Voluntary Purchases and Adverse Selection in the Market for Flood Insurance

Jacob T. Bradt[a,*], Carolyn Kousky[b], Oliver E. J. Wing[c]

[a] *John F. Kennedy School of Government, Harvard University, 79 John F. Kennedy St., Cambridge, MA 02138, United States*
[b] *Wharton Risk Center, University of Pennsylvania, 3819 Chestnut St., Suite 130, Philadelphia, PA 19104, United States*
[c] *School of Geographical Sciences, University of Bristol, Bristol, BS8 1SS, United Kingdom*

Abstract

Flood-related events are the most damaging natural hazard in the United States, yet many households at risk do not have flood insurance. Using detailed policy- and claims-level data from the National Flood Insurance Program (NFIP), we conduct a holistic analysis of the market for publicly provided flood insurance in the U.S., focusing on not only high-risk areas subject to an incomplete mandate requiring the purchase of insurance, but also lower risk areas where no such mandate exists. We are able to better understand determinants of demand for insurance in a setting with voluntary purchase and low take-up and therefore provide a more complete analysis of the market for flood insurance in the U.S. than previous work. In addition to exploring correlates of demand for flood insurance, this paper provides quasi-experimental estimates of households' willingness-to-pay for flood insurance and finds strong evidence to suggest the NFIP failing to utilize full information on flood risk leads to adverse selection in the program.

Keywords: Natural disasters, flood risk, insurance, adverse selection

JEL codes: G52, Q54, Q58

1. Introduction

Flood-related events are the most damaging natural hazard in the United States (U.S.), with flood-related economic losses totaling just under $300 billion in 2017, the largest flood-related loss year on record (NOAA National Centers for Environmental Information, 2021). Floods not only represent a significant share of current damages from natural disasters—accounting for more damage than wildfires, tornadoes, and earthquakes combined—but also are intensifying in many places due to various factors including the increased frequency and severity of storms and heavy rainfall events, as well as sea-level rise in coastal settings (Marsooli et al., 2019; Neumann et al., 2015; Prein et al., 2017; Strauss et al., 2016). When flood waters recede, homeowners are often left with significant amounts of damage. According to an annual survey by the Federal Reserve, most homeowners in the U.S. have very low levels of liquid savings—amounts that would be insufficient for covering post-disaster rebuilding (Federal Reserve Board of

*Corresponding author
Email address: `jbradt@g.harvard.edu` (Jacob T. Bradt)

Governors, 2020). Federal disaster aid—contrary to popular opinion—is quite limited and not consistently available.[1] While access to credit is helpful in financing recovery for some, it is often limited for many low- and moderate-income households (Collier and Ellis, 2020). As such, insurance is typically necessary for households to fully recover financially from disasters. Indeed, a growing body of evidence finds that those with insurance recover faster than those without insurance (Kousky, 2019). Having funds for repair and rebuilding is linked to positive outcomes in terms of emotional well-being, mental health, educational attainment, and the stability of families (McKnight, 2020).

Insurance is thus foundational to recovery, and yet many households at risk of flooding do not have insurance to cover damage from these events. For instance, only about 20% of households flooded in New York City during Hurricane Sandy had flood insurance at the time of landfall in 2012; a similar share of homes in the greater Houston area had flood insurance when Hurricane Harvey hit in 2017 (City of New York, 2013; Yu, 2017). Damage due to flooding is not usually covered under standard property insurance contracts. Since it was first introduced in the late 1960s, the National Flood Insurance Program (NFIP), housed within the Federal Emergency Management Agency (FEMA), has been the primary provider of flood insurance in the U.S., underwriting over 5 million residential policies annually. Recent estimates suggest this represents around 90 to 95% of all residential flood contracts written in the U.S. (Kousky et al., 2018a). Data that we use in this paper on the universe of NFIP insurance contracts indicate that only 3.9% of all housing units in the continental U.S. had an active flood insurance policy in 2019, despite the underpricing of many classes of policies. There is a clear flood insurance gap in the U.S., where many households who are exposed to flood risk do not have coverage.

This paper has two primary objectives. First, considering the important role of flood insurance in ensuring many households' financial well-being, this paper seeks to provide a holistic examination of those factors associated with demand for flood insurance. In so doing, we aim to provide a better understanding of the entire market for flood insurance: previous work examining flood insurance in the U.S. has focused almost exclusively on those areas that FEMA has designated as at highest risk of flooding. By extending this work to focus on the entire market for flood insurance under the NFIP, we are able to provide a more exhaustive set of findings which have important implications for policymakers, shedding light on those attributes of communities characterized by particularly high or low take-up of insurance. Second, this paper aims to better understand households' willingness-to-pay for insurance and the degree of asymmetrically used information in the market for flood insurance. Asymmetrical use of information describes a setting in which homeowners and insurers alike have access to information on relevant buyer attributes, but insurers do not use this information in setting prices.[2] The presence of asymmetrically used information has important implications: a specific first-order concern is that since the NFIP uses broad flood zones to set premiums instead of

[1]Certain federal disaster assistance programs are authorized when the President issues a disaster declaration. Such declarations tend not to be issued for smaller, localized floods. For larger disasters, the President can authorize one or both of two programs administered by the Federal Emergency Management Agency (FEMA): public assistance for local governments and/or individual assistance for households. When authorized, the individual assistance program offers grants up to $33,000, although the average is only a few thousand dollars. The primary federal assistance, however, is a loan from the Small Business Administration, but taking on additional debt is burdensome for some families, and some may be denied a loan due to insufficient credit scores and debt-to-income ratios (Collier and Ellis, 2020).

[2]We distinguish between asymmetrically used and asymmetric information throughout this paper. The former describes a situation in which insurers observe (or could feasibly observe) an attribute of insurance buyers that is correlated with insurance demand and subsequent risk experience

more granular, up-to-date measures of flood risk, the pool of insured households may be adversely selected. If the NFIP is indeed adversely selected, this has important implications for the likely impacts of pricing reforms.

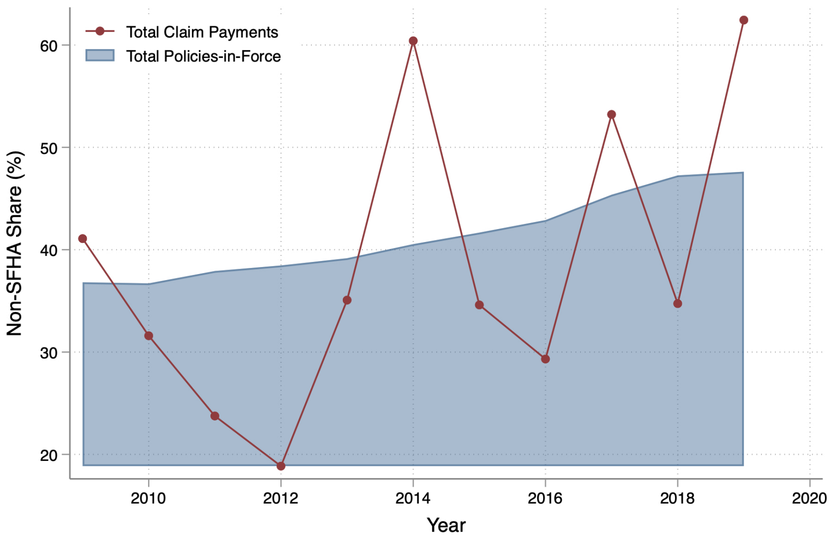
Since the early 1970s, there has been a federal requirement that high risk households in a FEMA-mapped 100-year floodplain—areas with a 1% or greater probability of experiencing a flood each year—with a federally-backed loan or a loan from a federally-regulated lender purchase flood insurance for the life of the loan. Despite the presence of this requirement, take-up of insurance in areas subject to the mandatory purchase requirement remains relatively low: we estimate that the nationwide share of households in 100-year floodplains purchasing insurance was 48.3% in 2019. When not mandated to have flood insurance—such as those without a qualifying loan or outside the 100-year floodplain—purchases of flood insurance are even more infrequent: we estimate that the nationwide take-up rate outside 100-year floodplains, for example, was 2.2% in 2019.

Flooding is not limited to the 100-year floodplain (also called the Special Flood Hazard Area or SFHA) despite the seemingly low demand for insurance in these areas. Understanding flood insurance demand in these lower risk areas where purchases are voluntary is therefore critical for consideration of households' financial recovery from flooding events. Using our data of flood claims between 1978 and 2019, we find that almost a third of all claims are from outside SFHAs. Many recent storms, including Hurricanes Katrina, Ike, Sandy, and Harvey, all led to flooding that extended outside of SFHAs. Beyond just catastrophic events that exceed the 1% annual inundation probability, FEMA SFHAs often do not capture rainfall-related flood risk and this risk is increasing in many places around the country, often outside of SFHAs. Recent modeling accounting for unmapped inland flood risk finds that there are 2.6–3.1 times more households in a 100-year floodplain than suggested by the SFHA designation on FEMA maps (Wing et al., 2018).

Ensuring households' financial resilience to increasing flood risk—including both rainfall flooding and catastrophic events, both of which impact areas outside SFHAs—requires a better understanding of what drives voluntary purchases of flood insurance in these areas where there is no purchase requirement. Surprisingly, there are pockets around the country where a non-trivial number of households purchase flood insurance outside of the SFHA. For example, in the state of Texas, around 70% of the NFIP policies are not in SFHAs and are thus voluntarily purchased. This does not imply that take-up rates outside of SFHAs are high, as there may simply be more households outside of SFHAs. That said, non-SFHA policies represent an important subset of those written under the NFIP. Figure 1 shows the share of all policies-in-force (PIF) and the total value of claim payments for which non-SFHA policies account over the previous decade. As Figure 1 demonstrates—despite low take-up outside SFHAs—these policies represent a large and growing share of overall NFIP policies-in-force, and have accounted for over 60% of the value of total claim payments in certain recent years. Clearly the voluntary segment of the NFIP is non-trivial and plays an important role in the market for flood insurance in the US. Despite the importance of these policies, the determinants of these

but does not to use this attribute in pricing. The latter describes a case where this attribute is unobservable from the insurer's perspective. As discussed in Section 4.4, the efficiency results are the same across these two situations; however, we believe that this is an important distinction in the context of flood insurance in the U.S. since the NFIP can feasibly—and indeed does—observe better measures of household inundation risks than those used in setting premiums.

Figure 1: The share of all policies-in-force (PIF) and the total value of claim payments for which non-SFHA policies account from 2009 to 2019.



Note: Policies are considered in-force each year if they are effective at any point during that year. See Section 3.1 for further details. Total Claim Payments refers to the total value of claim payments. Data are from the universe of residential NFIP policies and claims from 2009 to 2019.

voluntary purchases are largely unaddressed in the literature to date.

While prior research has examined the demand for flood insurance, these papers are unable to separately examine take-up rates in and outside SFHAs and are often unable to distinguish the impacts of the mandatory purchase requirement from drivers of voluntary demand. Nonetheless, these studies find that take-up rates are greater in areas of higher risk and that take-up rates and coverage levels increase with the education and income of homeowners and with higher home values (Atreya et al., 2015; Kousky, 2011; Kriesel and Landry, 2004; Landry and Jahan-Parvar, 2011). One prior study examines voluntary purchases of flood insurance using a survey of residents in Texas and Florida (Brody et al., 2017). The authors find that those voluntarily purchasing a flood policy are more likely to own more expensive homes and have higher levels of education.

Researchers have also found that after a serious flood event or a year with high flood damages, take-up rates for flood insurance increase, but the effect diminishes after a few years (Browne and Hoyt, 2000; Gallagher, 2014). Much of this increase, however, could be driven by a requirement that recipients of federal disaster aid purchase flood insurance. An examination of take-up rates for flood insurance after hurricanes finds that this requirement increased take-up rates by about 5%, with only an additional 1.5% increase in take-up rates not due to this requirement (Kousky, 2017). These flood policies may not be maintained, however: the bump in policies is gone three years after the disaster (Kousky, 2017).

