverage C - Personal Property	Coverage D - Loss of Use	Coverage E - Personal Liability	Coverage F - Medical Payments	Deductible
AAA SoCal 15-6084	Repl. cost	10% of Cov. A	75% of Cov. A	20% of Cov. A	$100,000	$1,000	$1,000
Allstate 21-1436	Repl. cost	10% of Cov. A	50% of Cov. A	20% of Cov. A	$100,000	$1,000	$1,000
Liberty Mutual 19-1562	Repl. cost	10% of Cov. A	50% of Cov. A	20% of Cov. A	$100,000	$1,000	$1,000
Nationwide 20-612	Repl. cost	10% of Cov. A	55% of Cov. A	No limit (24 months)	$100,000	$1,000	$1,000
State Farm 21-1404	Repl. cost	10% of Cov. A	75% of Cov. A	30% of Cov. A	$100,000	$1,000	$1,000
USAA 21-809	Repl. cost	10% of Cov. A	50% of Cov. A	20% of Cov. A	$300,000	$5,000	$1,000

Notes: Coverage A is assumed as the structure replacement cost, which is recommended by insurers. Deductible is assumed as $1,000. All other values are standard for each rate filing. Rate filing identifiers correspond to California Department of Insurance filing numbers, accessible through Web Access to Rate and Form Filings (WARFF).

We must also make assumptions about the homeowner characteristics. We assume that the customer has not had a recent claim, that they have been with the insurer for fewer than two years, and that they bundle their homeowners and automobile insurance policies.[50] Protective devices also

49. California Department of Insurance (CDI). (2024). Web Access to Rate and Form Filings (WARFF). http://www.insurance.ca.gov/0250-insurers/0800-rate-filings/0050-viewing-room.

50. As of 2015, 78 percent of consumers bundled their homeowners and auto policies. See: J.D. Power. 2015 US Household Insurance Study. https://www.jdpower.com/business/press-releases/2015-us-household-insurance-study.

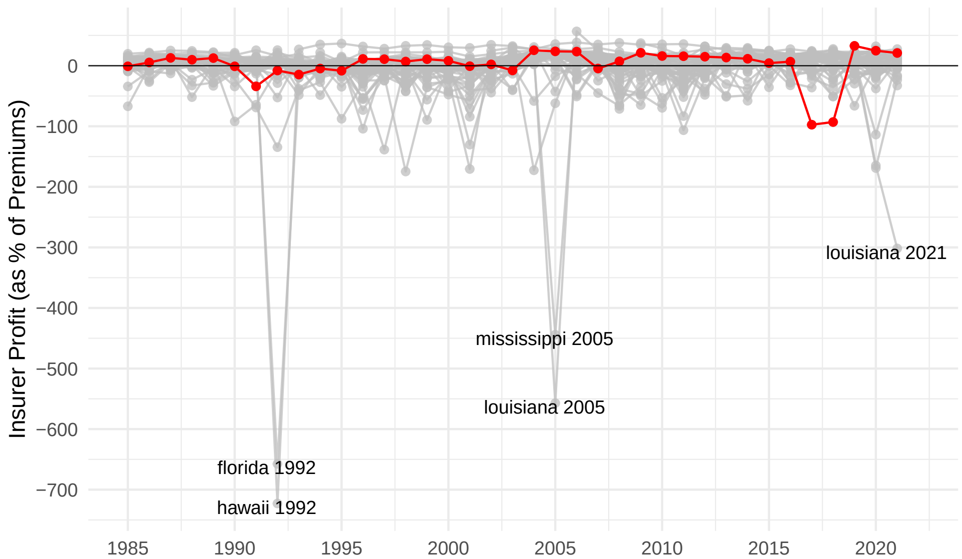
typically factor in to homeowners insurance; we assume the customer has smoke detectors, dead bolt locks, and fire extinguishers, but no burglar alarm (local or central-reporting), no central-reporting fire alarms, and no sprinklers. The homeowner is assumed to be 45 years old and married without children. Importantly, these characteristics are assumed constant when pricing every company's premiums.

Besides constructing premiums, we also use the rate filings to assess the granularity of each firm's wildfire risk measurement. The count of each firm's rating variables is reported in Table 1. The following variables were counted as wildfire variables: (i) Numeric wildfire hazard scores from firms such as CoreLogic or Verisk; (ii) Any custom territory, such as State Farm's grid ID, AAA Southern California's brush fire territories, or a territory that aggregates several zip codes; (iii) An administrative territory such as a zip code if there is wildfire-specific information; if there is only a general zip code factor which is not explicitly related to wildfire, it is not counted; (iv) Public protection class or any variable interacted with it, such as construction type $\times$ protection class. Table 1 is constructed using the following rate filings, which reflect the most recent filings as of 2021: State Farm 21-1404, Farmers 21-2410, CSAA/AAA NorCal 20-4189, Mercury 20-3267, Auto Club Enterprises/AAA SoCal 15-6084, Liberty Mutual 19-1562, Allstate 21-1436, USAA 21-809, Nationwide 20-612, Travelers 20-887, FAIR Plan 21-2452.

B Additional Results

B.1 Additional figures and tables

Appendix Figure 1: Profitability of homeowners insurers by state and year



Notes: Figure reports state-level summaries of insurer profitability in homeowners insurance. California is in red; all other states are in gray. The data represent profit on insurance transactions, which accounts for underwriting profit, investment gains, and federal taxes. Data are from National Association of Insurance Commissioners "Report on Profitability by Line by State" for the years 1985 to 2021. See Appendix A for a data description.

Appendix Table 2: Hypothetical cost-based pricing with asymmetric risk classification

Wildfire Score Range	Mean Wildfire AAL	Expected Profit from Pricing at Mean	Average Breakeven Price	Average Market Share	Homes
(0,20]	56	-14	166.88	0.495	12,208
(20,40]	162	-45	230.42	0.518	23,778
(40,60]	300	-100	825.12	0.467	19,883
(60,70]	290	-86	1,016.13	0.481	10,981
(70,80]	377	-121	2,863.79	0.315	7,793
(80,90]	439	-158	5,017.60	0.181	11,853
(90,100]	871	-575	9,382.48	0.001	5,491

Notes: Table reports additional outcomes for the example of cost-based pricing under asymmetric risk classification described in Section 4.4. Column (3) is expected profit per customer served for the firm segmenting on wildfire risk scores ("Firm A"). Column (4) summarizes the minimum price at which Firm A earns zero profits in each segment. These prices are calculated by successively increasing Firm A prices in each segment. Column (5) is the share of customers in each segment served by Firm A at the prices in Column (4). This exercise omits homes in grid cells where the dataset contains only one home. A well-known theoretical result that we reproduce in Section 7 is that when Firm B prices each house at expected cost and there are no switching costs, Firm A cannot break even except by serving the single highest-risk customer in each segment. Firm A does better in this example for two reasons: Firm B prices at the grid (not house) level and there are switching costs. Nevertheless, in some of the highest-risk segments, this simple example collapses to the degenerate outcome where Firm A charges a price equal to the maximum expected wildfire loss in the segment and serves one or zero customers.

B.2 Multi-way Bertrand simulations

This section extends the hypothetical Bertrand duopolies in Section 6.1.2 to allow for simultaneous competition between all six firms in the dataset. As in Table 4, we assume that customers choose the firm offering the lowest annual premium, but we now assume customers choose from all firms in the data. Table 3 reports results from estimating Equation 11 separately for each firm. The regression for each firm includes fixed effects for that firm's wildfire risk classification variables. As in the duopoly results in the main text, firms pricing on less granular information suffer a large winner's curse. Column (1) shows that there is no statistically significant difference in the homes that State Farm wins relative to homes in the same 1 kilometer grid cell that it loses to another firm. Columns (2) through (5) show that for Allstate, USAA, Liberty Mutual, and AAA Southern California, winning a customer reveals significant "bad news" relative to what that firm knew about the customer previously (that is, relative to other customers in the same class within the firm's risk classification method). The empirical magnitudes of the winner's curse for each firm are similar to those in the main text. In Column (6), there is a noisy zero result for the winner's curse for Nationwide. This may reflect the fact that Nationwide's high prices in high-risk segments lead to it winning very few customers in those segments, such that there are not enough homes won by Nationwide to yield an empirically meaningful estimate in this regression framework.[51] Those implied market shares are reported in Appendix Table 4. The first column shows the share of customers in the dataset won by each firm, and the second shows the share of customers with wildfire risk scores above 70 won by each firm.

51. Nationwide's low market shares also reflect the similar segmentation methods used by Allstate and Nationwide. In unreported regressions, we repeat this multi-way exercise omitting Allstate from the simulation and find a large winner's curse effect for Nationwide, in line with the other firms in the table.

Appendix Table 3: The winner's curse in stylized multi-way Bertrand

	State Farm	Allstate	USAA	Liberty Mutual	AAA SoCal	Nationwide
	(1)	(2)	(3)	(4)	(5)	(6)
win = TRUE	-0.027	0.563***	0.399***	0.404***	0.663***	-0.075
	(0.024)	(0.093)	(0.045)	(0.038)	(0.120)	(0.085)
State Farm segment FE	✓					
Allstate segment FE		✓				
USAA segment FE			✓			
Liberty Mutual segment FE				✓		
AAA SoCal segment FE					✓	
Nationwide segment FE						✓
Observations	95,295	95,295	95,295	93,976	55,606	95,295

Notes: Table reports results of six separate regressions following Equation 11 for the multiway competition described in Appendix Section B.2. AAA Southern California is included as an option for homes in its service territory. Standard errors are clustered by zip code.

Appendix Table 4: Market shares in stylized multi-way Bertrand

Firm	Overall	High Wildfire Scores
State Farm	0.14	0.20
Allstate	0.16	0.06
Nationwide	0.25	0.03
Liberty Mutual	0.15	0.27
USAA	0.21	0.27
AAA SoCal	0.10	0.17

Notes: Table reports share of customers won by each firm in the multiway competition described in Appendix B.2. "High wildfire scores" are wildfire score groups of 70 or higher. AAA Southern California is included as an option for homes in its service territory.

C Risk Load

This section describes how we approximate the risk load associated with covering the high wildfire hazard homes in the dataset, as presented in Equation 7. Our approach uses catastrophe model estimates of the variance of property-level losses. We use this information to calibrate an analytical "marginal surplus" risk load calculation (Kreps 1990). We first describe how we create a statewide pseudosample that allows us to explore the relationship between risk load and market share. We then present details on the marginal surplus calculation.

C.1 Constructing a statewide pseudosample

We calculate the risk load associated with successively larger portfolios of homeowners policies in high wildfire hazard areas of California. The following process generates a pseudosample that approximately replicates the full population of wildfire-threatened homes in California.