We contribute to the literature on the demand for flood insurance by conducting a holistic analysis of the market for publicly provided flood insurance in the U.S. Specifically, we separately examine policies that are inside SFHAs, and thus likely subject to the mandatory purchase requirement, and policies outside of SFHAs, and voluntarily purchased. We construct a Census tract-level panel dataset of take-up rates for NFIP residential policies that covers the contiguous U.S. between 2009 and 2019, calculating take-up rates for non-SFHA and SFHA households separately. We combine data on flood insurance policies and flood claims from FEMA with estimates of the number of households in and out of the SFHA calculated by combining remote sensing data, Census survey results, and FEMA floodplain boundaries. We integrate these data on the NFIP with novel data on flood risk provided through a partnership with the First Street Foundation; disaster declaration data from FEMA; and a range of sociodemographic, household, and geographic controls from the Census and other sources.

We begin our analysis of insurance demand by exploring those attributes which correlate with demand, both outside and within SFHAs. We then turn to estimating key parameters of interest: households' willingness-to-pay for insurance and the degree of asymmetrically used information in the market for flood insurance. To estimate demand elasticities, we make use of quasi-experimental, policy-induced variation in flood insurance premiums from two recent reform events to account for the potential endogeneity of prices in this setting. Given our access to both claims and policy data, we implement a conceptually intuitive approach to testing for the presence of asymmetrically used information on flood risk that does not require quasi-experimental variation: the "unused observables" test of Finkelstein and Poterba (2014), which we implement using a granular measure of inundation risk derived from the First Street Foundation National Flood Model (FSF-NFM). The logic of this test is that in a symmetric information environment, the insurer will condition prices on characteristics that are correlated with both demand for insurance coverage and the risk of experiencing a loss. If we observe correlations between insurance demand, ex-post measures of risk experience, and our FSF-NFM based measure of inundation risk, this is suggestive of asymmetrically used information on flood risk.

Our analysis offers three main findings. First, we find that households appear to be responsive to prices when opting into flood insurance under the NFIP, though this response is relatively inelastic. Our results suggest that a 1% increase in the average price of non-SFHA policies results in a 0.290% decrease in the probability that a non-SFHA household purchases insurance and that a 1% increase in the average price of SFHA policies results in a 0.327% decrease in the probability that an SFHA household purchases insurance. Second, we find that the market for flood insurance under the NFIP is likely characterized by some level of asymmetrically used information on flood risk. We find that a 1 percentage point increase in the probability that a household is at risk of flooding according to FSF-NFM increases the probability of purchasing insurance by 369% and the average claim amount by 60% outside of SFHAs and increases the probability of purchasing insurance by 48% and the average claim amount by 58% inside of SFHAs. We offer adverse selection as a plausible explanation for our finding of asymmetrically used information in the NFIP. Our third and final takeaway is that there appears to be substantial heterogeneity in insurance take-up across certain observable household attributes, including income and education, which provides an important roadmap for policymakers interested in increasing the financial well-being of households exposed to natural hazard risks by indicating those communities in which take-up rates are particularly low and where targeted efforts may be most needed.

The federal government holds the risk for NFIP flood policies. The approach to rate-setting under the NFIP has not been updated in decades and currently includes many cross-subsidies, as broad flood zones are used to set prices instead of more granular measures of flood risk. In addition, some policies receive discounted prices (for more detail on pricing, see Kousky et al. (2017)). For years where claims exceed the program's revenue, the NFIP has borrowing authority from the U.S. Treasury. Pricing in the program, however, is not designed to cover catastrophic losses, and high claims years beginning in 2005 and continuing with major storms in 2008, 2012, and 2017 have led the program to carry a debt exceeding \$20 billion despite Congressional approval for \$16 billion in debt forgiveness after Hurricane Harvey in 2017 (Horn and Webel, 2021). To address the program's fiscal solvency issues and better align incentives through premiums that reflect property-level risk, various reforms to the rate setting process have been proposed in recent years. As of July 2021, a particular set of reforms to modernize rate-setting—collectively referred to as "Risk Rating 2.0"—are planned to enter into force in October 2021 (Horn, 2021). Our results suggest that these reforms will likely mitigate the asymmetric use of information in the NFIP, potentially reducing adverse selection in the program. If rate increases under these reforms focus on the highest risk areas (note: not the SFHA, but the highest flood risk areas) leaving premiums for lower risk households unchanged (or lower), then it is likely that rate changes can improve the fiscal sustainability of the program and possibly increase overall insurance penetration.

The next section of our paper provides additional background on flood insurance in the U.S. We explain our data sources in Section 3 and then detail the empirical methods that we employ in studying flood insurance demand in Section 4. In Section 5 we present and discuss our results and in Section 6 we provide our primary conclusions.

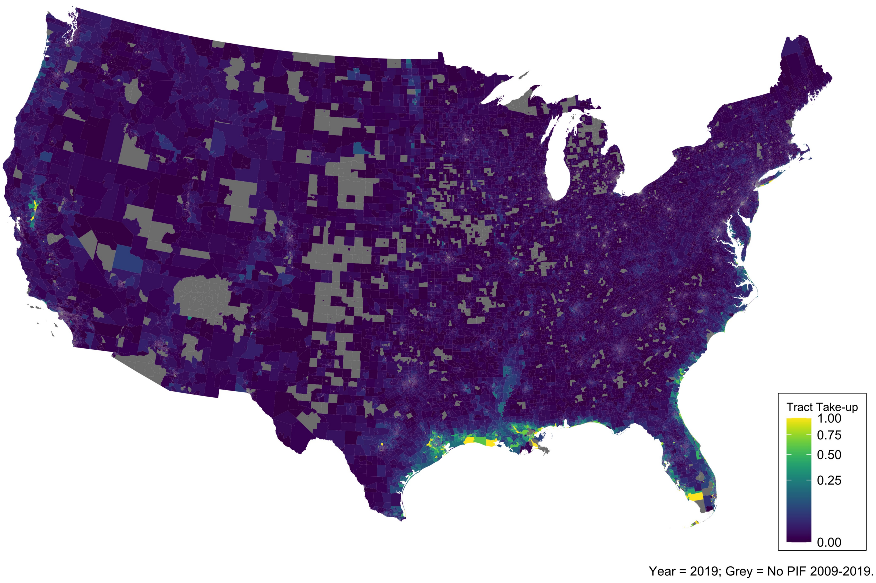
2. Flood Insurance in the United States

The NFIP was founded over 50 years ago in response to a lack of private sector flood insurance. Although there is a small, emerging private market for residential flood coverage today, best estimates suggest that the NFIP still writes 90 to 95% of residential flood policies in the United States (Kousky et al., 2018a). Communities can voluntarily opt-in to the NFIP. When the do, they must adopt minimum floodplain management regulations governing SFHAs; in exchange, all residents are eligible to purchase a policy through the NFIP. Today, over 22,000 communities participate in the program and there are over 5 million policies-in-force nationwide. The program is highly concentrated geographically. Figure 2 shows estimates of Census tract-level non-SFHA and SFHA take-up rates for 2019: take-up is higher in coastal settings, particularly in the Gulf and Atlantic Coasts, and certain inland waterways. For a detailed overview of the NFIP, see Kousky (2018).

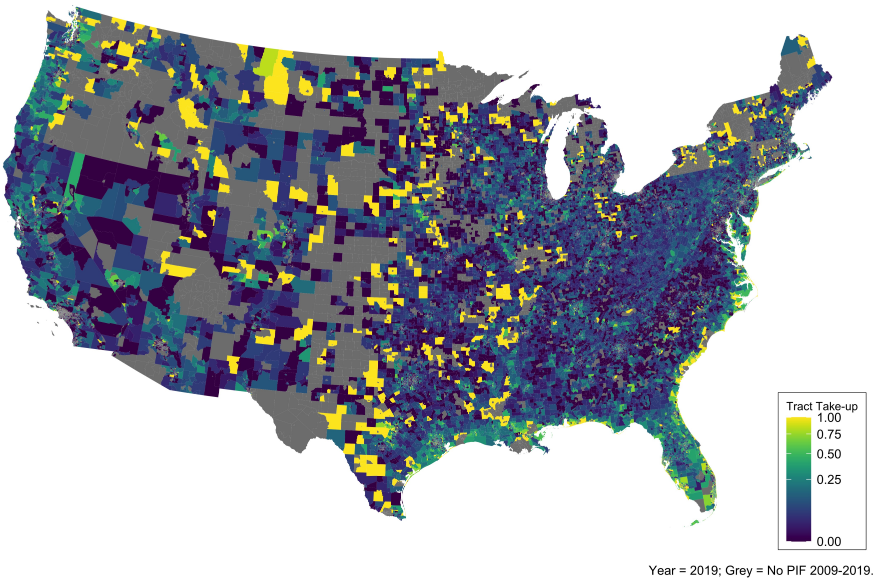
In addition to overseeing the NFIP and underwriting the majority of residential flood insurance policies in the U.S., FEMA is responsible for publishing flood insurance rate maps (FIRMs). The FIRMs demarcate the 100-year floodplain boundary, or the SFHA. Flood zones within SFHAs describe the nature of the underlying inundation

Figure 2: Estimated take-up rates for (a) non-SFHA and (b) SFHA households at the Census tract-level for the year 2019.

(a) Estimated tract-level non-SFHA take-up rates



(b) Estimated tract-level SFHA take-up rates



Note: Scales are constant across Figures 2(a) and 2(b). Grey indicates Census tracts for which there are no relevant policies-in-force (PIF) over the period for which we have data (2009-2019). Take-up rates are calculated by first estimating the number of PIF outside and within SFHAs at the tract-level on an annual basis and then using land use and population data to estimate the spatial distribution of households outside and within SFHAs at the tract-level. There are a large number of tracts for which we estimate full SFHA take-up in 2019, particularly in the center of the country as depicted in Figure 2(b). This is likely driven by measurement error in how we estimate the number of housing units within SFHAs: in many of these cases, the number of estimated housing units and policies-in-force within SFHAs is quite low, so any measurement error will likely result in larger variation in estimated take-up. See Section 3.1 for additional details.

risk, varying primarily based on whether an area is also prone to breaking waves (V zones), or not (A zones). NFIP premiums currently vary by flood zone, as well as other characteristics of the property. For properties with a favorable loss history in the X zone, there is also a lower cost flood policy, referred to as a preferred risk policy (PRP). The NFIP is in the process of modernizing their rate-setting practices, but the zone-based rating is still in use during the time period which we study (Kousky et al., 2017).

The NFIP is not just an insurance program; it also aims to support flood risk reduction and risk communication. One tool for doing this is the Community Rating System (CRS), created in 1990. The CRS is a voluntary program that rewards communities that take measures to reduce their flood risk or improve risk communication with points based on the specific actions or policies adopted. As communities accrue points, they move through levels of the program. At each new level, residents in the SFHA get a 5% reduction on NFIP premiums, up to a 45% reduction for the highest-achieving communities. Outside the SFHA, a 5% reduction is given for the for the first few levels of the program and a 10% reduction for the latter.

Early in the program's history, Congress instituted the mandatory purchase requirement requiring lenders to ensure flood insurance is purchased for properties in SFHAs that are secured by a federally-backed loan or loan made by federally-regulated lenders. Estimates suggest that as of 2019 63% of homeowners have a mortgage nationwide (Neal, 2019). Many states also have disclosure laws that mandate sellers inform potential buyers if the property is located in an SFHA. Beyond SFHAs, take-up rates tend to be very low: we estimate that the nationwide non-SFHA take-up rate in 2019 was 2.2%, compared to 48.3% within SFHAs. It has become an accepted truism that very few people chose to purchase flood insurance voluntarily.

As of early 2021, a set of reforms to the NFIP rating structure is set to enter into force on October 1, 2021. These reforms—collectively referred to as Risk Rating 2.0—will continue a recent trend in the phasing out of NFIP price discounts previously given to policyholders in older homes, which began with legislative action in the form of the Biggert-Waters Flood Insurance Reform Act of 2012 and continued with the Homeowner Flood Insurance Affordability Act of 2014 (Horn, 2021). Importantly, Risk Rating 2.0 will ensure that premiums for individual properties more closely align to their actual flood risk. Rather than using broad flood zones to set rates, premiums will be calculated based on the specific features of an individual property (Horn, 2021). This change will also modernize the approach to price-setting in the program, making use of modern data and models, as well as undo a regressive cross-subsidy in current pricing between low- and high-valued homes that occurs when insurance pricing does not reflect the value of the structure. Risk Rating 2.0 could lead to non-trivial premium changes for some properties both outside and within SFHAs, although any price increases will be phased in at a Congressionally mandated cap of no more than 18% per year. While certain properties might experience price increases, it is also possible that others may experience price decreases under Risk Rating 2.0. To consider the likely impacts of these reforms, we need to know not only how the demand for flood insurance and the cost of insuring homes varies with the price of insurance and underlying inundation risk, but also the risk profile of those properties experiencing price increases or price decreases, both outside and within the SFHA.

3. Data

Our analysis uses two publicly available administrative datasets on flood insurance policies and claims under the NFIP in combination with novel data on flood risk made available through a partnership with the First Street Foundation to construct a panel describing insurance demand at the Census tract-level. We supplement these datasets with data on previous exposure to flooding as well as sociodemographic attributes and household characteristics. The data cover all policies written and claims paid under the NFIP in the contiguous U.S. from 2009 to 2018. Table 1 provides summary statistics of the panel used in our analysis. The methods used to construct the main panel are described in the subsections that follow.

3.1. Flood Insurance Policy and Claims Data

We obtain data on flood insurance policies purchased and claims made through the NFIP from FEMA's Open Data Initiative.[3] Nationwide data are available at the individual policy- and claim-level for all policies issued since 2009 and for all claims dating back to 1978. We limit our analysis to residential policies in the contiguous U.S. In addition to several other important policy-specific attributes, the policy-level data include information on the following: (1) the FEMA flood zone of the policy; (2) policy start and end dates; (3) the Community Rating System (CRS) discount of the policy; and (4) the premium of the policy. The claim data include information on the status, value, and date of initiation for all claims made in the period covered in the data. All monetary values are adjusted to 2019 dollars throughout our analysis.

We restrict the sample of flood insurance policies and claims in several ways. We remove all commercial NFIP policies, leaving only residential buildings in our sample, including both single-family and multi-unit residences. We exclude all policies with coverage exceeding the maximum allowable coverage amount for the applicable occupancy type or with total coverage less than or equal to zero. Since the NFIP policy data include refunds, we drop all policies with a negative premium. Given that some premiums in the data do not align with the rate schedule published by FEMA for residential properties, we exclude those policies with a premium that is less than the first or greater than the ninety-ninth percentile of all premiums (in real terms) in each of the eleven years in our panel. We also remove non-SFHA policies with recorded CRS discounts that exceed the maximum allowable discount for these policies of 10%. Moreover, we drop all those claims with zero or negative reported payouts or damages.

Despite the availability of policy- and claim-level data from the NFIP, geographic identifying information in these data is limited: the finest geographic unit of analysis identified for each policy in the data is the Census tract. Moreover, the publicly available NFIP policy microdata do not provide a mechanism for explicitly linking policies to households over time, nor do these data offer a means to link claims to specific policies. In light of these two limitations—our inability to explicitly link policies over time and to claims and the limited geographic information—

[3] Data available through FEMA's Open Data Initiative may be accessed at https://www.fema.gov/openfema.

we choose to study demand for flood insurance at an aggregate level. Specifically, we focus our analysis on studying insurance take-up rates at the Census tract-level, where we examine take-up separately outside and within SFHAs.

To calculate annual non-SFHA and SFHA take-up rates for each Census tract, we first calculate the number of non-SFHA and SFHA policies-in-force (PIF) by year for each tract. There is no standard definition for what qualifies a policy as being "in-force" each year, so we count policies in our restricted sample as in-force in a given year if they become effective at any point during that year. While this is likely a relatively expansive definition of annual PIF and may therefore be considered an upper bound on the true value of this statistic, we believe this provides a close approximation to the truth and apply this definition consistently throughout.

In addition to annual non-SFHA and SFHA PIF estimates, calculating non-SFHA and SFHA take-up rates requires tract-level estimates of the number of residential units within and outside of SFHAs. Importantly, since we include all residential policies in our annual PIF counts, we must account for the possibility of multiple policies being in-force for a given residential property: we therefore estimate counts of residential units, not residential properties, both within and outside SFHAs at the Census tract-level. Since we do not have access to a nationwide, spatially-referenced dataset of residential units, we combine several publicly available datasets to approximate the number of residential units within and outside of SFHAs at the tract-level. We first use a process known as dasymetric mapping to estimate the spatial distribution of the population in the contiguous U.S. Specifically, this process involves using detailed land cover data to distribute aggregate population counts more accurately over space: given that certain land use classifications are more habitable than others, this allows for a more precise approximation of the spatial distribution of populations than, say, assuming populations—and housing units—are uniformly distributed within aggregate geographies. The estimated spatial population distribution which we use herein is based on the 2011 National Land Cover Dataset and 2010 Decennial Census population counts.[4] We then intersect this estimated spatial population distribution with FEMA's National Flood Hazard Layer as of February 2021 and 2010 Census tracts boundaries to estimate the population share of each Census tract in the contiguous U.S. that falls within SFHAs. Finally, we multiply these estimates of the population share falling within SFHAs by Census tract-level estimates of the number of residential units for the years 2009 to 2019 taken from the U.S. Census Bureau's American Community Survey (ACS), 5-year estimates. These census tract SFHA and non-SFHA residential unit counts are then used to estimate the annual SFHA and non-SFHA take-up rates for each tract-year from 2009 to 2019.[5]

[4] We use the estimated dasymetric allocation of population available through the U.S. Environmental Protection Agency's (EPA) EnviroAtlas. This is estimated by distributing the 2010 Decennial Census population data using land cover classes from the 2011 National Land Cover Dataset (NLCD). The NLCD provides nationwide data on land cover at a 30-meter resolution and is based on satellite data collected and analyzed in 2011. The NLCD land use classes are divided into five categories based on habitability. These data are then intersected with 2010 Census block population data to assign populations to 30-meter grids. Additional information on the EPA's EnviroAtlas is available at https://www.epa.gov/enviroatlas.

[5] There are several important limitations to this approach to estimating the non-SFHA and SFHA housing unit counts at the Census tract-level that are worth noting. First, the assumption that the ratio between population and housing units is constant is strong. In practice there is significant heterogeneity in the relationship between population and the number of housing units over space: certain areas are more likely to have fewer individuals per housing unit than others. Furthermore, while it is possible to account for variation in the total number of housing units at the Census tract-level using annual estimates from the ACS, this approach does not account for variation over time in the underlying estimated spatial population distribution given that we construct static estimates of this distribution based on data from 2010. This approach nevertheless offers a marked improvement over alternative assumptions such as spatial uniformity of populations over aggregate geographies.

There are two important ways in which we further restrict our sample of Census tract take-up rates. First, in certain instances, our approach estimating the number of housing units within and outside of SFHAs at the tract-level produces zero-valued counts. This may arise in several cases: the estimated spatial population distribution may indicate that no individuals reside either inside or outside of SFHAs—or both—for a given tract; there may be no overlap between a given tract boundary and the NFHL; or a tract may not include any residential units. We drop such observations where we estimate zero housing units, since the take-up rate is undefined in these cases. Importantly, this means that there may be tracts where we analyze non-SFHA demand but not SFHA demand or SFHA demand but not non-SFHA demand, though the former is far more common than the latter given the more limited geographic coverage of SFHAs. Additionally, we drop observations for which there are no policies-in-force each year to avoid having our results be driven by a relatively small number of tract-year observations with little overall flood insurance penetration. We believe that this is an important step as it removes areas which are arguably outside of the market for flood insurance under the NFIP, either because there are no enrolled communities within the given geography or inundation risk is trivial.

We aggregate several additional NFIP policy variables to the Census tract-level for non-SFHA and SFHA policies to capture factors that influence the demand for flood insurance. These include the average CRS discount within and outside the SFHA each tract-year and the average total premium cost (premium plus fees net discounts) within and outside the SFHA each tract-year. Note that the individual-level policies data only provide the total premium across all policies under a single contract. We therefore divide the total premium by the number of policies on a single contract to get the average policy-level premium for multi-unit residences with multiple PIF. This policy-level premium measure is what we use when aggregating to the Census tract-level for non-SFHA and SFHA policies.

The individual-level claims data are used to construct two different measures of the cost of insuring the pools of non-SFHA and SFHA policyholders at the tract-level. The first is the annual probability that a given in-force policy experiences a claim, which we calculate separately for areas outside and within SFHAs at the tract-level by first calculating the number of non-SFHA and SFHA claims in our processed claims data for each Census tract-year and dividing this by the relevant tract-year-level PIF estimate. The second is the average claim payment per $1000 of coverage, conditional on a claim occurring, which we calculate by first normalizing each claim payment by the coverage amount of the associated policy—which is included in the claims microdata—and then averaging this for non-SFHA claims and SFHA claims by Census tract and year.

3.2. Measures of Unpriced Flood Risk

We hypothesize that one potential driver of voluntary purchase of flood insurance is household awareness of flood risk not depicted in FEMA FIRMs. If homeowners do indeed have access to information about their inundation risk beyond what is used in setting rates under the NFIP and condition their insurance decisions on this private information, then we would expect to observe a correlation between both demand for insurance and ex-post measures of cost and some measure that proxies this private information.

FEMA FIRMs serve two main purposes. The first is to delineate the boundaries of flood zones, including the extent of SFHAs, thereby communicating inundation risk to households. The second related purpose is to determine the price at which households can purchase insurance under the NFIP—and whether a household is subject to the mandatory purchase requirement. There are several reasons why it is plausible that homeowners may have access to better information on inundation risk than what is communicated in FIRMs. First, the FEMA flood maps do not provide complete coverage. Roughly 60% of the continental United States is currently mapped, yet, even within this 60%, larger rivers tend to be focused on to the detriment of smaller streams and much of the mapping is approximate (Wing et al., 2017, 2018). Further, riverine and coastal flood hazards tend to be the focus of most FIRM studies. While FEMA flood studies sometimes include shallow flooding in areas where a clearly defined channel does not exist, generally, they tend to not fully document areas of potential pluvial or surface water flooding. Residents may be aware of this flood risk through prior experience, communication with other residents, or information provided by other sources beyond FEMA maps.

To proxy for flood risk not captured in FEMA maps, we use the FSF-NFM (Bates et al., 2021), which builds on the modelling of Fathom-U.S. set out in Wing et al. (2017). Like the FEMA NFHL, it delineates the 1-in-100-year flood zone, but FSF-NFM takes advantage of recent advances in remote-sensing data and computational capacity to produce flood maps comprehensively across the country. This top-down flood mapping method leverages available terrain, hydrography, hydrology, and protection databases available in the U.S. to generate fluvial, pluvial, and coastal inundation models with complete coverage of the continental U.S. Rainfall and tidal surge measurements from the National Oceanic and Atmospheric Administration (NOAA) alongside river flows from the U.S. Geological Survey (USGS) are used as inputs to an inundation model which routes the extreme flows through channels defined by the USGS National Hydrography Dataset (where relevant) and over the land surface represented by the ~30 m resolution USGS National Elevation Dataset. Known flood defence structures are assimilated from the U.S. Army Corps of Engineers National Levee Database, alongside hundreds of locally sourced adaptation measures. The model thus produces seemless estimates of the 100-year (and other recurrence probability) floods across the contiguous U.S. This contrasts with the bottom-up approach used by FEMA, where individual local-level models are constructed and stitched together to produce a nationwide map. Many areas remain unmapped by FEMA—for instance, only 33% of river reaches are modelled—and FEMA maps have long been criticized as not incorporating updated data and models in a timely fashion. As such, up to three times as many people are expected to reside in the 100-year flood zone estimated by FSF-NFM as compared with FEMA estimates (Association of State Floodplain Managers, 2020; Wing et al., 2018). Studies show that where high-quality FEMA maps exist, FSF-NFM produces a similar map of 100-year flood extent (Wing et al., 2017, 2018). For more details on the FSF-NFM, see (Bates et al., 2021).