We start from our sample of 100,000 homes in 400 high-hazard zip codes. We augment these data with summary counts of total homes and high-hazard (score > 50) homes for all California zip codes from CoreLogic. In the course of this merge, we calculate that the 400 zip codes in our detailed sample account for about 75 percent of the total homes in California with wildfire scores above 50. We then define 800 zip code $\times$ hazard (score above/below 50) bins. In each of these bins, we draw homes from that bin with replacement until we reach the total number of homes reported for that bin in the zip code totals data. This "builds up" a sample that approximately matches the wildfire risk distribution for the true population of homes in these zip codes.[52] We can then study the relationship between risk load and market share by calculating risk loads associated with covering various fractions of these homes (10 percent, 20 percent, etc.). This exercise assumes that insurers equalize their market shares across zip codes, which may be a reasonable approximation given the diversification benefits of such a strategy.

C.2 Marginal surplus calibration

Our wildfire catastrophe model data report the mean (AAL), standard deviation, and 99th, 99.6th, and 99.8th percentiles of annual wildfire losses for each property in the dataset. We do not observe information about the covariance of losses across properties. We assume a conservatively high degree of correlation given our interest in benchmarking approximate upper bounds on risk loads. Let each home i's exposure to wildfire in a given year be determined by a Bernoulli random variable W_i with success probability w_i. Conditional on exposure to a wildfire, destruction of home i is a Bernoulli random variable D_i with probability d_i so that home i's annual probability of destruction is $w_i d_i$. Let $Cor(W_i, W_j)$ be the correlation in annual wildfire exposure between homes i and j. Assume d is independent across homes. One can show that the correlation in annual losses across homes i and j is

$$Cor(W_i D_i, W_j D_j) = \frac{d_i d_j \sqrt{w_i(1-w_i)}\sqrt{w_j(1-w_j)}}{\upsilon_i \upsilon_j},$$

where

$$\upsilon_i = \sqrt{w_i d_i(1-d_i) + d_i^2 w_i(1-w_i) + w_i(1-w_i)(d_i)(1-d_i)}.$$

52. The need to stratify by high/low instead of simply resampling at the zip level arises because the process originally used to select these 100,000 homes was not a simple random sample.

See Section C.3 for the proof. We take d_i to be 0.5 for all i based on destruction probabilities conditional on exposure reported in Baylis and Boomhower (2022). We assume conservatively high annual exposure probabilities of 1 percent for all homes.[53] We define a 20-by-20 km grid over the state of California and assume that the correlation in annual wildfire occurrence for two homes in the same grid cell is 0.5, and that this correlation is zero across grid cells. These assumptions imply that $Cor(W_iD_i, W_jD_j)$, the correlation in realized losses, is 0.25 for homes in the same grid cell and 0 for homes in different grid cells.

Kreps (1990) derives the minimum necessary premium to add a portfolio of risks to an existing book of insurance contracts as:

$$\rho_i(\Omega_j) = l_i + a_i + \phi'_{ij}(\Omega_j),$$

where l_i is the average annual loss of the contracts being added, $\phi'_{ij}(\Omega_j)$ is the change in risk load from the contracts being added, and a_i is a constant that captures administrative costs. Let S_j be the standard deviation of losses from the existing book of business, S'_{ij} the standard deviation of losses from the combined book of business, C_{ij} the correlation of losses on the new contracts with the existing book of business, y a market cost of capital, and z a distribution statistic that reflects the firm's "acceptable probability of ruin." For example, if annual losses are distributed normally, setting $z = 2.65$ implies the firm will cover losses in 99.6 percent of years, enough for a 1-in-250 year event. Kreps (1990) shows that:

$$\phi'_{ij}(\Omega_j) = \underbrace{\frac{y}{1+y}}_{\text{capital cost}} \times \underbrace{z}_{\text{“distribution statistic”}} \times \underbrace{\frac{(2S_jC_{ij} + \sigma_i)\sigma_i}{S_j + S'_{ij}}}_{\text{change in s.d. of firm's losses}} .$$

The final term is the change in the standard deviation of the firm's losses after adding the new contracts to the portfolio. It can be derived by starting from the identity $Var(X+Y) = Var(X) + Var(Y) + 2Cov(X,Y)$:

$$(S'_{ij})^2 = S_j^2 + \sigma_i^2 + 2C_{ij}S_j\sigma_i$$
$$(S'_{ij})^2 - S_j^2 = \sigma_i(\sigma_i + 2C_{ij}S_j)$$
$$(S'_{ij} + S_j)(S'_{ij} - S_j) = \sigma_i(\sigma_i + 2C_{ij}S_j)$$
$$(S'_{ij} - S_j) = \frac{\sigma_i(\sigma_i + 2C_{ij}S_j)}{(S'_{ij} + S_j)}.$$

Given an assumed correlation structure across homes, the data allow us to calculate σ_i, the standard deviation of wildfire losses for our portfolio of high wildfire hazard homes, by summing standard deviations within and then across grid cells, again using the variance sum rule. We calibrate S_j to match the distribution of losses for an insurer with a countrywide portfolio of homeowners and automobile insurance policies using aggregate loss statistics from the Insurance Information Institute (III).[54] We assume that this insurer faces another catastrophe peril (e.g., hurricane, fire

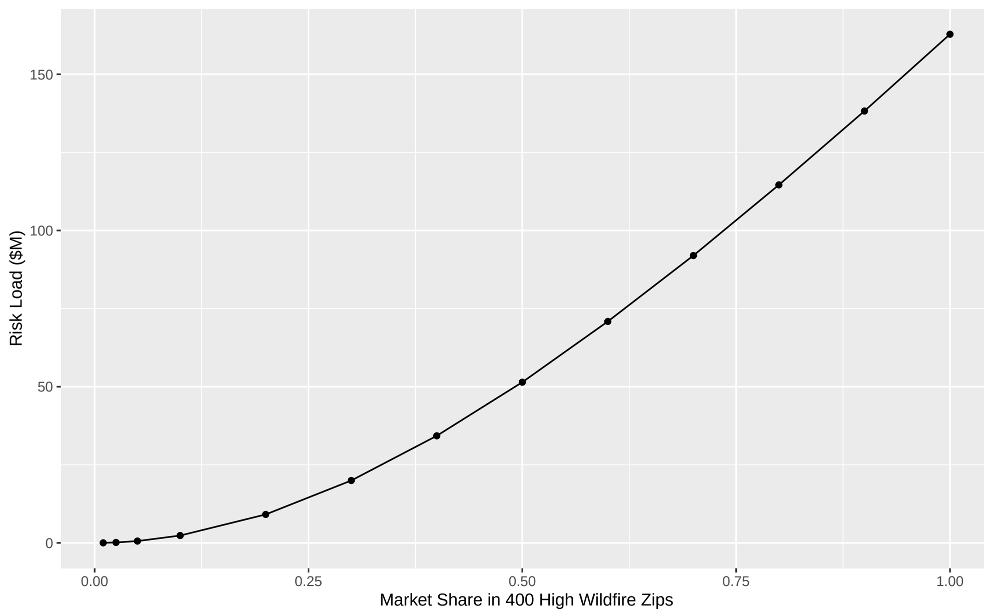
53. Based on data in Buechi et al. (2021), we estimate that the average annual share of acres in California that experience a wildfire was 0.3 percent to 0.7 percent in each decade between 1979 and 2018.

54. This hypothetical insurer has 4.5 million U.S. homeowners policies and 4.5 million U.S. auto insurance policies. Based on III data on claims frequency and severity, homeowners insurance claims are an iid binomial process with

following earthquake) with expected annual losses of $200 million and coefficient of variation of 2.5.[55] We see these assumed hurricane/earthquake exposures as conservative for several reasons (1) we only assume one other peril type; (2) the hurricane AAL values we can observe in 10K filings are net of reinsurance (and so do not represent the full AAL); (3) we choose a coefficient of variation on the small side of reported values. We assume that wildfire losses are independent of all other types of losses. We set z equal to 2.65, so that the firm is required to be able to meet its obligations up to a 1-in-250-year loss total. We assume the market return y on capital is 0.15.

Appendix Figure 2 shows the relationship between risk load $\phi(\Omega_j)$ and market share. There is a convex relationship, reflecting the shrinking diversification benefit as the portfolio becomes relatively more concentrated in wildfire risk.

Appendix Figure 2: Risk load versus market share in wildfire areas



Notes: Figure shows total risk load expenses due to the addition of a portfolio of California wildfire risks to a national property insurance portfolio. Horizontal axis shows fraction of total homes that the insurer covers in 400 zip codes that represent most high-hazard California homes. Vertical axis is risk load in millions of dollars. See text for details and assumptions.

Appendix Table 5 reports details of the marginal surplus calculations. A 1 percent market share in the 400 high-hazard zip codes would represent about 28,000 insurance policies and $6 million in expected losses. The standard deviation of wildfire losses (σ_i) would be $8.9 million. The standard deviation of losses on the rest of the portfolio (S_j) is $550.1 million, and the addition of these 28,000 wildfire risks to the portfolio has only a small effect on the overall standard deviation (S'_{ij}). The

annual claim probability of 5.92 percent and loss amount per claim of $15,091, while auto insurance claims are an iid binomial process with annual claim probability of 1.1 percent and loss amount per claim of $18,204. See: https://www.iii.org/table-archive/21296 and https://www.iii.org/fact-statistic/facts-statistics-auto-insurance.

55. These numbers are calibrated to financial statements from real firms: Zurich Insurance Group (Farmers) reports an AAL of $192 million for North America hurricane losses (see Zurich Insurance Group, Annual Report 2022, page 136). The coefficient of variation is informed by publicly available catastrophe model predictions for hurricane wind losses from insurer filings with the Florida Commission on Hurricane Loss Projection Methodology. See page 205 of the RMS filing "North Atlantic Hurricane Models: Version 23.0 (Build 2250), May 19 2023" and page 191 of the CoreLogic filing "Florida Hurricane Model 2023, April 24, 2023 Version."

resulting risk load is less than $100,000 dollars or less than 1 cent per dollar of wildfire AAL. The size of this risk load increases quickly with the number of wildfire policies. At a 20 percent market share, the risk load is 8 cents per dollar of AAL, or $16 per policy. The rate of increase in risk load with market share eventually slows as the variance of the portfolio becomes dominated by wildfire losses.

A key insight is that the marginal change in risk load associated with adding a high-hazard parcel to a firm's book of business depends not only on the riskiness of the property, but also its covariance with other properties in the firm's risk portfolio. Firms with a larger share of business outside of California will have lower surplus requirements because the correlation between a high-hazard property in California and insured risks outside of California is lower.