The First Street Foundation uses its flood model to calculate a risk score ranging from 1 to 10 for 142 million properties in the contiguous U.S.[6] This "Flood Factor" captures both the likelihood and the depth of flooding to which a property is at risk as estimated by FSF-NFM. Increasing Flood Factors imply higher flood risk and an increased likelihood of experiencing more severe flooding depths: a property with a Flood Factor of 1 is unlikely to experience any flooding, whereas a property with a Flood Factor of 10 is quite likely to experience relatively extreme flooding over a 30-year cumulative period. We generate measures of inundation risk implied by FSF-NFM for areas both outside and within SFHAs for each Census tract using these Flood Factors given their holistic nature. Specifically, we calculate the probability that a household within a Census tract will experience any flooding according to FSF-NFM, for both areas outside and within SFHAs. To do so, we generate counts of the number of properties with a Flood Factor greater than 1 at the Census tract-level for non-SFHA and SFHA properties and normalize these counts by the estimated number of non-SFHA and SFHA housing units for each tract.[7] While Flood Factors were not available to the public during the entire period of our study, we use them as a proxy for better flood risk information, which might be available to households from experience or better local sources of information.

We also use another approach to measuring perceived flood risk by including prior rainfall events at the tract-level. Specifically, we construct a count of the number of days in the year prior to each tract-year in our panel with precipitation amounts in exceedance of the seventy-fifth percentile for each tract: daily precipitation amounts are retrieved for all measurement stations reported in NOAA's Global Historical Climatology Network (GHCN) database for the contiguous U.S. and nationwide daily precipitation and historical precipitation layers are generated by inverse distance weighting and the number of annual seventy-fifth percentile exceedance days is calculated for all tract-years in our panel. Soil permeability is also included in our model as saturated hydraulic conductivity measurements sourced from the Natural Resources Conservation Service's (NRCS) SSURGO database.

3.3. Prior Flood Experience

There are several tract-level geographic attributes which plausibly influence flood insurance purchase behavior besides flood risk exposure. Since the number of flood insurance purchases tend to be much greater in coastal communities, we control for coastal location by constructing a binary variable which indicates whether a Census tract borders the coast as defined by the U.S. Geological Survey. While the FSF-NFM accounts for storm surge and tidal flooding, two important forms of flooding in coastal areas that play a large role in determining inundation risk in these areas, there are other factors associated with coastal tracts that may impact demand for insurance, such as risk salience. Moreover, it is plausible that individuals residing in coastal tracts are systematically different from those residing in non-coastal tracts, particularly in terms of their risk preferences and risk awareness.[8]

[6] The First Street Foundation calls this risk score a property's "Flood Factor." More information on the Flood Factor is available at https://floodfactor.com/.

[7] It is important to note that our normalization is with respect to the estimated number of housing units both within and outside SFHAs at the Census tract-level, which we calculate according to the method discussed in Section 3.1. As a result, this probability is likely an underestimate, particularly where a large share of properties with Flood Factors greater than 1 have multiple units.

[8] While using a coastal indicator may not allow us to allocate any estimated relationship to these different potential differences between coastal and non-coastal areas that are not accounted for by our other covariates, we believe that it is nonetheless important to account for unobserved

Previous work examining demand for flood insurance finds that take-up is higher after a serious flood event or a year with high flood damages, but the effect dies out after a few years (Browne and Hoyt, 2000; Gallagher, 2014). Much of this increase, however, could be driven by a requirement that recipients of federal disaster aid purchase flood insurance. We obtain data on flood-related federal disaster declarations, including disaster declarations due to coastal storms, hurricanes, severe storms, and other flooding events, from FEMA's Open Data Initiative. To account for the potential for federal disaster aid to drive observed changes in take-up following federal disaster declarations, we also collect data on the authorization of FEMA Individual Assistance (IA) grants for each disaster in our restricted declaration sample.[9] IA grants are not authorized for all disasters: from 2004 to 2011 only 45% of disaster declarations received authorization for the issuance of IA grants (Kousky, 2017). Conditioning on whether an area included in a disaster declaration also receives FEMA IA funding allows us to determine to what extent the previously-observed pattern of increased take-up following disaster declarations is driven by the requirement for IA recipients to procure insurance as a condition of receiving funding. Unfortunately, we are only able to construct counts of approved IA grants at the level of a federal disaster declaration, which may cover multiple counties. These approved IA grant counts likely mask substantial heterogeneity in the number of approved IA grants across counties, let alone Census tracts within affected counties. As a result, we analyze both disaster declaration-level counts of IA grants as well as whether IA is approved for a given disaster declaration when examining the impact of disaster aid on Census tract-level.

3.4. Homeowner and Household Attributes

As discussed in Section 1, numerous studies suggest that demand for insurance and other defensive investments are related to demographic and socioeconomic factors. We use sociodemographic data from the U.S. Census Bureau's American Community Survey, 5-year estimates at the Census tract-level for each tract-year in our panel.[10] This includes median household income, median home value, median year the home was constructed, the percent of households with a mortgage, the share of the population that is non-white, the share of the population under 35, and the total population. These variables are, unfortunately, at a Census tract-level and thus cannot capture any differential sorting patterns in and out of SFHAs.

heterogeneity across coastal and non-coastal areas. To understand whether this coastal indicator obscures ways that preferences, risk, and/or salience affect take-up, we estimate our primary results (Table 2) excluding the coastal indicator and find that coefficient estimates are nearly identical to those reported in Table 2.

[9]Following the declaration of a federal emergency or disaster, FEMA may provide aid in three forms: Individual Assistance (IA), Public Assistance (PA), and Hazard Mitigation Grant Program (HMGP). IA is the primary means of providing direct financial and in-kind federal aid to affected individuals and households and the largest category of IA for individual households is the Individuals and Households Program (IHP). The IHP provides financial and direct assistance to eligible individuals and households who due to a disaster have uninsured or under-insured expenses. IA grants are not available for all disasters; states must specifically request that IA be authorized when requesting a that a federal disaster declaration be issued, and FEMA evaluates whether a given disaster declaration request merits authorizing use of IA funding. Homeowners who receive federal disaster assistance for a flood-damaged home or personal property located within an SFHA are required to purchase flood insurance and maintain coverage for as long as the building exists. In practice, this requirement is difficult to enforce (Webster, 2019).

[10]We choose to use American Community Survey 5-year estimates as opposed to 1-year or 3-year estimates since the 1-year estimates are unavailable for certain low population geographies, the 3-year estimates were discontinued in 2013, and the 5-year estimates provide a high degree of precision. The downside to the 5-year estimates is the limited within unit (tract) variation over time.

3.5. Final Panel Dataset

Summary statistics for our final panel dataset are provided in Table 1. All values are reported either at the Census-tract or Census-tract-zone (non-SFHA of SFHA) level. The mean number of PIF in SFHAs at the Census tract-level in our sample is around 49 and the mean outside SFHAs is 34; both have substantial variation across tract-years. Take-up rates tend to also be higher inside SFHAs than outside SFHAs—by an order of magnitude on average. Importantly, these averages are across tract-years in our sample and therefore differ from prior estimates of nationwide take-up rates. Unsurprisingly, average premiums at the tract-year level are higher for SFHA policies than non-SFHA policies and there is far greater variation across tract-years for the former than the latter.

The probability of a policy experiencing a claim is higher within SFHAs than outside SFHAs: we estimate that 1.2% of policies outside of SFHAs experience a claim in a given year compared with 1.7% of policies within SFHAs for the average tract in our panel. It is important to note however that this estimate may be an upper bound on the true claim probability as it does not account for the potential for a given policy to experience more than one claim in a given year. As indicated by the minimum, median, and maximum values for tract-level non-SFHA and SFHA claim probabilities reported in Table 1, these variables exhibit substantial positive skew: while 95% of tract-year observations estimated to have claim probabilities under 10% for both non-SFHA and SFHA policies, we estimate maximum claim probabilities of 100% for a small subset of tract-years, both for non-SFHA and SFHA policies.[11] The probability that a property within SFHAs has a Flood Factor greater than 1 (i.e., is likely to experience flooding of some kind) is 66%, whereas outside SFHAs this probability is 16%.[12]

Tracts average around 2 flood-related federal disaster declarations and 35,000 approved IA grants in the previous 5 years across all tract-years in our sample (2009-2019). Note that despite reporting tract-year summary statistics in Table 1, the number of disaster declarations varies at the county-year-level in our data and the count of approved IA grants in the previous 5 years varies by disaster declaration, which can cover multiple counties. Both disaster declaration and IA grant counts exhibit significant positive skew: while the median values are 1 and 0, the maximum values are 12 and over 1 million respectively. Naturally, there exists substantial overlap between those areas experiencing a greater number of disaster declarations and those areas with a large number of approved IA grants: tracts within

[11] These instances where we estimate particularly high claim probabilities (e.g., claim probabilities of 100%) are indeed infrequent: the 99th percentiles of our estimated non-SFHA and SFHA claim probabilities are 0.33 and 0.5 across all tract-years in our final dataset, respectively. These high estimates may in-part be driven by the fact that our method produces reasonable upper bounds on the true ex-post probability of experiencing a claim; however, there is reason to believe that we produce reasonable estimates of claim probabilities in these cases. Tract-years exceeding the 95th percentile of estimated claim probabilities for both non-SFHA and SFHA policies are geographically smaller in terms of land area, have larger populations, and experience on average one more federal disaster declaration compared with all other tract-year observations. Geographically-concentrated tracts are likely to have a relatively concentrated spatial distribution of policies, resulting in a higher estimated claim probability conditional on experiencing a flooding event. Moreover, a tract experiencing additional disaster declarations is likely to increase the number of policies within that tract experiencing a claim, which aligns with the observation that tract-years with a higher estimated claim probability tend to experience more disaster declarations on average.

[12] While this is lower than the overlap between the FSF-NFM and FEMA SFHAs estimated in Bates et al. (2021), it is near the lower range of "Critical Success Index" values reported for validation of the model against high-quality local models and the entire catalog of FEMA 1% annual probability flood maps. This underprediction is likely due to the fact that this normalization of the number of residential properties with Flood Factors greater than 1 is with respect to the estimated number of housing units both within and outside SFHAs at the Census tract-level, which we calculate according to the method discussed in Section 3.1. As a result, this probability is likely an underestimate, particularly where properties with multiple units have Flood Factors greater than 1.

Table 1: Summary statistics for Census tract-level panel dataset.

	Mean	St. Dev.	Min	Median	Max
NFIP Attributes					
SFHA Take Up	0.20	0.28	0.00	0.09	1.00
nSFHA Take Up	0.02	0.09	0.00	0.00	1.00
SFHA PIF	48.87	254.05	0.00	1.00	13,713.00
nSFHA PIF	33.95	137.36	0.00	5.00	7,106.00
SFHA Claim Probability	0.02	0.09	0.00	0.00	1.00
nSFHA Claim Probability	0.01	0.07	0.00	0.00	1.00
Avg. CRS Discount SFHA	0.03	0.07	0.00	0.00	0.45
Avg. CRS Discount nSFHA	0.00	0.01	0.00	0.00	0.10
Avg. Policy Cost SFHA (2019 USD)	706.65	887.29	0.00	543.23	35,756.96
Avg. Policy Cost nSFHA (2019 USD)	436.12	257.10	0.00	444.79	8,595.26
Homeowner Attributes					
Pct. of Pop. with College Degree	0.29	0.19	0.00	0.24	1.00
Pct. of Pop. under 35	0.46	0.10	0.00	0.45	1.00
Unemployment Rate	0.08	0.05	0.00	0.07	1.00
Minority Pct. of Pop.	0.26	0.24	0.00	0.17	1.00
Total Population	4,435.13	2,078.65	0.00	4,158.00	72,041.00
Household Attributes					
Median HH Income (2019 USD)	65,266.91	31,709.94	2,499.00	58,168.43	297,918.32
Median Home Value (2019 USD)	259,758.64	213,863.96	8,341.68	189,198.51	2,157,289.39
Median Home Construction Age	42.00	16.75	1.00	40.00	79.00
Pct. of HH with a Mortgage	0.64	0.15	0.00	0.66	1.00
Geography Attributes					
Number of High Precipitation Days	57.77	22.16	1.00	60.50	136.00
DD: All Rel. Cumul. 5-year Lag	1.90	1.88	0.00	1.00	12.00
Total IA Count Cumul. 5-year Lag	34,608.31	120,264.53	0.00	0.00	1,000,652.00
Coastal Tract	0.15	0.36	0.00	0.00	1.00
Total Tract Area Share: Water	0.02	0.05	0.00	0.00	1.00
Soil Permeability (μm/s)	21.65	22.16	0.25	12.71	126.89
SFHA Pr(FF>1)	0.66	0.37	0.00	0.80	1.00
nSFHA Pr(FF>1)	0.16	0.20	0.00	0.09	1.00
Observations			733,033		

Note: This table summarizes our panel for the full set of Census tracts in our data, including tract-years for which there are no policies-in-force. In most of our analysis, we drop observations for which there are no policies-in-force each year to avoid having our results be driven by a relatively small number of tract-year observations with little overall flood insurance penetration. See Section 3.1 for a discussion of the sample restrictions we apply when analyzing demand.

counties exceeding the 95th percentile of cumulative 5-year IA grants average over 4 disaster declarations in the previous 5-year period whereas all other tracts average under 2 disaster declarations in the previous 5 years. Moreover, the number of cumulative IA grants is positively correlated with population: those areas with particularly high levels of approved IA grants also tend to have larger populations.