Relatedly, the marginal surplus also depends on exposure to other catastrophe perils, as illustrated by panel B of Appendix Table 5. This exercise illustrates how diversification in insurers' overall risk profile can generate meaningful variation in marginal surplus costs. To generate these results, we focus on hurricanes and assume a high-hazard market share of 10 percent. As with the baseline scenario of panel A, we assume a coefficient of variation of hurricane losses of 2.5, but vary the insurer's expected hurricane losses. If an insurer has no exposure to hurricane risk (such that the variability of non-wildfire losses is driven exclusively by non-catastrophe homeowners and auto losses), the variability of wildfire losses dominates the overall portfolio and the risk load can reach 48 cents per dollar of AAL. Adding hurricane exposure significantly decreases this risk load. Given the larger magnitude of hurricane risk compared to wildfire risk across the United States, an insurer that is adequately capitalized against hurricane losses is well-positioned to manage wildfire risk. More generally, these calculations illustrate how variation in insurers' overall risk profile can generate meaningful variation in marginal surplus costs.

C.3 Derivation of correlation in losses

Let W_i and W_j be Bernoulli random variables with success probabilities w_i and w_j and correlation $Cor(W_i, W_j)$. Let D_i and D_j be independent Bernoulli random variables with success probabilities d_i and d_j. The correlation of realized losses W_iD_i and W_jD_j is, by definition,

$$Cor(W_iD_i, W_jD_j) = \frac{Cov(W_iD_i, W_jD_j)}{\sqrt{Var(W_iD_i)}\sqrt{Var(W_jD_j)}}.$$

Bohrnstedt and Goldberger (1969) shows that the covariance term in the numerator has an asymptotic approximation as

$$\begin{aligned} Cov^*(W_iD_i, W_jD_j) = & \, E[W_i]E[W_j]Cov(D_i, D_j) + E[W_i]E[D_j]Cov(W_j, D_i) \\ & + E[W_i]E[D_j]Cov(W_j, D_i) + E[D_i]E[D_j]Cov(W_i, W_j). \end{aligned}$$

The Ds are independent, so this simplifies to:

$$\begin{aligned} &= E[D_i]E[D_j]Cov(W_i, W_j) \\ &= d_i d_j Cor(W_i, W_j) sd(W_i) sd(W_j) \\ &= d_i d_j Cor(W_i, W_j)\sqrt{w_i(1-w_i)}\sqrt{w_j(1-w_j)}. \end{aligned}$$

Goodman (1960) derives an exact formula for the variances in the denominator:

$$\begin{aligned} Var(W_iD_i) &= E[W_i]^2Var(D_i) + E[D_i]^2Var(W_i) + Var(W_i)Var(D_i) \\ &= w_i^2 d_i(1-d_i) + d_i^2 w_i(1-w_i) + w_i(1-w_i)d_i(1-d_i). \end{aligned}$$

Combining yields

$$Cor(W_iD_i, W_jD_j) = \frac{d_i d_j \sqrt{w_i(1-w_i)}\sqrt{w_j(1-w_j)}}{\upsilon_i \upsilon_j},$$

where

$$\upsilon_i = \sqrt{w_i d_i(1-d_i) + d_i^2 w_i(1-w_i) + w_i(1-w_i)(d_i)(1-d_i)}.$$

Appendix Table 5: Approximate risk load from marginal surplus

Panel A: Risk load versus market share in California wildfire zip codes

Market Share (%)	Policies	Wildfire AAL ($M)	σ_i ($M)	S_j ($M)	S'_{ij} ($M)	Risk Load ($M)	Average Risk Load per AAL ($)	Marginal Risk Load per AAL ($)
1	27,825	6.0	8.9	550.07	550.14	0.02	0.004	NA
5	139,722	30.1	44.0	550.07	551.82	0.59	0.020	0.029
10	279,612	60.6	88.2	550.07	557.08	2.36	0.039	0.058
20	559,379	120.4	174.8	550.07	577.18	9.11	0.076	0.113
30	839,123	181.1	262.5	550.07	609.50	19.97	0.110	0.179
50	1,398,690	301.7	438.0	550.07	703.17	51.44	0.170	0.283
100	2,797,560	603.8	876.3	550.07	1,034.66	162.81	0.270	0.405

Panel B: Sensitivity of risk load to other catastrophe perils

Hurricane AAL ($M)	Policies	Wildfire AAL ($M)	σ_i ($M)	S_j ($M)	S'_{ij} ($M)	Risk Load ($M)	Average Risk Load per AAL ($)	Marginal Risk Load per AAL ($)
0	279,612	60	87.59	8.54	88.01	26.70	0.44	0.48
50	279,612	60	87.66	137.76	163.29	8.57	0.14	0.21
100	279,612	60	87.55	275.13	288.73	4.57	0.08	0.11
200	279,612	61	88.15	550.07	557.08	2.36	0.04	0.06
300	279,612	60	87.82	825.04	829.71	1.57	0.03	0.04
500	279,612	60	87.76	1,375.03	1,377.82	0.94	0.02	0.02

Notes: The table reports the marginal surplus and associated risk load required to cover various shares of homes in wildfire areas of California for a hypothetical countrywide property insurer. See Section 4.2 and Appendix C for details and assumptions.

D Instrumenting for Measurement Error in Segment-Mean Wildfire Risk

This section elaborates on the two-stage least squares (2SLS) specification for β_{jk} as described in Section 5. We first show that sampling variation in structure-level wildfire risk generates biased β_{jk} estimates when we observe few homes per territory. We then show that instrumenting for territory mean risk with an auxiliary risk measure alleviates this issue.

D.1 Measurement error

To estimate β_{jk} for a given firm, we calculate average wildfire risk and insurance premium in each of the firm's wildfire pricing segments. These estimates are potentially susceptible to bias from measurement error in the reconstruction of segment-mean wildfire risk. Measurement error in segment-mean price is less of an issue both because a firm's prices do not vary within its pricing segments (see Figure 4) and because the classical errors-in-variables model implies measurement error in the dependent variable is less important. The mean wildfire risk for a territory k in the sample can be represented as $\tilde{l}_k = l_k + \nu_k$, where ν_k is a mean-zero sampling error whose variance is inversely proportional to sample size. An OLS regression of p_{jk} on $\tilde{l}_k$ will yield a biased estimate of β_{jk}, with the size of the bias depending on ν_k and thus the number of segments that the firm uses for pricing. For example, Liberty Mutual uses 393 wildfire segments, meaning that we observe 100,000/393 = 254 homes per segment on average; while Nationwide uses 30 wildfire segments, meaning that we observe 3,333 homes per segment on average. Appendix Table 6 shows that OLS estimates of β_{jk} for firms using more than about one hundred pricing segments are notably smaller than the other estimates, with values substantially below one. This suggests measurement error in segment-mean risk due to sampling variation is causing attenuation in the OLS estimates for firms with many segments.

D.2 Instrumenting for measurement error

A common approach to address measurement error is to instrument for the mismeasured regressor with auxiliary data on the mismeasured quantity. In this setting, we observe a zip code level measure from the California Department of Insurance (CDI) that summarizes the distribution of wildfire risk for all homes in the zip code.[56] Appendix Table 7 reports first-stage regressions of segment-mean wildfire risk on this auxiliary measure. The dependent variable in these regressions is the regression-adjusted segment mean wildfire AAL that comes from Equation 9. The independent variable is the weighted mean of CDI zip code risk for the zip codes containing the homes that we observe in the segment. CDI zip-code risk summaries strongly predict the mean risk in the data, with first stage F statistics exceeding 60 for all but the insurer with the least-sophisticated wildfire risk pricing.

Appendix Table 6 shows OLS and 2SLS estimates of β_{jk} for each firm. The OLS and 2SLS estimates are similar for firms with few territories (Allstate, Nationwide) and different for firms with many territories (USAA, Liberty Mutual, State Farm). This is true for the relationship between the firm's own prices and territory risk (panel A) and State Farm's prices and each firm's territory-level mean risk (panel B). The exception is AAA Southern California. For this firm, the IV and OLS specifications differ meaningfully even though the firm only has 19 territories. Both slope estimates

56. These reports aggregate survey information from all major insurers, were required by SB 824. See Appendix A for discussion.

Appendix Table 6: OLS versus IV price gradients

Firm Segmentation	Allstate		Nationwide		USAA		Liberty Mutual		AAA SoCal		State Farm	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Fringe Firm Prices												
Segment-Mean AAL	2.67***	2.87***	3.57***	3.42***	0.67***	2.17***	0.65**	2.46***	-0.32***	-1.44**	0.36***	1.10
(SE)	(0.36)	(0.15)	(0.61)	(0.60)	(0.13)	(0.31)	(0.16)	(0.37)	(0.44)	(1.08)	(0.03)	(0.14)
State Farm Prices												
Segment-Mean AAL	0.83**	0.95	0.96	0.95	0.48***	1.01	0.45***	0.91	1.09	1.99		
(SE)	(0.08)	(0.14)	(0.16)	(0.12)	(0.08)	(0.20)	(0.09)	(0.16)	(0.34)	(0.98)		
F-statistic		255.6		61.1		62.1		88.8		9.1		1694.2
Zip Code FE	Yes	Yes	Yes	Yes	No	No	No	No	No	No	No	No
Number of Segments	94	94	30	30	323	323	393	393	19	19	11946	11946

Notes: Table reports estimates of β_{jk} from ordinary least squares and a two-step estimation procedure that first calculates regression-adjusted mean prices and wildfire AAL by segment following Equations 9 and 10, and then regresses these segment-level prices against segment-level mean wildfire AAL. The first-step estimation of segment means includes zip code fixed effects for Allstate and Nationwide, where there is cross-cutting variation in zip codes and wildfire segments. The second-step regression of mean price on mean wildfire AAL is estimated by two-stage least squares when there are many territories to remove measurement error in segment-mean wildfire AAL due to sampling variation. Standard errors are calculated by bootstrapping the full estimation procedure 500 times.

Appendix Table 7: IV First Stage Results

Segmentation for Instrument	Allstate Scores	Nationwide Scores	USAA Territories	Liberty Mutual Territories	AAA SoCal Territories	State Farm Grid Cells
Segment-Mean Risk Coefficient	599.68***	434.95***	276.05***	293.22***	221.55***	327.99***
(SE)	(37.51)	(55.63)	(35.02)	(31.11)	(73.51)	(7.97)
Zip Code FE	Yes	Yes	No	No	No	No
Number of Segments	94	30	323	393	19	11946

Notes: This table reports estimates of a regression of regression-adjusted segment-level prices against segment-level mean wildfire risk. This is the first stage of the two-stage least squares approach explained in the text. Standard errors are calculated by bootstrapping the estimation procedure 500 times. "Zip Code FE" indicates whether zip code fixed effects are included in the initial estimation of regression-adjusted segment-level prices. This is only possible for firms that segment on wildfire risk scores, since there is cross-cutting variation in wildfire risk score and zip code.

for AAA Southern California also have a negative sign. We note that the first stage relationship for AAA is weak, with a first-stage F statistic of about 9. The AAA Southern California results may also be related to the general difference in observed wildfire pricing behavior for AAA Southern California throughout the paper.

E Alternative Explanations for Empirical Risk Price Gradients

The empirical results of Sections 5 and 6 are consistent with a story in which insurers using less granular risk classification strategies set wildfire risk premiums above segment-average risk measures to account for adverse selection. This evidence is highly suggestive, but not dispositive. There are other possible explanations for the relatively steep risk pricing gradients we observe among firms operating at an information disadvantage. Premium markups in high-risk market segments could additionally represent a response to uncertain loss probabilities, differences in risk load across insurers, or differences in risk assessment modeling.

E.1 Alternative models of wildfire risk

The CoreLogic CAT modeling estimates are one of several available sources of wildfire risk exposure information. Alternative sources include other proprietary CAT models, such as those generated by Verisk or RMS, and publicly provided estimates from sources like the U.S. Forest Service. Differences in CAT model projections could potentially explain some disparities in risk pricing that we observe across firms.

Notably, State Farm uses both the CoreLogic and Verisk AIR models of wildfire risk;[57] The correspondence we document between segment-average CoreLogic AAL estimates and risk pricing in Section 5.1 suggests that the Verisk AIR model generates similar relative differences in AAL across market segments. Still, the possibility remains that disparities in risk pricing could be due to CAT model differences.

We cannot access proprietary AAL metrics generated by other leading providers. We can, however, use some publicly available metrics to assess the robustness of our findings to alternative measures of wildfire risk exposure. The U.S. Forest Service (USFS) has developed a model of Risk to Potential Structures (RPS), a gridded data product which estimates the annual probability of destruction from wildfire, for a representative house, at any point in the United States (Scott et al. 2024). For each home in the dataset, we calculate a rough USFS-based "pseudo-AAL" by multiplying the USFS RPS annual loss probability by the structure reconstruction cost from CoreLogic.

These USFS-based pseudo-AALs require a number of caveats that limit the degree to which they can be interpreted literally as dollar-denominated loss predictions. Unlike the CoreLogic wildfire catastrophe model or another insurance pricing model, the USFS RPS model is not designed for insurance use or for single-location analysis. The USFS model assumes that all homes are equally vulnerable to destruction during a wildfire, regardless of structure characteristics such as roof, year built, etc. Furthermore, the underlying wildfire model is not designed to represent fire spread in developed urban areas, which includes many of the neighborhoods in our study, relying instead on ad hoc post-processing and interpolation ("oozing") for fire spread in these areas. Finally, the USFS RPS model includes only a limited, not-locally-explicit treatment of fire suppression (firefighting). For all of these reasons, the USFS-based pseudo-AALs provide a potentially useful alternative characterization of *relative* risk across properties, but are unlikely to generate reliable predictions of the expected *level* of losses.[58]

57. See, for example, State Farm rate filing 21-1404 to CDI, April 2021.

58. The USFS notes that the dataset "provides information about communities' relative wildfire risk profile" and that "As a national-scale project with a focus on communities, Wildfire Risk to Communities has limitations. The data are not locally calibrated and are not fine scale—they are best for considering risk as aggregated across a community and are not designed for considering risk at the individual home scale." See https://wildfirerisk.org/about/faq/ and Scott et al. (2024) for details.

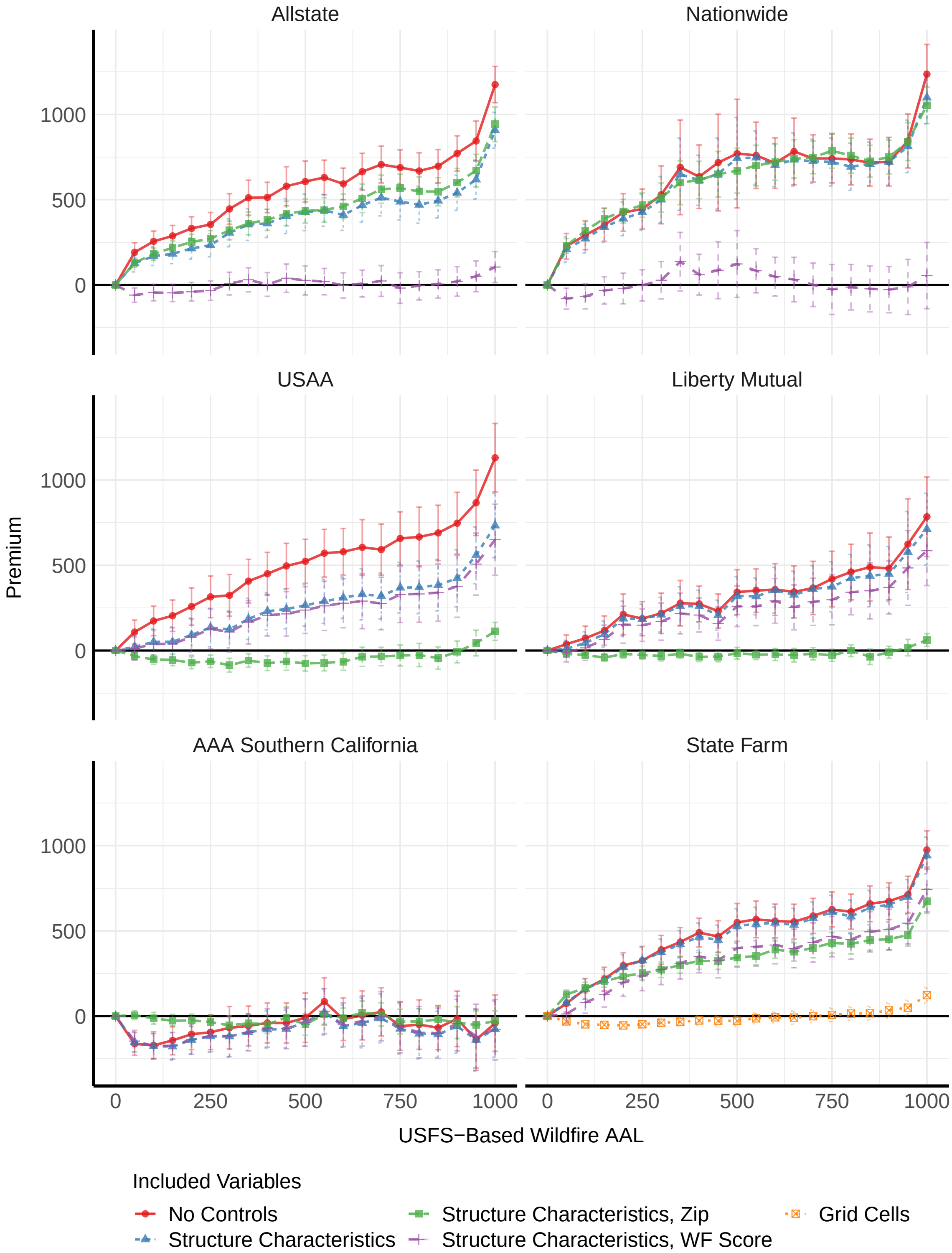
We reproduce the pricing and adverse selection results of the main text using the USFS pseudo-AALs in place of the CoreLogic AALs. Appendix Figure 3 reproduces the variance decomposition of Figure 4 using the USFS-based pseduo-AAL. As in the main text, State Farm's prices display a clear information advantage. There is a strong residual relationship between State Farm prices and USFS-based AAL even after controlling for zip code or wildfire risk score segment dummies.

Appendix Figure 4 reproduces the illustrative cost-based benchmark under symmetric and asymmetric risk classification from Figure 3. As in the main text, a firm that classifies customers by wildfire risk scores and has a competitor who classifies customers by 1 kilometer grid cells serves a highly adversely selected customer population. In the highest wildfire risk segments, the USFS pseudo-AALs for the customers served by the firm with less sophisticated pricing are many times higher than the overall mean in the segment.

Appendix Table 8 reproduces the risk price gradients reported in Figure 5 using the USFS-based pseudo-AALs as the measure of predicted wildfire risk. Because of the limitations on the USFS pseudo-AAL for predicting the level of expected loss, we focus on the relative difference in implied pricing behavior for Allstate and Nationwide relative to State Farm. As in the main text, the estimated price gradients for Allstate and Nationwide are substantially steeper than for State Farm. The bottom row shows that regardless of which risk measure we use, the Allstate price gradient is about 3.3 times steeper than the State Farm price gradient, and the Nationwide price gradient is about 3.8 times steeper than the State Farm price gradient.