4. Empirical Methods

We now describe the approach we employ to empirically analyze the market for flood insurance in the U.S. We begin by describing our main demand measure: aggregate take-up rates. Next, we summarize our approach to studying correlates of demand for flood insurance and then outline the methods we use to estimate key parameters of interest: households' willingness-to-pay for insurance and the degree of asymmetrically used information in the market for flood insurance.

As previously noted, homeowners located in a FEMA-mapped 100-year floodplain—areas referred to as SFHAs—with a federally-backed loan or a loan from a federally-regulated lender are mandated to purchase flood insurance for the life of the loan. With perfect compliance, the tract-level take-up rate within SFHAs should be highly correlated with the percentage of mortgages subject to this requirement. Along with households that own their own homes, however, there are also those with a mortgage but not from a lender subject to the requirement; for these homeowners in the SFHA, purchase of flood insurance would be voluntary. Still, these homeowners are likely to at some point have received information on their flood risk from lenders or directly from sellers during the transaction process due to state disclosure laws, which might influence their decision differently than homeowners outside the SFHA. Moreover, the premium schedule differs substantially across these two segments of the NFIP: premiums within SFHAs are higher and vary coarsely with property-level risk, whereas outside SFHAs, premiums vary relatively little and are generally substantially lower than those within SFHAs. We therefore examine aggregate take-up inside and outside SFHAs at the Census tract-level separately. This allows us to not only adequately account for the different decision problems that homeowners within and outside SFHAs face in our aggregate, tract-level analysis of demand, but also better understand the drivers of insurance demand in a setting in which the purchase decision is fully voluntary.

4.1. Extensive Margin Demand

The demand measure that we use is the annual take-up rate for non-SFHA and SFHA residential properties at the Census tract-level. While our decision to focus on aggregate take-up rates is motivated by the two limitations of the underlying NFIP policies and claims microdata discussed in Section 3.1—the inability to explicitly link policies over time and to claims and the limited geographic information available at the policy- and claim-level—doing so has a helpful feature. We can interpret take-up rates as aggregate or average choice probabilities for individuals within a given geography, where the choice is over whether to purchase insurance.

Specifically, let $d_{it} = 1$ if household $i \in \mathcal{J}_j^z$ purchases insurance in period t and 0 otherwise, where $\mathcal{J}_j^z$ is the set of households in zone $z \in \{nSFHA, SFHA\}$ in Census tract j. The take-up rate in zone z in tract j in year t, y_{jt}^z, is therefore given by:

$$y_{jt}^z = \mathbb{E}[d_{it}] = Pr(d_{it} = 1), \forall i \in \mathcal{J}_j^z \tag{1}$$

We can estimate the above expectation as $y_{jt}^z = \frac{1}{|\mathcal{J}_j^z|} \sum_{i \in \mathcal{J}_j^z} d_{it}$, where $|\mathcal{J}_j^z|$ indicates the number of households in set $\mathcal{J}_j^z$. This corresponds to the method used to calculate take-up rates described in section 3.1. The zonal take-up rate for a given Census tract, therefore, models the aggregate extensive margin purchase decision for the associated population of homeowners; it tells us for a random draw from the relevant pool of homeowners, what is the probability that the homeowner chooses to purchase insurance.[13]

4.2. Examining Correlates of Demand

We are interested in better understanding how extensive margin demand relates to several observable factors, including sociodemographic characteristics, household attributes, inundation risk measures, and insurance policy features. As a result, we are interested in modeling the aggregate extensive margin choice probabilities given in Eq. (1) conditional on a set of observables, $\boldsymbol{X}^z_{jt}$. Generally, we model this aggregate conditional choice probability using a linear model:

$$y^z_{jt} = \boldsymbol{X}^{z\,\prime}_{jt}\boldsymbol{\beta}^z + \theta^z_j + \delta^z_{st} + (\lambda^z_c \times t) + \epsilon^z_{jt} \tag{2}$$

where y^z_{jt} is the take-up rate in zone $z \in \{nSFHA, SFHA\}$ in Census tract j in year t, $\boldsymbol{X}^z_{jt}$ is a vector of observable characteristics, θ^z_j is a tract fixed effect, δ^z_{st} is a state-by-year fixed effect, and $(\lambda^z_c \times t)$ is a county-year linear trend. As indicated by the use of superscripts, we estimate models of non-SFHA and SFHA take-up separately across Census tracts. We cluster standard errors at the Census tract-level throughout our analysis to allow for arbitrary dependence of ϵ^z_{jt} over time within tracts and, unless otherwise specified, estimate the linear model described by Eq. (2) using ordinary least squares (OLS).

We include a relatively rich set of fixed effects and time trends to control for unobserved effects on take-up across tracts, over time, and across tracts over time. We include tract-level fixed effects, θ^z_j, to control for time-invariant mean differences in take-up across Census tracts. State-by-year fixed effects, δ^z_{st}, are included to account for unobserved state-specific factors that might affect take-up over time. Examples of such unobserved state-by-year variation might include changes in state-level macroeconomic conditions or changes in state-level requirements on flood risk disclosure. Finally, we include county-by-year time trends to account for county-specific trends, such as changes to building codes which affect development in floodplains and which tend to occur at a more local level.

The observable attributes, $\boldsymbol{X}^z_{jt}$, include measures which vary at different scales. These include time-invariant variables which only vary across Census tracts and time-varying variables which differ either across groups of tracts (counties) or Census tracts. In certain cases where we aim to estimate the relationship between time-invariant attributes and tract-level take-up, we replace the tract fixed effects in Eq. (2) with county-level fixed effects since time-invariant attributes are perfectly collinear with the former. Moreover, we include county-year time trends to account for differ-

[13]While focusing on take-up rates provides important information about aggregate demand, it does not speak to the intensive margin—how much coverage individuals purchase conditional on choosing to purchase insurance. Our processed claims data suggest that the intensive margin decision is perhaps second order: conditional on purchasing insurance, homeowners appear to acquire coverage amounts that exceed expected losses. Examining over 2 million residential claims from 1978 to 2019, we find that 93.4% of claims payments are less than the coverage amount. Moreover, claim amounts account for 22.1% of total coverage amounts on average. Wagner (2019) finds similar intensive margin purchase behavior in a similar sample. It is also the case that those who purchase insurance as a requirement of their mortgage have mandated coverage levels, and thus limited flexibility in this choice. These facts suggest that focusing on extensive margin demand produces more important insights about the market for flood insurance in the U.S. as this accounts for the bulk of the heterogeneity in consumer decision-making in this setting.

ences in groups of tracts over time as opposed to county-year fixed effects since key variables of interest including disaster declaration counts and IA funding amounts vary across county-years as opposed to at the tract-year level.

4.3. Estimating Demand Elasticities

The price elasticity of demand for flood insurance is an important parameter for policymakers and practitioners alike. Given that many of the potential reforms to the NFIP involve changes to the premium schedule, obtaining valid estimates of this parameter is important to better understand the likely impacts of these reforms. Indeed, understanding the price sensitivity of the current pool of insured households under the NFIP has important implications for the likely effects of the planned Risk Rating 2.0 reforms. We can re-write Eq. (2) to give the main estimating equation that we use to estimate price elasticities:

$$y^z_{jt} = \alpha^z p^z_{jt} + \tilde{\boldsymbol{X}}^{z\prime}_{jt}\tilde{\boldsymbol{\beta}}^z + \theta^z_j + \delta^z_{st} + (\lambda^z_c \times t) + \epsilon^z_{jt} \tag{3}$$

where p^z_{jt} is the average price of insurance in zone z in Census tract j in year t; α^z is the parameter of interest, which measures the average effect of a unit increase in flood insurance premiums on take-up; and $\tilde{\boldsymbol{X}}^z_{jt}$ are non-price controls.

In studying demand for flood insurance, there are reasons to be concerned with the endogeneity of price. Insurance policies that are, all else equal, more desirable, such as by offering greater coverage, are more expensive. In addition, policies for high risk properties are more expensive while at the same time, the higher risk may also increase demand. As a result, price will capture aspects of both the desirability of the product and risk levels, which would bias our estimated coefficient on price. We attempt to account for the potential endogeneity of price both outside and within the SFHA by explicitly specifying a reduced form for the price of insurance.

While these endogeneity concerns are clear in the case of demand for insurance within SFHAs—where we observe significant heterogeneity in inundation risks and premiums—they are arguably less applicable in the case of demand outside the SFHA since premiums for these policies are effectively not risk-based. We nonetheless seek to account for the potential endogeneity of non-SFHA premiums in order to explicitly examine whether or not this should be a concern. We discuss further the conceptual arguments why price endogeneity may be less of a concern in the case of non-SFHA demand alongside our results, which we report in Section 5.3.

We assume a linear reduced form for the first stage price regression. To endogenize price, we first need to construct valid instruments. Given the differences in rate-setting across non-SFHA and SFHA policies, we need to identify and construct two separate instruments, one for each of the price variables. To instrument average non-SFHA price, we exploit the exogenous variation in the relative price between primary and non-primary residences introduced by the Homeowner Flood Insurance Affordability Act (HFIAA) of 2014, which introduced a new surcharge to aid in the fiscal sustainability of the NFIP. The amount of this HFIAA surcharge, which began to be assessed on all NFIP policies starting in 2015, is \$25 for primary residences and \$250 for all other properties (i.e., for non-primary residences).[14] Whereas the bulk of historical policy and legislative reforms to the NFIP rate-setting structure apply only to SFHA policies, the differential HFIAA surcharge applies to all policies, both outside and within the SFHA.