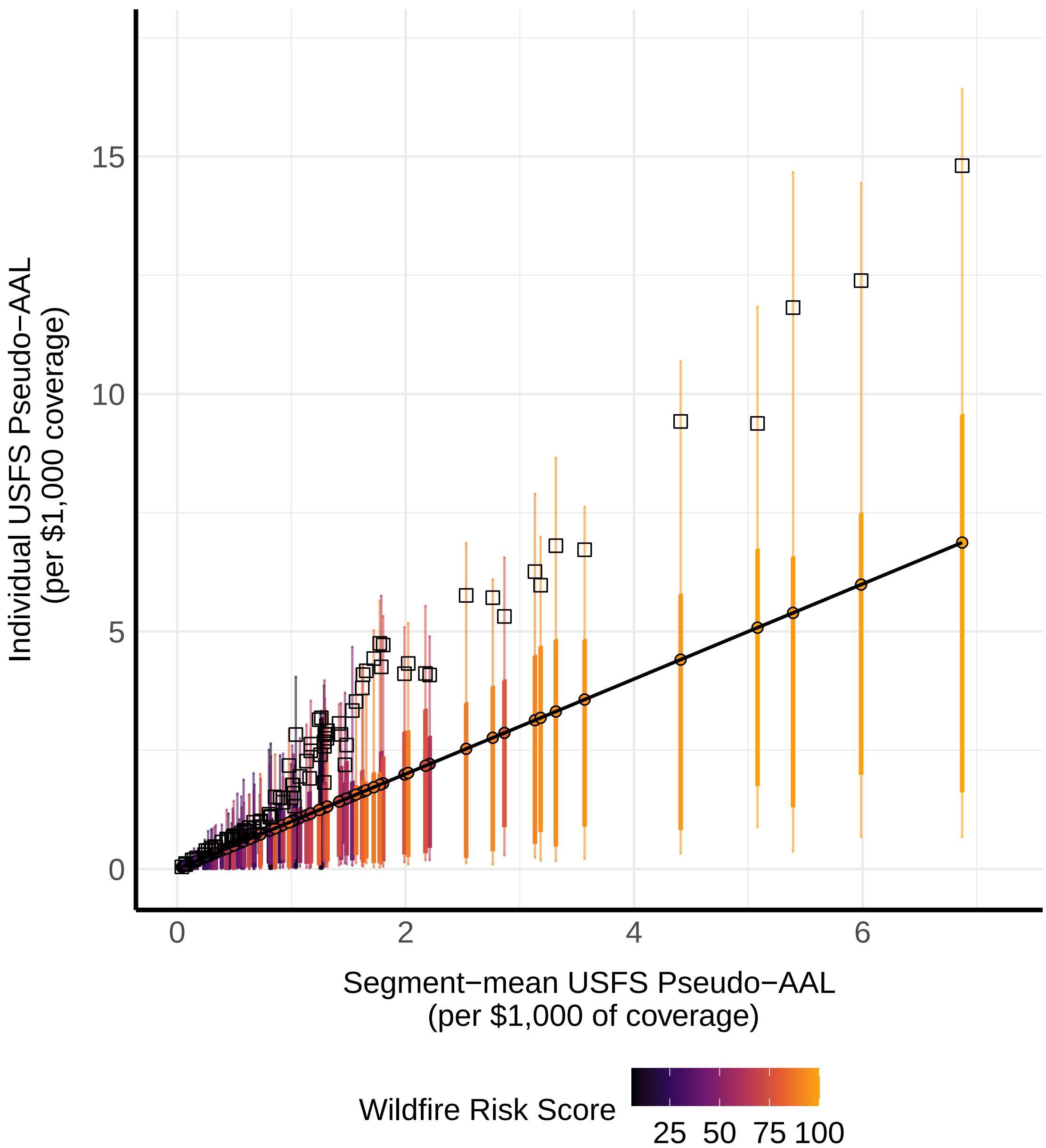
Appendix Table 9 compares the the winner's curse in Bertrand duopoly when using the USFS-based (panel A) vs. CoreLogic (panel B) wildfire AALs. We estimate the regression with AAL in levels, instead of logs as in the main text, because there are some zero values for the USFS-based AAL. Compared to the CoreLogic AALs, the magnitude of the winner's curse is substantially larger using the USFS-based AALs. One reason for this difference is that the wildfire risk is higher on average for all homes using the USFS-based AAL, as evidenced by the dependent variable mean of \$509 in panel A vs. \$302 in panel B. Moreover, the degree of adverse selection is also higher in panel A. Using the example of Allstate, the USFS-based AALs imply that customers won by Allstate are 329/509 = 65% costlier than those won by State Farm. Using the CoreLogic AALs implies a difference of 116/302 = 38%.

Appendix Figure 3: Validating firms' reported pricing variables, using USFS AAL



Notes: Constructed similarly to Figure 4, using USFS-based wildfire AAL instead of CoreLogic wildfire AAL.

Appendix Figure 4: Illustrative per-customer costs under symmetric and asymmetric risk classification, using USFS pseudo-AAL



Notes: Box plots show the distribution of individual USFS pseudo-AAL within each wildfire risk score. Whiskers show the 10th and 90th percentiles; notches show the interquartile range; circular markers are means. Horizontal axis plots the mean of USFS pseudo-AAL in each wildfire risk score. The square markers show Firm A's expected per-customer pseudo-AAL in the illustrative asymmetric classification example in Section 4.4.

Appendix Table 8: Implied Wildfire price gradients using USFS pseudo-AALs

	Panel A: Allstate		Panel B: Nationwide	
	CoreLogic	USFS Pseudo-AAL	CoreLogic	USFS Pseudo-AAL
(A) Price Gradient, Challenger	2.67	0.97	3.57	1.54
(B) Price Gradient, State Farm	0.83	0.28	0.96	0.38
Ratio (B)/(A)	3.20	3.42	3.71	4.08

Notes: Reproduces the exercise in Figure 5 using USFS pseudo-AAL. Regressions include zip code fixed effects.

Appendix Table 9: The winner's curse in stylized Bertrand duopoly, alternative wildfire risk models

Panel A: USFS-based wildfire AAL

	USFS pseudo-AAL				
	Allstate	Nationwide	USAA	Liberty Mutual	AAA SoCal
	(1)	(2)	(3)	(4)	(5)
1[Win]	333.58***	339.65***	373.73***	384.80***	359.94***
	(44.75)	(52.85)	(45.52)	(57.30)	(58.58)
Structure Characteristics	✓	✓	✓	✓	✓
Allstate Segments FE	✓				
Nationwide Segments FE		✓			
USAA Segments FE			✓		
Liberty Segments FE				✓	
AAA SoCal Segments FE					✓
Observations	95,295	95,295	95,295	93,976	55,606
R^2	0.23	0.17	0.34	0.37	0.13
Dependent variable mean	519.15	519.15	519.15	523.14	509.89
Fraction Won	0.53	0.55	0.50	0.44	0.50

Panel B: CoreLogic wildfire AAL

	Wildfire AAL				
	Allstate	Nationwide	USAA	Liberty Mutual	AAA SoCal
	(1)	(2)	(3)	(4)	(5)
1[Win]	115.71***	75.51***	138.75***	82.37***	96.73***
	(25.51)	(27.43)	(17.12)	(14.42)	(19.73)
Structure Characteristics	✓	✓	✓	✓	✓
Allstate Segments FE	✓				
Nationwide Segments FE		✓			
USAA Segments FE			✓		
Liberty Segments FE				✓	
AAA SoCal Segments FE					✓
Observations	95,295	95,295	95,295	93,976	55,606
R^2	0.19	0.16	0.44	0.48	0.20
Dependent variable mean	302.21	302.21	302.21	299.09	235.61
Fraction Won	0.53	0.55	0.50	0.44	0.50

Notes: Follows Table 4, using alternative models for the dependent variable. Regressions are estimated with the AAL in levels, rather than logs as in the main text, because there are some zero values for the USFS-based AAL.

E.2 Ambiguity loading

Premium markups in high-risk market segments could additionally represent a response to uncertain loss probabilities. A series of studies has shown that insurers demand premiums in excess of expected costs when loss probabilities are ambiguous (Dietz and Niehörster 2021; Kunreuther et al. 2009; Kunreuther et al. 1995). More recent work has estimated ambiguity loads – i.e. the extra insurance premium due to ambiguity – and shown how these can vary with the insurer's degree of ambiguity aversion (see, for example, Dietz and Niehörster 2021).

Climate change increases the ambiguity in climate risk projections (Moore 2024). If insurers are adjusting their premiums to account for this ambiguity, this could help explain the steep wildfire risk pricing gradients we estimate. We cannot fully rule out this explanation. However, to fully rationalize the significant differences in risk pricing gradients we observe across firms, we would need to assume that firms using more granular risk information are significantly less ambiguity-averse. We also note that ambiguity loading and the winner's curse are not necessarily distinct mechanisms. Mumpower (1991) and Kunreuther et al. (1995) discuss how ambiguity loading may be a response to the winner's curse, as firms anticipate that uncertain loss probabilities will increase the dispersion of predicted losses and thus offered prices.

E.3 Risk load differences

Insurers must purchase reinsurance or hold large capital reserves to be ready to pay out significant claims when a natural disaster strikes. Firms with higher costs of capital, more limited access to reinsurance, or a less-diversified catastrophe risk portfolio will face higher capital reserve costs. For example, Appendix Table 5 shows how risk loads vary with levels of hurricane exposure. Thus, one possible explanation for the differences in the risk gradients we observe could be underlying differences in risk load costs.

In 2022, State Farm, Allstate, Liberty Mutual, Nationwide, and USAA were all among the top ten writers of property insurance in the United States.[59] These insurers look quite similar along dimensions we can readily observe. Consider, for example, a comparison between the two largest writers of homeowners multi-peril insurance: State Farm and Allstate. These two firms account for approximately 9 percent and 5 percent of the U.S. property and casualty insurance market, respectively. In 2022, they had similar loss ratios (60 percent for State Farm and 63 percent for Allstate). They also have similar shares of premiums written in states with high hurricane risk. For example, in 2022, the high-risk state of Florida was home to 4 percent of State Farm's written premiums and 3.9 percent of Allstate's. Although these comparisons are far from comprehensive, similarities along the dimensions we can observe do not provide evidence that differences in rate schedules are explained by differences in risk load costs.

59. Source: NAIC data, sourced from S&P Global Market Intelligence, Insurance Information Institute.

F Proofs and Results for Equilibrium Model

F.1 Proof of Proposition I

The best response of the dominant firm to a price p^F set by the fringe firms is the set of prices that maximize profits. For a given risk level l, the dominant firm can take all customers at a maximum price of $p^F - \delta$ or sell only to its current customers (a share α of the market) at a maximum price of $p^F + \delta$. Any price below $p^F - \delta$ reduces profits and any price above $p^F + \delta$ loses all customers. The low and high prices yield identical profits at $\tilde{l}$, defined by:

$$\pi^D = p^F - \delta - \tilde{l} = \alpha(p^F + \delta - \tilde{l}). \tag{21}$$

Equation 21 is rearranged as $\tilde{l} = p^F - \delta\frac{1+\alpha}{1-\alpha}$. For $0 < \alpha < 1$, $|\frac{d\pi^D}{dl}|$ is greater for the low price than the high price strategy and, therefore, the low price (high price) strategy yield higher profits for $l < \tilde{l}$ ($l > \tilde{l}$). We next show that $\tilde{l} > 0$, which implies that $p^D = p^F - \delta$ is the best response for $l \in [0, \tilde{l})$. To establish this, we will (a) solve for p^F, the price that makes fringe profits zero, and (b) show that the condition $\bar{l} > \delta\frac{1+\alpha}{1-\alpha}$ implies $\tilde{l} > 0$. The price p^F satisfies:

$$\pi^F = \frac{1-\alpha}{2}(p^F - \tilde{l})^2 - \frac{1}{2}(l^* - p^F)^2 + \frac{\alpha}{2}\delta^2 = 0. \tag{22}$$

Solving for p^F yields $p^F = l^* - \delta\sqrt{\frac{1+3\alpha}{1-\alpha}}$, which is result (1) of the Proposition. If $p^F > \bar{l}$, then:

$$l^* - \delta\sqrt{\frac{1+3\alpha}{1-\alpha}} > \bar{l}; \tag{23}$$

$$\bar{l} > \delta\sqrt{\frac{1+3\alpha}{1-\alpha}}, \tag{24}$$

where $l^* - \bar{l} = \bar{l}$. The second inequality is established by showing that $\frac{1+\alpha}{1-\alpha} > \sqrt{\frac{1+3\alpha}{1-\alpha}}$ and combining this with the condition $\bar{l} > \delta\frac{1+\alpha}{1-\alpha}$. Having established that $p^F > \bar{l}$, we have from the definition of $\tilde{l}$:

$$p^F = \tilde{l} + \delta\frac{1+\alpha}{1-\alpha} > \bar{l}; \tag{25}$$

$$\tilde{l} > \bar{l} - \delta\frac{1+\alpha}{1-\alpha} > 0, \tag{26}$$

where the last inequality uses $\bar{l} > \delta\frac{1+\alpha}{1-\alpha}$. We have shown the first part of result (2) of the proposition: The price $p^D = p^F - \delta$ is the dominant firm's best response for $l \in [0, \tilde{l})$. The second part of result (2) is established from Equation 21, where it was shown that $p^D = p^F + \delta$ yields the highest profits for $l > \tilde{l}$. The price $p^D = p^F + \delta$ maximizes profits for values of $l > \tilde{l}$, yielding positive profits. Dominant firm profits, $\pi^D = p^F + \delta - l$, are decreasing in l and equal to zero at $l_1 = p^F + \delta$. Therefore, $p^D = p^F + \delta$ is a dominant firm's best response for $l \in [\tilde{l}, l_1)$, the second part of result (2). Finally, for values of $l > l_1$, dominant firm profits are negative and it maximizes profits by not selling any insurance policies. It achieves this by setting a sufficiently high price, $p^D > p^F + \delta$, so that all customers are served by the competitive fringe. Thus, we have established the third part of result (2): $p^D > p^F + \delta$ is the dominant firm's best response for $l \in [l_1, l^*]$.

F.2 Proof of Proposition II

Under the market equilibrium in Proposition I, the dominant firms earns positive profits (Equation 14) and the fringe earns zero profits (Equation 15). Total profits for the dominant and fringe firms with technology adoption (T) can be written:

$$\begin{aligned}\pi^T = \int_0^{\tilde{l}} (p^F - \delta - l)\frac{1}{l^*}dl + \alpha \int_{\tilde{l}}^{l_1} (p^F + \delta - l)\frac{1}{l^*}dl \\ + (1-\alpha)\int_{\tilde{l}}^{l_1} (p^F - l)\frac{1}{l^*}dl + \int_{l_1}^{l^*} (p^F - l)\frac{1}{l^*}dl.\end{aligned} \tag{27}$$

Similarly, profits under the initial conditions (0) can be written:

$$\begin{aligned}\pi^0 = \int_0^{\tilde{l}} (p_0 - l)\frac{1}{l^*}dl + \alpha \int_{\tilde{l}}^{l_1} (p_0 - l)\frac{1}{l^*}dl \\ + (1-\alpha)\int_{\tilde{l}}^{l_1} (p_0 - l)\frac{1}{l^*}dl + \int_{l_1}^{l^*} (p_0 - l)\frac{1}{l^*}dl.\end{aligned} \tag{28}$$

Solve the integrals in Equations 27 and 28 and evaluate the inequality $\pi^T > \pi^0$. After cancelling common terms, obtain:

$$\bar{p} = (p^F - \delta)\frac{\tilde{l}}{l^*} + \alpha(p^F + \delta)\frac{l_1 - \tilde{l}}{l^*} + (1-\alpha)p^F\frac{l_1 - \tilde{l}}{l^*} + p^F\frac{l^* - l_1}{l^*} > p_0. \tag{29}$$

F.3 The effect of market share on the value of information

We consider how the *VOI* for the dominant firm varies with its market share. By differentiating Equation 14 with respect to α, rearranging, and making use of the definitions $p^F = l^* - \delta\sqrt{\frac{1+3\alpha}{1-\alpha}}$ and $\tilde{l} = p^F - \delta\frac{1+\alpha}{1-\alpha}$, we find that *VOI* is increasing in market share ($\frac{\partial VOI}{\partial \alpha} > 0$) if and only if:

$$\tau < \frac{1}{2} + \left(\frac{1+3\alpha}{1-\alpha}\right)^{0.5}, \tag{30}$$

where $\tau = \frac{\bar{l}}{\delta}$. The variable τ measures the relative magnitude of switching costs, specifically the size of δ compared to the average AAL ($\bar{l}$) in the segment. For $\alpha < 1$, the right-hand side of Equation 30 is increasing in α and has an asymptote at infinity as α goes to one. Thus, for a given value of τ, there is always a market share α at which *VOI* is increasing in the market share.

F.4 Results for the equilibrium model with imperfect risk information

The model of Section 7.3 assumes that the information technology gives the dominant firm perfect information about the risk of each property. In practice, even the best catastrophe models provide risk estimates measured with error and, in some cases, somewhat coarse information (e.g., risk bins). In this section, we consider how the market equilibrium changes when the dominant firm adopts an imperfect information technology.

We assume that within the risk segment, the information technology assigns consumers to subsegments and that this assignment process has some error. Specifically, there are M risk groups, each containing an equal number of properties, and the technology provides a noisy signal for property i denoted $m_i \in M$. We assume that if t_i is the true risk assignment for the property, then:

$$\begin{aligned} p(m_i = t_i) &= b \\ p(m_i = m) &= \frac{1-b}{M-1}, \quad m \neq t_i. \end{aligned} \tag{31}$$

In Equation 31, each property is assigned to the correct group with probability b. For an incorrectly assigned property, there is an equal probability that the true risk group is any one of the other groups. We assume the assignment process must be better than random, $b > \frac{1-b}{M-1}$, to justify costly investment in the information technology.

We derive the market equilibrium for the case in which the risk model groups customers within the risk segment into $M = 2$ groups, a low-risk group with $l \in [0, \bar{l})$ and a high-risk group with $l \in [\bar{l}, l^*)$. We continue to assume a uniform distribution for l and allow the model's signal $m = \{low, high\}$ to be imperfect. From Equation 31, $P(m_i = t_i) = b$ and $P(m_i \neq t_i) = 1 - b$, where t_i is the true risk assignment. For the model to be better than random at assigning properties, $b > 0.5$.

Bayes' rule gives the conditional probability that the true risk is some value x given the signal m_i:

$$\begin{aligned} P(l_i = x|m_i) &= \frac{b}{\bar{l}}, \quad x \in t_i; \\ P(l_i = x|m_i) &= \frac{1-b}{\bar{l}}, \quad x \notin t_i. \end{aligned} \tag{32}$$

Using Equation 32, we can write the expected risk of a property i conditional on m_i as:

$$\begin{aligned} E[l_i|m_i = low] &= \bar{l}(1.5 - b); \\ E[l_i|m_i = high] &= \bar{l}(0.5 + b). \end{aligned} \tag{33}$$

Notice that as the signal becomes less precise, the two expected values converge. At $b = 0.5$, the expected risk equals the mean risk $\bar{l}$ for both signals.

The dominant firm can only distinguish properties as low or high risk according to the signal m_i and, thus, must charge the same price to customers within these groups. If p^F is the price charged by the fringe firms, then a low/high pricing strategy by the dominant firm is as follows:

$$\begin{aligned} p^D &= p^F - \delta, \quad m_i = low; \\ p^D &= p^F + \delta, \quad m_i = high. \end{aligned} \tag{34}$$

Following the logic of the basic model, the dominant firm will always want to charge as high a price as possible when it undercuts the fringe firms and takes the whole market at $p^F - \delta$ and when it sells to only its loyal customers at $p^F + \delta$.

Under the low/high pricing strategy, the dominant captures all of the customers with a low-risk signal. This means that fringe firms will sell only to $(1-\alpha)$ of the customers with a high-risk signal.

A fringe firm's expected profits are, thus,

$$\pi^F = (1-\alpha)(p^F\bar{l} - \bar{l}^2(\frac{1}{2}+b)). \tag{35}$$

In Equation 35, we assume that the number of properties in the risk segment is sufficiently large such that the risk model generates each signal for $\bar{l}$ customers. Setting fringe profits equal to zero, we obtain the fringe price as $p^F = \bar{l}(\frac{1}{2}+b)$. The dominant firm's expected profits are then given by:

$$\begin{aligned}\pi^D &= (p^F-\delta)\bar{l} - (\frac{3}{2}-b)\bar{l}^2 + \alpha(p^F+\delta)\bar{l} - \alpha(\frac{1}{2}+b)\bar{l}^2 \\ &= \bar{l}^2(2b-1) - \delta\bar{l}(1-\alpha).\end{aligned} \tag{36}$$

There are other pricing strategies that the dominant firm could use, characterized by combinations of low $(p^F-\delta)$ and high $(p^F+\delta)$ prices set for customers with low- and/or high-risk signals. We focus on the alternative pricing strategy in which the dominant firm sets a high price $(p^F+\delta)$ for customers with low- and high-risk signals. It can be shown that if the low/high strategy is more profitable in expectation than this "high-price" alternative, then it dominates all other potential pricing strategies. The expected profits for the high price strategy are simply:

$$\pi^D = 2\delta\bar{l}\alpha. \tag{37}$$

The dominant firm's profits are independent of b because, for any value of b, equal numbers of customers are mis-categorized as low- and high-risk. When the dominant firm uses the high-price strategy, the fringe price is $p^F = \bar{l}$.

Appendix Table 10: Critical values of $\tau = \bar{l}/\delta$ above which a low/high pricing strategy is more profitable than the high-price strategy

	α				
b	0.1	0.2	0.3	0.4	0.5
1.00	1.10	1.20	1.30	1.40	1.50
0.95	1.22	1.33	1.44	1.56	1.67
0.90	1.38	1.50	1.63	1.75	1.88
0.85	1.57	1.71	1.86	2.00	2.14
0.80	1.83	2.00	2.17	2.33	2.50
0.75	2.20	2.40	2.60	2.80	3.00
0.70	2.75	3.00	3.25	3.50	3.75
0.65	3.67	4.00	4.33	4.67	5.00
0.60	5.50	6.00	6.50	7.00	7.50
0.55	11.00	12.00	13.00	14.00	15.00
Basic model	1.22	1.50	1.86	2.33	3.00

Comparing profits in Equations 36 and 37, we find that the low/high pricing strategy is more profitable when:

$$\tau = \frac{\bar{l}}{\delta} > \frac{1+\alpha}{2b-1}. \tag{38}$$

Appendix Table 10 shows the critical values of τ for a range of α and b values. For comparison, the bottom row of the table provides the critical τ value for the low/high pricing strategy to be optimal in the perfect information model. We see that as long as δ is not too large relative to the mean risk in the segment, which would imply a small value of τ, the low/high pricing strategy is still optimal even when the risk information is imperfect.