The introduction of this surcharge therefore provides exogenous variation in the renewal price of insurance outside of SFHAs: tracts with a greater share of non-primary residences outside SFHAs are likely to observe greater increases in the average non-SFHA price of insurance as a result of these reforms. Given that flood insurance for non-primary residences is relatively more expensive than primary residences following the reforms, it is possible that a non-trivial number of non-primary residences dropped their insurance policies following the reforms. To account for the endogeneity of the share of non-primary residences outside the SFHA following the 2014 legislation, we use a fixed, pre-reform measure of the share of non-primary residences outside the SFHA for each Census tract interacted with an indicator for the year being after the 2014 reforms to instrument non-SFHA price:

$$Z_{jt}^{nSFHA} = 1[t \geq 2015] \times (\%nonPrimary_{j,t=2014}^{nSFHA})$$

where $1[t \geq 2015]$ equals 1 for tract-year observations that are 2015 or later and $(\%nonPrimary_{j,t=2014}^{nSFHA})$ is the number of non-primary residences outside SFHAs in tract j in 2014 as a share of total residences outside SFHAs in tract j. We use Z_{jt}^{nSFHA} to instrument for price in estimating Eq. (3)—for non-SFHA take-up—by two-stage least squares. The identifying assumption underlying the use of this instrument is that the HFIAA surcharge is the only factor that differentially affects non-primary and primary residences in 2015 conditional on our covariates:

$$\mathbb{E}\left[Z_{jt}^{nSFHA} \times \epsilon_{jt}^{nSFHA} \middle| \tilde{\boldsymbol{X}}_{\boldsymbol{jt}}^{\boldsymbol{nSFHA}}, \theta_j^{nSFHA}, \delta_{st}^{nSFHA}, (\lambda_c^{nSFHA} \times t)\right] = 0$$

The logic underlying the use of this instrument is that tracts with a greater share of non-primary residences outside SFHAs pre-HFIAA reforms will face on average higher non-SFHA prices following this policy than tracts with a smaller share of non-primary residences outside SFHAs. That said, the 2014 law also introduced several concurrent reforms for which it is difficult to account, and which may result in a violation of the exclusion restriction. However, to our knowledge, none of these additional reform components altered prices systematically for non-primary and primary residences. Moreover, we are aware of no other cases of reform to non-SFHA premiums or other instances of quasi-experimental variation in the price of non-SFHA flood insurance that would provide valid alternative instruments.[15]

To instrument the average SFHA price, we exploit exogenous variation in the relative price of insurance between pre- and post-FIRM homes mandated by the Biggert-Waters and HFIAA reforms. Specifically, Biggert-Waters and the HFIAA increased the relative price of insuring pre-FIRM homes—older homes that were built before FEMA produced flood maps for a community and which as a result had historically seen discounted premiums

[14] For a detailed description of recent legislative reform to the NFIP, including the HFIAA of 2014, see: https://www.fema.gov/flood-insurance/rules-legislation.

[15] We examine flood insurance rates dating back to 1984 using archived editions of the NFIP manuals and find relatively little policy-induced variation in rate setting beyond the differential surcharge for primary and non-primary homes introduced in the HFIAA of 2014 and the change in the relative price between pre-and post-FIRM homes introduced by Biggert-Waters.

within SFHAs—relative to post-FIRM homes by phasing out discounted premiums beginning in 2013. Wagner (2019) demonstrates that the exogenous variation provided by this policy reform provides a valid instrument when estimating an individual-level model of flood insurance demand. The analogous, aggregate-level variation which we use to instrument average SFHA price is the share of post-FIRM policies in a tract each year interacted with an indicator for the year being post-reform. Given that pre-FIRM households face on average higher prices following these reforms, these households are more likely to drop their insurance policy in the post-reform period. To account for the endogeneity of the share of post-FIRM policies following the 2012 and 2014 reforms, we use a fixed, pre-reform measure of the share of post-FIRM policies for each tract interacted with an indicator for the year being after the 2012 reform to instrument SFHA price:

$$Z_{jt}^{SFHA} = 1[t \geq 2013] \times (\%postFIRM_{j,t=2012}^{SFHA})$$

where $1[t \geq 2013]$ equals 1 for tract-year observations that are 2013 or later and $(\%postFIRM_{j,t=2012}^{SFHA})$ is the number of post-FIRM residences within SFHAs in tract j in 2012 as a share of total residences within SFHAs in tract j. We use Z_{jt}^{SFHA} to instrument for price in estimating Eq. (3)—for SFHA take-up—by two-stage least squares. The identifying assumption underlying the use of this instrument is that the change in relative prices faced by pre- and post-FIRM homes is the only factor differentially affecting these two household types in 2013 conditional on our covariates:

$$\mathbb{E}\left[Z_{jt}^{SFHA} \times \epsilon_{jt}^{SFHA} \middle| \tilde{\boldsymbol{X}}_{\boldsymbol{jt}}^{\boldsymbol{SFHA}}, \theta_j^{SFHA}, \delta_{st}^{SFHA}, (\lambda_c^{SFHA} \times t)\right] = 0$$

The logic underlying the use of this instrument is that tracts with a greater share of pre-FIRM homes in the pre-reform period will face on average higher SFHA prices following the new laws than tracts with a smaller share of pre-FIRM homes.

4.4. Testing for Asymmetrically Used Information

The coarse nature of the NFIP rate-setting structure—particularly outside of SFHAs—and the documented potential for outdated or missing risk information in FEMA flood maps discussed in Section 3.2 means that NFIP prices do not fully reflect all flood risk. If indeed homeowners have access to information about their inundation risk that is not reflected in the premiums charged under the program, then there is reason to be concerned that individuals might sort into and out of the NFIP selectively based on this information: individuals who know that they are a high-inundation risk household may select into the NFIP, whereas individuals who know that they are a low-risk household may forego insurance altogether.

Given the important implications of this potential for adverse selection into the NFIP, we are interested in determining whether this market is characterized by asymmetrically used information about risk.[16] Given our access

[16]Following Finkelstein and Poterba (2014), we emphasize that the test we employ cannot distinguish between asymmetric information—

to both claims and policy data, we implement a conceptually intuitive approach to testing for asymmetrically used information: the "unused observables" test of Finkelstein and Poterba (2014). The logic of this test is that in a symmetric information environment—an environment in which insurers experience no frictions to observing relevant buyer attributes—insurance contracts should be conditioned on characteristics that are correlated with both demand for insurance coverage and the risk of experiencing a loss. As a result, if we find that a characteristic describing the pool of insured households is correlated with demand and a measure of the ex-post risk of loss conditional on prices, this suggests that the insurer does not condition prices on this attribute, which in turn implies the existence of asymmetrically used information (Finkelstein and Poterba, 2014).

We partition our vector of observable characteristics, $\boldsymbol{X}_{\boldsymbol{jt}}^{z\prime} = \begin{bmatrix} \boldsymbol{D}_{\boldsymbol{jt}}^{z\prime} & \boldsymbol{W}_{\boldsymbol{jt}}^{z\prime} \end{bmatrix}$. Let $\boldsymbol{D}_{\boldsymbol{jt}}^{z}$ be a set of attributes used by the NFIP to classify buyer risk types and let $p_{jt}^{z} = f(\boldsymbol{D}_{\boldsymbol{jt}}^{z})$, where $f(\cdot)$ is the NFIP's pricing function. Let $\boldsymbol{W}_{\boldsymbol{jt}}^{z}$ be a set of household attributes which are not used by the NFIP in setting premiums. The estimating equations for the unused observable test are therefore:

$$y_{jt}^{z} = \alpha^{z} p_{jt}^{z} + \boldsymbol{D}_{\boldsymbol{jt}}^{z\prime}\boldsymbol{\eta}^{z} + \boldsymbol{W}_{\boldsymbol{jt}}^{z\prime}\boldsymbol{\gamma}^{z} + \theta_{j}^{z} + \delta_{st}^{z} + (\lambda_{c}^{z} \times t) + \epsilon_{jt}^{z} \tag{4}$$

$$c_{jt}^{z} = \sigma^{z} p_{jt}^{z} + \boldsymbol{D}_{\boldsymbol{jt}}^{z\prime}\boldsymbol{\pi}^{z} + \boldsymbol{W}_{\boldsymbol{jt}}^{z\prime}\boldsymbol{\xi}^{z} + \theta_{j}^{z} + \delta_{st}^{z} + (\lambda_{c}^{z} \times t) + \varepsilon_{jt}^{z} \tag{5}$$

where c_{jt}^{z} is an ex-post measure of the insured populations' risk experience.[17] In practice, we estimate Eq. (5) with two different measures of ex-post cost discussion in Section 3.1: the annual probability that a given policy-in-force experiences a claim and the average claim payment per \$1000 of coverage, conditional on a claim occurring. Rejecting the null hypothesis that $\begin{bmatrix} \boldsymbol{\gamma}^{z\prime} & \boldsymbol{\xi}^{z\prime} \end{bmatrix} = \boldsymbol{0}$ implies a rejection of symmetric information, regardless of the signs on each element of these coefficient vectors.

In our primary implementation of the version of the test characterized by Eqs. (4) and (5), we use our FSF-NFM-derived risk measure as the unused observable, so $\boldsymbol{W}_{\boldsymbol{jt}}^{z}$, $\boldsymbol{\gamma}^{z}$, and $\boldsymbol{\xi}^{z}$ become scalars. As we discuss in Section 3.2, we believe that our FSF-NFM-derived measure serves as a reasonable proxy for inundation risk not measured in the FEMA maps. A positive relationship between this FSF-NFM measure and both take-up and ex-post cost implies asymmetrically used information and is suggestive of adverse selection; a negative relationship similarly implies asymmetrically used information but suggests the potential for advantageous selection. It is worth noting that the

information to which only households have access—and asymmetrically used information. The case of asymmetrically used information is one in which both sides of the market, households and insurers, have access to information as there are no barriers to the insurer observing the relevant buyer attributes; however, the insurer chooses to not condition policy prices on this information. The result is a market equilibrium with efficiency attributes that resemble those of a market in which sellers are prevented from observing a given buyer attribute (Finkelstein and Poterba, 2014). We believe that while asymmetric information is possible in this setting, FEMA generally has access to better risk information than that contained in FIRMs; for instance, FEMA works with catastrophe modelling firms for the purposes of reinsuring its insurance pool and will use these catastrophe modeling firms for the purposes of setting rates under Risk Rating 2.0. Thus, FEMA can in theory access better information on risk when setting prices. While there is potential for both asymmetric and asymmetrically used information in this setting, the case for asymmetrically used information is clear: asymmetrically used information might arise from the coarse premium structure which varies relatively little with true inundation risk, particularly outside of SFHAs.

[17] By "ex-post" we mean some measure that captures a homeowner's risk experience after deciding to purchase insurance. To measure ex-post risk experiences, we use two measures: the annual probability that a given policy-in-force experiences a claim and the average claim payment per \$1000 of coverage. These are ex-post in that they describe a policyholder's risk experience each year conditional on purchasing insurance that year.

unused observables test cannot distinguish between selection and moral hazard; however, we believe that the potential for moral hazard is limited in this setting. We discuss this issue further alongside our results in Section 5.4.

5. Empirical Results and Discussion

5.1. Correlates of Insurance Demand

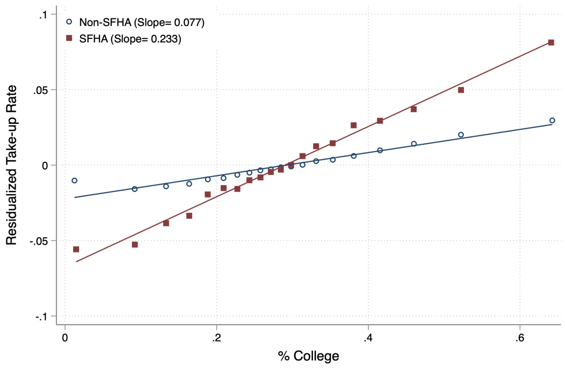
We first examine non-parametric estimates of the conditional expectation functions (CEFs) of non-SFHA and SFHA take-up rates, showing the conditional relationships between these insurance demand variables and a set of observables which we identify ex ante as plausible factors driving flood insurance take-up. Figure 3 shows the non-parametric estimates of the CEFs of non-SFHA and SFHA take-up rates conditional on (a) the share of the tract population with a college degree, (b) the natural logarithm of tract median household income, (c) the share of the tract population below the age of 35—a proxy for age—and (d) the share of households with a mortgage. The estimated CEFs are also conditional on our FSF-NFM-derived measure of inundation risk (which are zone specific), county-level fixed effects, state-by-year fixed effects, and county-by-year time trends to fully capture the effect of each of the covariates of interest after accounting for potential risk-based sorting across tracts as well a rich set of unobserved factors. Figure 3 plots residualized non-SFHA and SFHA take-up rates to facilitate comparisons across the two; the mean tract-level take-up rates in our estimating sample are 0.026 and 0.293 for non-SFHA take-up and SFHA take-up, respectively.