F.5 Results for the equilibrium model with regulation

We show that the zero profit and average price isoclines in Equations 19 and 20, respectively, are upward sloping at the unconstrained equilibrium, as shown in panel (a) of Figure 9. From Equation 15, equilibrium profits for the fringe firms are given by:

$$\pi^F = \frac{1}{2}\Big\{(1-\alpha)(p^F - \tilde{l})^2 - (l^* - p^F)^2 + \alpha\delta^2\Big\} = 0. \tag{39}$$

Applying the Implicit Function Theorem, we obtain:

$$\frac{dp^F}{d\tilde{l}} = -\frac{\pi^F_{\tilde{l}}}{\pi^F_{p^F}} = \frac{(1-\alpha)(p^F - \tilde{l})}{(1-\alpha)(p^F - \tilde{l}) + l^* - p^F} > 0, \tag{40}$$

where $p^F - \tilde{l} > 0$ and $l^* - p^F > 0$ from the results $\tilde{l} = p^F - \delta\frac{1+\alpha}{1-\alpha}$ and $p^F = l^* - \delta\sqrt{\frac{1+3\alpha}{1-\alpha}}$. From Equation 17, we have:

$$M = \bar{p} - \eta(p^F - \delta) - (1-\eta)(p^F + \delta) = 0, \tag{41}$$

where $\eta = \frac{\tilde{l}}{p^F - \delta}$ at the unconstrained equilibrium. Applying the Implicit Function Theorem, we obtain:

$$\frac{dp^F}{d\tilde{l}} = -\frac{M_{\tilde{l}}}{M_{p^F}} = \frac{2\delta(p^F - \delta)}{2\delta\tilde{l} + (p^F - \delta)^2} > 0. \tag{42}$$

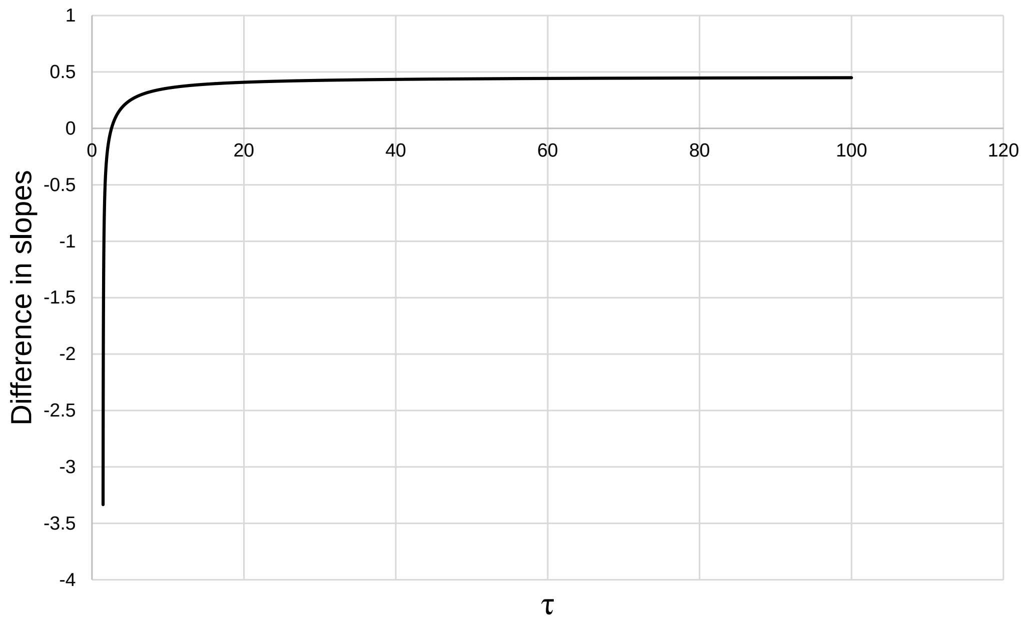
Numerical analysis is used to show that at the unconstrained equilibrium the relative magnitudes of the isocline slopes in Equations 40 and 43 are indeterminate. It can be shown that scaling δ and l^* by a constant factor leaves the slopes in Equations 40 and 43 unchanged. Therefore, we need only consider how the slopes vary with $\tau = \frac{\bar{l}}{\delta}$. Appendix Figure 5 shows that for $\alpha = 0.2$, the difference in slopes can be positive or negative. Except for small values of τ, the slope of the zero profit isocline is greater than that of the average price isocline.

An increase in $\bar{p}$ shifts up the average price isocline in Equation 17, as shown in panel (b) of Figure 9. Fixing p^F and applying the Implicit Function Theorem to Equation 41, we obtain:

$$\frac{d\tilde{l}}{d\bar{p}} = -\frac{M_{\bar{p}}}{M_{\tilde{l}}} = -\frac{[\tilde{l} + \alpha(p^F + \delta - \tilde{l})]^2}{2\delta\alpha(p^F + \delta)} < 0. \tag{43}$$

The decline in $\tilde{l}$ for fixed p^F gives the upward shift in the average price isocline depicted in Figure 9.

Appendix Figure 5: Difference in isocline slopes for different values of τ



Notes: Figure shows the slope of the zero profit isocline in Equation 40 minus the slope of the average price isocline in Equation 43 for $\alpha = 0.2$ and different values of $\tau = \frac{\bar{l}}{\delta}$. Results show that the zero profit isocline is steeper than the average price isocline at the unconstrained equilibrium for values $\tau \geq 2.6$.

F.6 Results for the equilibrium model with information provision

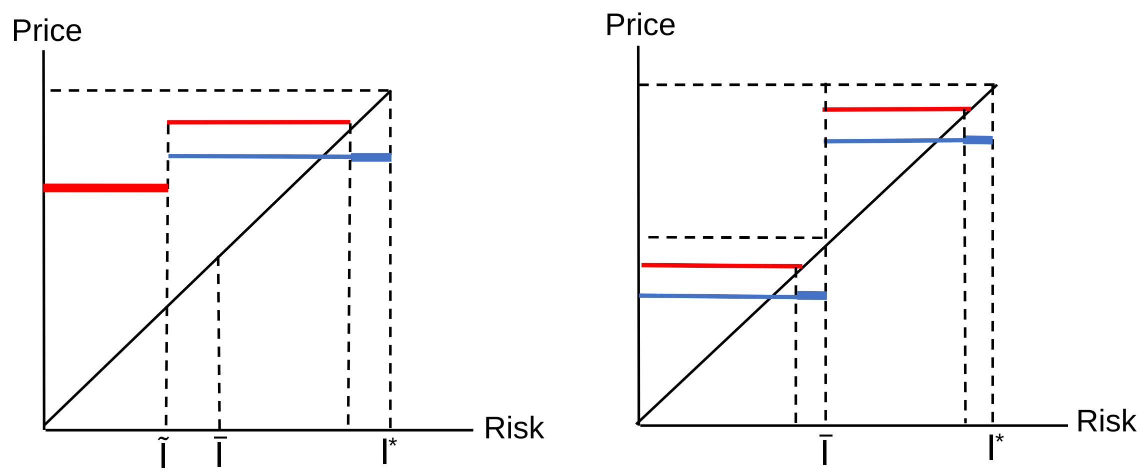
Panels (c) and (d) of Figure 8 depict the case in which $\tilde{l} > \bar{l}$ under the original equilibrium. However, if $\tilde{l} < \bar{l}$ initially, then with information provision there will not exist values of $\tilde{l}$ below which the dominant firm wants to use the low price strategy. To see this, consider that under the original equilibrium, $\tilde{l} = l^{max} - \delta\sqrt{\frac{1+3\alpha}{1-\alpha}} - \delta\frac{1+\alpha}{1-\alpha}$. This value is unaffected by the division of the risk segment, implying that it will fall outside of the interval $[\bar{l}, l^*]$ defining the new, higher-risk segment. Since the new, lower-risk segment is symmetric, $\tilde{l}$ will also fall outside of $[0, \bar{l})$. Rather than a two-part pricing strategy, the equilibrium in each segment will be defined by a p^F that makes the fringe profits zero, as before, and a price of $p^F + \delta$ set by the dominant firm that retains its current customers; see Appendix Figure 6.

We use numerical analysis to explore how the overall average price changes in this case. We set $\alpha = 0.2$ and $\bar{l} = 50$ and calculate the overall average price with and without information provision for δ values ranging from 18 to 34. The lower value of δ yields an $\tilde{l}$ under the original equilibrium just below $\bar{l}$ and the upper value yields an $\tilde{l}$ just above zero. Appendix Figure 7 shows that information provision reduces average prices for all values of δ.

Appendix Figure 6: Effects of information provision on market equilibrium: $\tilde{l} < \bar{l}$ under the original equilibrium

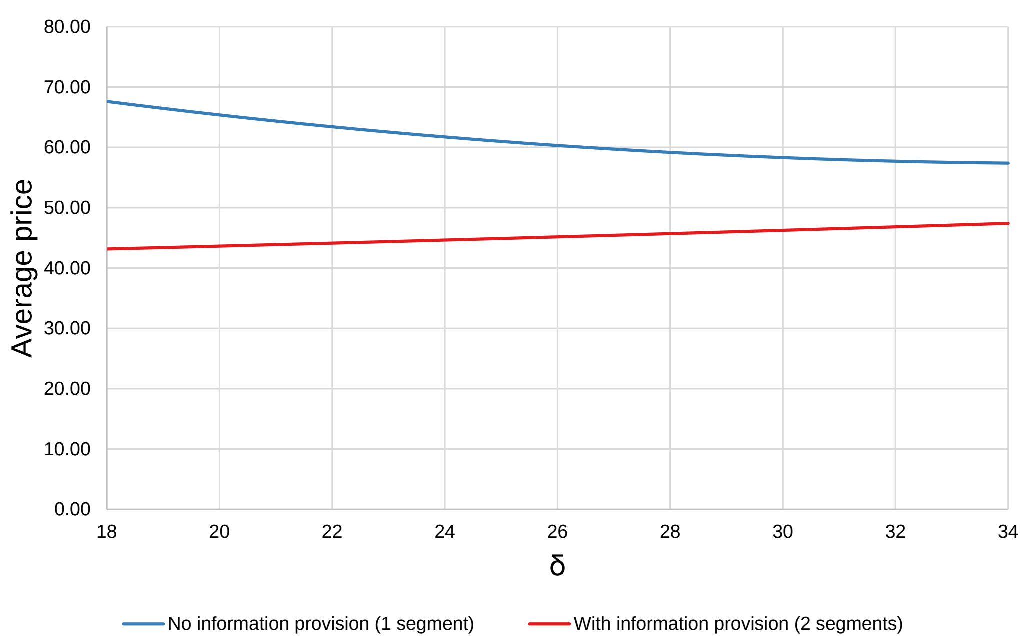
(a) Baseline case

(b) Effect of information provision



Notes: The two panels show how equilibrium prices change when new information is provided to the competitive fringe firms. Panel (a) is the original case in which fringe firms know only that properties are distributed according to $U(0, l^*)$. In panel (b), the firms can distinguish whether properties are distributed according to $U(0, \bar{l})$ or $U(\bar{l}, l^*)$.

Appendix Figure 7: Difference in overall average prices with information provision



Notes: Figure shows differences in overall average prices with information provision (see Appendix Figure 6). Parameters values are $\alpha = 0.2$ and $\bar{l} = 50$. The δ parameter is varied between 18, which yields an $\tilde{l}$ value just below $\bar{l}$ under the original equilibrium, and 34, which yields an $\tilde{l}$ value just above zero. As shown, information provision reduces the overall average price for all δ values.

Appendix References

Abraham, Kenneth S. 1985. "Efficiency and Fairness in Insurance Risk Classification." *Virginia Law Review,* 403–451.

Aizawa, Naoki, and Ami Ko. 2023. *Dynamic Pricing Regulation and Welfare in Insurance Markets.* National Bureau of Economic Research.

Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *The Quarterly Journal of Economics* 84 (3): 488–500.

Baylis, Patrick W, and Judson Boomhower. 2022. *Mandated vs. Voluntary Adaptation to Natural Disasters: The Case of U.S. Wildfires.* National Bureau of Economic Research.

Bohrnstedt, George W., and Arthur S. Goldberger. 1969. "On the Exact Covariance of Products of Random Variables." *Journal of the American Statistical Association* 64 (328): 1439–1442.

Borenstein, Severin. 1989. "The Economics of Costly Risk Sorting in Competitive Insurance Markets." *International Review of Law and Economics* 9 (1): 25–39.

Born, P., and W.K. Viscusi. 2006. "The Catastrophic Effects of Natural Disasters on Insurance Markets." *Journal of Risk Uncertainty* 33:55—72.

Born, Patricia H., and Barbara Klimaszewski-Blettner. 2013. "Should I Stay or Should I Go? The Impact of Natural Disasters and Regulation on U.S. Property Insurers' Supply Decisions." *The Journal of Risk and Insurance* 80 (1): 1–36.

Buechi, Hanna, Paige Weber, Sarah Heard, Dick Cameron, and Andrew J Plantinga. 2021. "Long-term trends in wildfire damages in California." *International Journal of Wildland Fire* 30 (10): 757–762.

Cather, David A. 2018. "Cream Skimming: Innovations in Insurance Risk Classification and Adverse Selection." *Risk Management and Insurance Review* 21 (2): 335–366.

Chamberlin, John R. 1985. "Assessing the Fairness of Insurance Classifications." *Research in Law and Economics* 7:65–87.

Cignarale, T., Joel Laucher, Ken Allen, and L. Landsman-Smith. 2019. *The Availability and Affordability of Coverage for Wildfire Loss in Residential Property Insurance in the Wildland-Urban Interface and Other High-Risk Areas of California: CDI Summary and Proposed Solutions.* California Department of Insurance.

Crocker, Keith J., and Arthur Snow. 1986. "The Efficiency Effects of Categorical Discrimination in the Insurance Industry." *Journal of Political Economy* 94 (2): 321–344.

Dietz, Simon, and Falk Niehörster. 2021. "Pricing ambiguity in catastrophe risk insurance." *The Geneva Risk and Insurance Review* 46 (2): 112–132.

Dixon, Lloyd, Flavia Tsang, and Gary Fitts. 2018. *The Impact of Changing Wildfire Risk on California's Residential Insurance Market: A Report for California's Fourth Climate Change Assessment.* California Natural Resources Agency.

Einav, Liran, and Amy Finkelstein. 2011. "Selection in Insurance Markets: Theory and Empirics in Pictures." *Journal of Economic Perspectives* 25 (1).

Einav, Liran, Amy Finkelstein, and Neale Mahoney. 2021. "The IO of selection markets." In *Handbook of Industrial Organization,* 5:389–426. 1. Elsevier.

Einav, Liran, Mark Jenkins, and Jonathan Levin. 2012. "Contract Pricing in Consumer Credit Markets." *Econometrica* 80 (4): 1387–1432.

Eling, M., I. Gemmo, and D. Guxha. 2024. "Big Data, Risk Classification, and Privacy in Insurance Markets." *Geneva Risk Insurance Review* 49:75–126.

Engelbrecht-Wiggans, Richard, Paul R Milgrom, and Robert J Weber. 1983. "Competitive bidding and proprietary information." *Journal of Mathematical Economics* 11 (2): 161–169.

First Street Foundation. 2022. *5th National Risk Assessment Fueling the Flames.* First Street Foundation.

Ge, Shan. 2022. "How Do Financial Constraints Affect Product Pricing? Evidence from Weather and Life Insurance Premiums." *Journal of Finance* 77 (1): 449–503.

Goodman, Leo A. 1960. "On the Exact Variance of Products." *Journal of the American Statistical Association* 55 (292): 708–713.

Gourevitch, J.D., C. Kousky, and Y. Liao. 2023. "Unpriced climate risk and the potential consequences of overvaluation in U.S. housing markets." *Nature Climate Change.*

Handel, Ben, Igal Hendel, and Michael D Whinston. 2015. "Equilibria in Health Exchanges: Adverse Selection Versus Reclassification Risk." *Econometrica* 83 (4): 1261–1313.

Heiss, Florian, Daniel McFadden, Joachim Winter, Amelie Wuppermann, and Bo Zhou. 2021. "Inattention and switching costs as sources of inertia in Medicare part D." *American Economic Review* 111 (9): 2737–2781.

Hendricks, Kenneth, and Robert H Porter. 1988. "An empirical study of an auction with asymmetric information." *American Economic Review,* 865–883.

Honka, Elisabeth. 2014. "Quantifying search and switching costs in the U.S. auto insurance industry." *The RAND Journal of Economics* 45 (4): 847–884.

Hoy, Michael. 1982. "Categorizing Risks in the Insurance Industry." *The Quarterly Journal of Economics* 97 (2): 321–336.

Iglesias, Virginia, Jennifer K. Balch, and William R. Travis. 2022. "U.S. Fires Became Larger, More Frequent, and More Widespread in the 2000s." *Science Advances* 8 (11): eabc0020.

Jaffee, Dwight M, and Thomas Russell. 1997. "Catastrophe Insurance, Capital Markets, and Uninsurable Risks." *Journal of Risk and Insurance,* 205–230.

Jin, Yizhou, and Shoshana Vasserman. 2021. *Buying Data from Consumers: The Impact of Monitoring Programs in U.S. Auto Insurance.* National Bureau of Economic Research.

Keys, Benjamin J, and Philip Mulder. 2024. *Property Insurance and Disaster Risk: New Evidence from Mortgage Escrow Data.* National Bureau of Economic Research.

Klein, Robert W., and Paul R. Kleindorfer. 1999. *The Supply of Catastrophe Insurance Under Regulatory Constraints.* Center for Financial Institutions Working Paper 99-25. Wharton School Center for Financial Institutions, University of Pennsylvania.

Koijen, Ralph S. J., and Motohiro Yogo. 2015. "The Cost of Financial Frictions for Life Insurers." *American Economic Review* 105 (1): 445–75.

Kousky, Carolyn, and Roger M Cooke. 2009. *The Unholy Trinity: Fat Tails, Tail Dependence, and Micro-correlations.* Resources for the Future 09-36.

Kreps, Rodney. 1990. "Reinsurer Risk Loads From Marginal Surplus Requirements." In *Proceedings of the Casualty Actuarial Society,* 77:196–203. Citeseer.

Kunreuther, Meszaros, Hogarth, and Spranca. 1995. "Ambiguity and underwriter decision processes." *Journal of Economic Behavior & Organization* 26 (3): 337–352.

Kunreuther, Michel-Kerjan, Doherty, Grace, Klein, and Pauly. 2009. *At War with the Weather: Managing Large-Scale Risks in a New Era of Catastrophes.* The MIT Press.

Kunreuther, Howard. 1996. "Mitigating Disaster Losses Through Insurance." *Journal of Risk and Uncertainty* 12:171–187.

Kunreuther, Howard C, and Erwann O Michel-Kerjan. 2011. *At War with the Weather: Managing Large-Scale Risks in a New Era of Catastrophes.* MIT Press.

Marcoux, Kendra, and Katherine R. H. Wagner. 2024. *Fifty Years of U.S. Natural Disaster Insurance Policy.* Technical report 10431. CESifo Working Paper.

Milgrom, Paul R, and Robert J Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica: Journal of the Econometric Society,* 1089–1122.

Moore, Frances. 2024. *Learning, Catastrophic Risk, and Ambiguity in the Climate Change Era.* National Bureau of Economic Research.

Mumpower, Jeryl L. 1991. "Risk, Ambiguity, Insurance, and the Winner's Curse." *Risk Analysis* 11 (3): 519–522.

Oh, Sangmin, Ishita Sen, and Ana-Maria Tenekedjieva. 2024. *Pricing of Climate Risk Insurance: Regulation and Cross-Subsidies.* Federal Reserve Board Finance and Economics Discussion Series.

Radner, Roy. 2003. "Viscous Demand." *Journal of Economic Theory* 112 (2): 189–231.

Scott, Joe H., Gregory K. Dillon, Michael N. Callahan, Melissa R. Jaffe, Kevin C. Vogler, Julia H. Olszewski, Eva C. Karau, Mitchell T. Lazarz, Karen C. Short, Karin L. Riley, Mark A. Finney, and Isaac C. Grenfell. 2024. *Wildfire Risk to Communities: Spatial Datasets of Landscape-Wide Wildfire Risk Components for the United States, Second Edition.* Forest Service Research Data Archive.

Shi, Linda, and Susanne Moser. 2021. "Transformative Climate Adaptation in the United States: Trends and Prospects." *Science* 372 (6549): eabc8054.

Staiger, Douglas, and James H. Stock. 1997. "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65 (3): 557–586.

Stone, James M. 1973. "A Theory of Capacity and the Insurance of Catastrophe Risks (Part I)." *Journal of Risk and Insurance,* 231–243.

Summers, J. K., A. Lamper, C. McMillion, and L. C. Harwell. 2022. "Observed Changes in the Frequency, Intensity, and Spatial Patterns of Nine Natural Hazards in the United States from 2000 to 2019." *Sustainability* 14 (7).

Taylor, Reid, Madeline Turland, and Joakim A. Weill. 2023. *Climate Change and the Regulation of a Crashing Insurance Market.* Working paper.

U.S. Department of the Treasury. 2023. *Insurance Supervision and Regulation of Climate-Related Risks.* Federal Insurance Office, U.S. Department of the Treasury.

Viscusi, W Kip, Joseph E Harrington Jr, and David EM Sappington. 2018. *Economics of Regulation and Antitrust.* MIT press.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data.* MIT press.

Wylie, David, Natee Amornsiripanitch, John Heilbron, and Kevin Zhao. 2024. *Who Bears Climate-Related Physical Risk?* Federal Reserve Bank of Philadelphia.