Several interesting correlations emerge in Figure 3. First, take-up of flood insurance appears to positively correlate with education, household income, and age in both the non-SFHA and SFHA segments of the market. This positive relationship is larger in absolute terms for SFHA take-up; however, given differences in levels across these two segments, the relationship between these attributes and insurance demand is stronger in relative terms outside of SFHAs. Unsurprisingly, there is little to no relationship between the share of households with a mortgage and non-SFHA take-up and a positive relationship between SFHA take-up and this variable: this is likely a function of the mandatory purchase requirement driving insurance purchase decisions on the extensive margin in SFHAs. Interestingly, there does appear to be a non-linearity in the relationship between SFHA take-up and the share of households with a mortgage: the positive relationship is largest in magnitude at lower shares of households with a mortgage, tapering off towards the upper end of the empirical support of this variable. While we cannot draw firm conclusions as to what drives this observed non-linearity, potential explanations include differences in the magnitude of information spillovers which affect the salience of flood insurance or differences in lender enforcement capacity in areas with low versus high shares of the population subject to the mandatory purchase requirement.

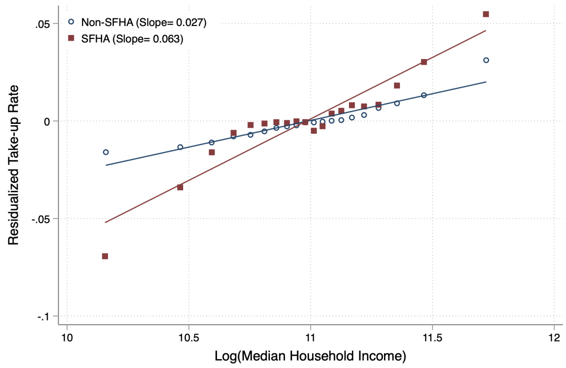
To fully explore correlations between non-SFHA and SFHA take-up at the Census tract-level and the different attributes we observe, we estimate several iterations of Eq. (2) controlling for unobserved heterogeneity at different levels for both take-up rates. Specifically, we estimate versions of Eq. (2) with county-level fixed effects to allow us to estimate coefficients on our time invariant observables as well as tract-level fixed effects to more robustly account

Figure 3: Binned scatterplots showing correlations between Census tract-level non-SFHA and SFHA take-up rates and four sociodemographic and household characteristics of interest.

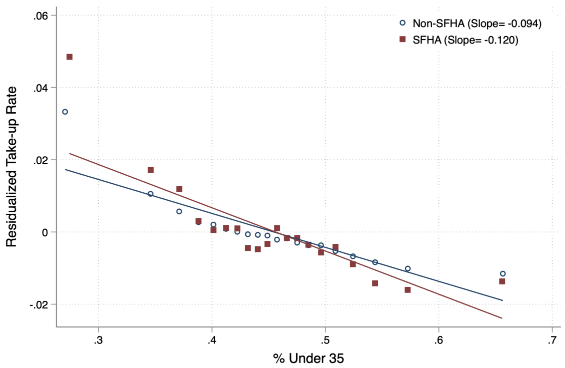
(a) % Population with a College Degree



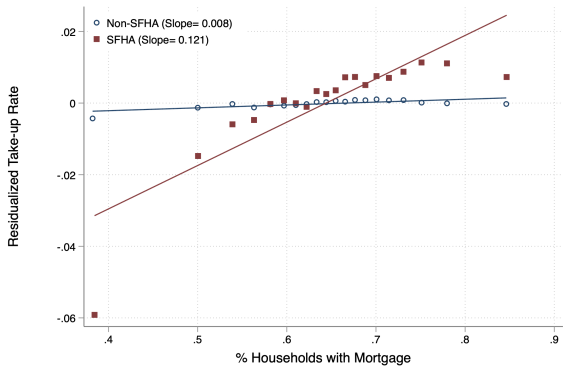
(b) Log(Median Household Income)



(c) % Population under 35



(d) % Households with Mortgage



Note: CEFs are conditional on the probability of a house having a First Street Foundation Risk Score greater than 1; county and state-by-year fixed effects; and county-specific time trends. Plotted take-up rates are residualized on the controls to remove systematic differences between non-SFHA and SFHA take-up rates to facilitate comparison across the two. The mean tract-level take-up rates in our estimating sample are 0.026 and 0.293 for non-SFHA take-up and SFHA take-up, respectively. The four sociodemographic and household characteristics of interest are taken from the American Community Survey, 5-year estimates and are measured at the Census tract-level.

for unobserved unit-level heterogeneity in estimating parameters on our time varying measures. Table 2 reports the main set of results estimating Eq. (2).

Results from our preferred specifications are reported in columns 2, 4, 6, and 8 of Table 2. These are our preferred specifications as they include county-by-year time trends, which accounts for differences across groups of similar Census tracts over time. When examining time-invariant attributes, we focus on columns 2 and 6; when examining time-variant attributes, we focus on columns 4 and 8.

There are several interesting correlations worth noting in the results reported in Table 2. Examining homeowner and household attributes, Table 2 confirms the results in Figure 3: education, population age, and household income all correlate positively with take-up. The average age of construction of a home exhibits a negative relationship with take-up and the median home value relates positively with take-up, both outside and within SFHAs. We continue to observe a strong, positive relationship between the share of households with a mortgage and take-up within SFHAs.

Several of the estimated coefficients on the time-invariant geographic attributes are worth noting. The share of a tract's total area that includes water correlates positively with take-up both outside and within SFHAs. This makes intuitive sense as all else equal areas with more water are likely to be at higher risk of flooding. There is a strong, positive relationship between a Census tract being coastal and take-up of flood insurance both outside and within SFHAs. Given that we include county fixed effects in the specifications containing coastal tract indicators, this implies that Census tracts directly along the coast have higher non-SFHA and SFHA take-up compared to tracts that do not border the coast within the same coastal county. There is a positive correlation between the probability that a household will experience any flooding according to FSF-NFM and take-up, with a particularly strong, positive relationship within SFHAs.

The number of federal disaster declarations experienced by a Census tract in the previous five years does appear to positively affect take-up outside of SFHAs; however, within SFHAs we estimate a null effect. It could be that a disaster that causes flooding beyond SFHAs makes the risk salient and leads to more purchases, but within SFHAs, due to the mandatory purchase requirement, take-up is already fairly high among those inclined to purchase. To account for the potential for federal disaster aid to drive observed changes in take-up due to the requirement that recipients of federal aid maintain insurance, we also include the cumulative number of FEMA IA grants approved at the disaster-level in the previous five years when estimating Eq. (2). Interestingly, areas with higher numbers of FEMA IA grants in the previous five-year period tend to have higher take-up outside of SFHAs, but lower take-up inside of SFHAs. Given this somewhat counterintuitive finding inside of SFHAs, we explore the relationship between disaster declarations, federal disaster aid, and take-up in greater detail. Moreover, since our data on the count of approved IA grants is only at the disaster declaration-level and therefore likely masks substantial heterogeneity in the level of approved IA grants at the Census tract-level, we focus on the impact of authorizing IA funding rather than the level of IA grants approved when exploring the effect of disaster aid further.

We examine the effect of disaster declarations and the authorization of IA funding for a given declaration over

Table 2: Results regressing the Census tract-level outside SFHA take-up rate (columns 1-4) and the Census tract-level within SFHA take-up rate (columns 5-8) on a set of covariates of interest.

	nSFHA Take-up				SFHA Take-up			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Homeowner Attributes								
Pct. of Pop. with College Degree	0.029***	0.028***	0.012***	0.008***	0.142***	0.143***	0.014	0.025***
	(0.003)	(0.003)	(0.002)	(0.002)	(0.013)	(0.013)	(0.009)	(0.009)
Pct. of Pop. under 35	-0.046***	-0.046***	-0.018***	-0.016***	0.031*	0.031*	-0.035***	-0.034***
	(0.004)	(0.004)	(0.003)	(0.003)	(0.016)	(0.017)	(0.010)	(0.010)
Unemployment Rate	0.008	0.010*	-0.003	0.000	-0.070***	-0.081***	0.011	-0.002
	(0.006)	(0.006)	(0.002)	(0.002)	(0.024)	(0.025)	(0.010)	(0.010)
Minority Pct. of Pop.	-0.011***	-0.012***	0.000	0.002	-0.027***	-0.026***	-0.007	0.005
	(0.002)	(0.002)	(0.001)	(0.001)	(0.009)	(0.009)	(0.008)	(0.008)
Log(Total Population)	-0.002**	-0.002**	0.020***	0.015***	-0.021***	-0.021***	0.025***	0.030***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.005)	(0.005)
Household Attributes								
Log(Median HH Income)	0.000	0.000	0.001	0.001	0.012**	0.012**	0.008***	0.004
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.006)	(0.003)	(0.003)
Log(Median Home Value)	0.016***	0.016***	0.003***	0.003***	0.036***	0.036***	0.020***	0.014***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.005)	(0.005)	(0.003)	(0.003)
Median Home Construction Age	-0.000***	-0.000***	-0.000*	-0.000***	0.002***	0.002***	-0.001***	-0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pct. of HH with a Mortgage	0.004	0.005	-0.001	0.000	0.155***	0.155***	0.024***	0.021***
	(0.003)	(0.003)	(0.001)	(0.001)	(0.011)	(0.012)	(0.006)	(0.005)
Geography Attributes								
Number of High Precipitation Days	-0.001	-0.003*	-0.001	-0.003***	0.013	0.010	0.004	-0.004
	(0.001)	(0.001)	(0.000)	(0.000)	(0.009)	(0.009)	(0.003)	(0.003)
DD: All Rel. Cumul. 5-year Lag	0.000***	0.000***	0.001***	0.001***	0.001**	-0.000	0.001***	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Total IA Count Cumul. 5-year Lag	0.002***	0.004***	0.002***	0.005***	-0.004***	-0.004***	-0.002*	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Coastal Tract	0.023***	0.023***			0.081***	0.081***		
	(0.002)	(0.002)			(0.006)	(0.006)		
Total Tract Area Share: Water	0.055***	0.055***			0.143***	0.145***		
	(0.009)	(0.009)			(0.024)	(0.024)		
Soil Permeability	0.002***	0.002***			0.009***	0.009***		
	(0.000)	(0.000)			(0.001)	(0.001)		
nSFHA Pr(FF>1)	0.089***	0.089***						
	(0.004)	(0.004)						
SFHA Pr(FF>1)					0.135***	0.136***		
					(0.005)	(0.005)		
NFIP Policy Attributes								
Avg. CRS Discount nSFHA	0.282***	0.321***	-0.083***	-0.059***				
	(0.040)	(0.043)	(0.009)	(0.009)				
Avg. Policy Cost nSFHA	0.014***	0.014***	-0.015***	-0.013***				
	(0.002)	(0.002)	(0.001)	(0.001)				
Avg. CRS Discount SFHA					0.181***	0.194***	-0.041***	-0.039**
					(0.025)	(0.026)	(0.015)	(0.016)
Avg. Policy Cost SFHA					-0.017***	-0.018***	-0.002***	-0.003***
					(0.001)	(0.001)	(0.001)	(0.001)
County FEs	✓	✓			✓	✓		
Tract FEs			✓	✓			✓	✓
County × Year Trends		✓		✓		✓		✓
State × Year FEs	✓	✓	✓	✓	✓	✓	✓	✓
Observations	609,889	609,889	612,134	612,134	370,497	370,497	372,475	372,475

Note: p values are calculated using a two-sided t-test and heteroskedasticity-robust standard errors clustered at the Census tract-level; $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Heteroskedasticity-robust standard errors clustered at the Census tract-level are reported in parentheses.

time since the initial declaration. Specifically, we estimate the following equation:

$$y^z_{jt} = \sum_{\tau=0}^{\tau=5} (\alpha^z_\tau (PDD_{c,t-\tau} \times IA_{c,t-\tau}) + \gamma^z_\tau PDD_{c,t-\tau}) + \tilde{\boldsymbol{X}}^{z\prime}_{jt}\tilde{\boldsymbol{\beta}}^z + \theta^z_j + \delta^z_{st} + (\lambda^z_c \times t) + \epsilon^z_{jt} \tag{6}$$

where $PDD_{c,t-\tau} = 1$ if the county c in which Census tract j is located experiences at least one federal disaster declaration in year $t-\tau$ and 0 otherwise; $IA_{c,t-\tau} = 1$ if the county c in which Census tract j is located experiences at least one federal disaster declaration for which IA funding is authorized in year $t-\tau$ and 0 otherwise; and $\tilde{\boldsymbol{X}}^z_{jt}$ is the set of non-disaster declaration and IA funding controls. Note that we do not estimate separate coefficients on the $IA_{c,t-\tau} = 1$ indicator; since that only counties under a federal disaster declaration are eligible for IA funding—which means that $IA_{c,t-\tau} = (PDD_{c,t-\tau} \times IA_{c,t-\tau})$—these are equivalent to α^z_τ.

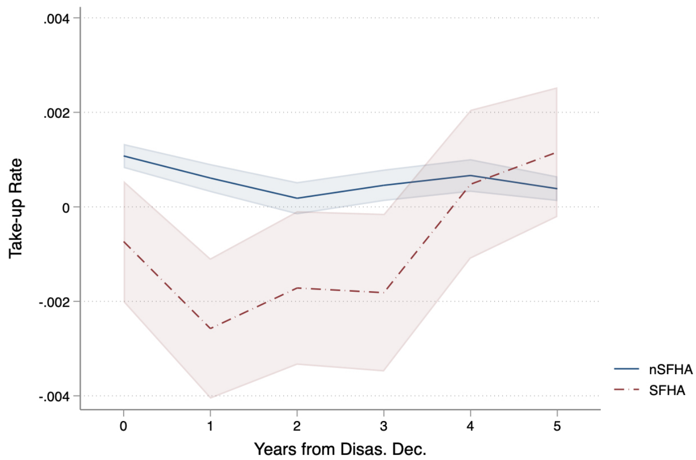
Figure 4a plots the results estimating Eq. (6) excluding the interaction terms, thereby showing the estimated difference in take-up over time from the declaration between Census tracts experiencing a federal disaster declaration and Census tracts not experiencing a disaster declaration. Similar to our results in Table 2, we observe a modest increase in non-SFHA take-up for Census tracts experiencing declarations that declines over time and an initial decrease in SFHA take-up that also decreases in magnitude over time and is close to zero for many of the years.

Figure 4b shows the estimated difference in take-up over time from a federal disaster declaration between Census tracts experiencing a federal disaster declaration for which FEMA IA program funding is authorized and Census tracts experiencing a federal disaster declaration for which IA funding is not authorized. Figure 4b confirms the finding from Table 2 that authorization of IA funding increases take-up of insurance outside of SFHAs, though the effect appears to diminish beginning three years after the occurrence of such a declaration, as found in previous studies (Kousky, 2017). Since many recipients of disaster aid are given a three-year group flood policy, this could explain the results; the group flood policy is not renewed as a standard flood policy once it expires.[18] In the case of SFHA policies, we estimate an initial increase in take-up in the year after a federal disaster declaration receiving IA funding authorization occurs, but beginning two years after such an authorization, the positive effect on take-up diminishes, ultimately becoming negative. This could be driven by some households in SFHAs that received aid but had not previously been insured being forced to purchase coverage and/or it could be related to the salience of the disaster. Either way, the evidence suggests that the requirement that households receiving IA funding within SFHAs purchase and maintain flood insurance may not be well-enforced in the years post-disaster.

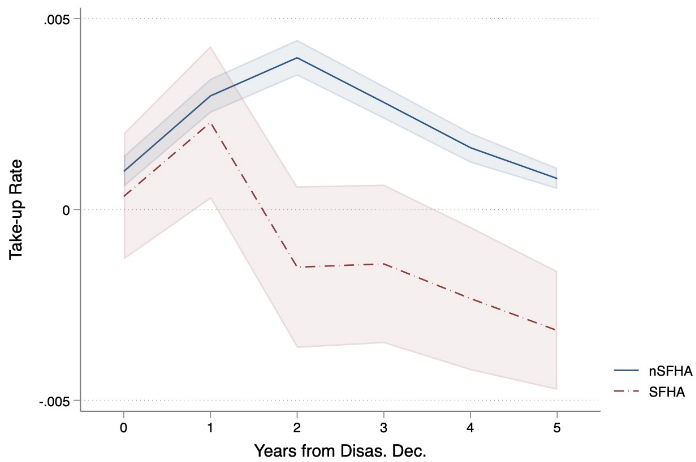
[18]Households experiencing a federal disaster declaration that suffer damage from flooding and have no flood insurance coverage receive a Group Flood Insurance Policy (GFIP) when claiming disaster assistance for their home or personal belongings. Households receiving GFIPs are given coverage for a three-year period beginning 60 days after the date of the federal disaster declaration and the premium is covered by part of the FEMA disaster assistance grant they receive. Households are notified of the termination of their GFIP at the end of the three-year coverage period; however, households are responsible for obtaining and maintaining an individual flood insurance policy on their own following the termination of their GFIP (Federal Emergency Management Agency, 2017).

Figure 4: Estimated relationships between (a) federal disaster declarations and take-up and (b) FEMA Individual Assistance program authorizations, conditional on a federal disaster declaration over time.

(a) Estimated difference over time in take-up between Census tracts experiencing a federal disaster declaration and Census tracts not experiencing a disaster declaration.



(b) Estimated difference over time in take-up between Census tracts experiencing a federal disaster declaration for which FEMA Individual Assistance (IA) program funding is authorized and those experiencing a federal disaster declaration for which IA funding is not authorized.



Note: Figure 4(a) shows the estimated relationship between non-SFHA and SFHA take-up and the occurrence of a federal disaster declaration over time. This relationship is estimated by regressing non-SFHA and SFHA Census tract-level take-up rates on binary variables indicating whether a tract experienced a federal disaster declaration in year $t - \tau$, where $\tau \in \{0, 1, 2, 3, 4, 5\}$; the full set of time-varying, non-disaster declaration control variables listed in Table 2; and a set of tract and state-by-year fixed effects and county-by-year trends. Figure 4(b) shows the estimated relationship between non-SFHA and SFHA take-up and whether a tract experiencing a federal disaster declaration also receives IA program funding authorization over time. This relationship is estimated by regressing non-SFHA and SFHA Census tract-level take-up rates on binary variables indicating whether a tract experiencing a federal disaster declaration also received IA funding in year $t - \tau$, where $\tau \in \{0, 1, 2, 3, 4, 5\}$; the full set of time-varying, non-disaster declaration control variables listed in Table 2; and a set of tract and state-by-year fixed effects and county-by-year trends. 95% confidence intervals are constructed using heteroskedasticity-robust standard errors clustered at the Census tract-level.

5.2. Robustness Checks

We explore the robustness of these results in two ways. First given that our dependent variables are bounded between zero and one, we also estimate a non-linear version of the model described by Eq. (2) using a fractional probit. One potential issue with the linear model described by Eq. (2) is that the assumption that the effect of any covariate is constant over its observed values is perhaps tenuous, unless the range of values over which each covariate is observed is very limited (Papke and Wooldridge, 1996). Since the outcomes in which we are interested are not defined outside the interval between zero and one, the assumption of constant relationships between the covariates and the outcome variables will potentially result in predictions outside that interval. We therefore re-estimate the primary results in Table 2 assuming an alternative functional form for the conditional expectation of take-up, specifically using a fractional probit. Appendix A discusses this fractional probit in detail and provides the estimated marginal effects that we estimate under this alternative model. Overall, we conclude that our results are qualitatively robust across the two models.

We also explore the robustness of our results to errors introduced by our approach to estimating take-up rates. We note several limitations to our approach to estimating the non-SFHA and SFHA housing unit counts at the Census tract-level in Section 3.1. First, the assumption that the ratio between population and housing units is constant is strong. Furthermore, this approach does not account for variation over time in the underlying estimated spatial population distribution given that we construct static estimates of this distribution based on data from 2010. We therefore re-estimate our primary results using annual tract-level non-SFHA and SFHA PIF counts to determine how robust our qualitative findings are to the way in which we calculate take-up rates. Appendix B provides additional details on this robustness check and reports the main results. We find that our results are overall qualitatively robust to using PIF instead of take-up rates as a measure of demand.

5.3. Demand Elasticities

We begin our analysis endogenizing price by examining the first stage for both variables, tract-level average non-SFHA and SFHA premiums. Figure 5a shows the estimated marginal effect of our instrument (the number of policies for non-primary homes outside SFHAs as a share of total policies outside SFHAs in a Census tract in 2014) on average non-SFHA premiums over time. This relationship is estimated from a regression of the average non-SFHA premium on the share of non-SFHA policies for non-primary homes in a tract in 2014 interacted with year indicators; the full set of non-price, time-varying controls included in columns (3) and (4) of Table 2; and a set of tract fixed effects, state-by-year fixed effects, and county time trends. As we can see from Figure 5a, conditional on the covariates, the differential HFIAA surcharge for non-primary residences introduced in 2015 appears to have a strong, positive effect on average non-SFHA premiums in tracts with high pre-HFIAA non-primary shares of non-SFHA policies. Note that Figure 5a reports the estimated marginal effect of a unit increase in the 2014 share of non-SFHA homes that are non-primary residences—i.e., an increase of 100 percentage points. Interestingly, the estimated marginal effect in the first full year following the differential surcharge—2015—is just under \$100, considerably less than the \$225 difference in the HFIAA surcharge for non-primary and primary residences. We interpret this as suggesting a degree of attrition of non-primary residences from the pool of non-SFHA policies after the introduction of the surcharge, validating our use of a pre-reform baseline.

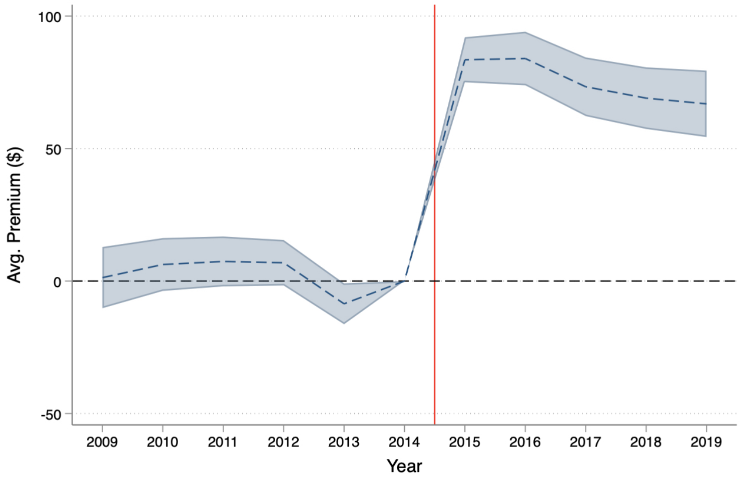
Figure 5b shows the evolution of the relationship between our SFHA instrument for price (the number of post-FIRM policies within SFHAs as a share of total policies within SFHAs in a Census tract in 2012) and the average SFHA premium over time. This relationship is estimated from a regression of the average SFHA premium on the share of post-FIRM policies in a tract in 2012 interacted with year indicators; the full set of non-price, time-varying control variables included in columns (7) and (8) of Table 2; and a set of tract fixed effects, state-by-year fixed effects, and county time trends. As we would expect since the price of post-FIRM policies relative to pre-FIRM policies decreases following the Biggert-Waters reforms, there exists a clear inverse relationship between the average SFHA premium and the 2012 share of post-FIRM policies from 2013 through 2018, particularly when compared to the pre-Biggert-Waters and HFIAA period. Examining Figure 5, both instruments appear to have strong first stages.

Table 3 reports estimates for the coefficient on average premiums in Eq. (3) estimated by both OLS and two-stage least squares. In the case of both non-SFHA and SFHA take-up, the price coefficients estimated by OLS appear to be biased upwards, towards zero. Accounting for price endogeneity results in an order of magnitude decrease in the estimated price coefficient in the case of SFHA demand; however, in the case of non-SFHA take-up, the degree to which the coefficient estimated by OLS is biased upward appears muted: while the two-stage least squares estimate for the coefficient on non-SFHA price is larger in magnitude than that estimated by OLS, we cannot reject the null hypothesis that these coefficients are equal. In line with the apparent muted bias when ignoring the potential for price endogeneity in the non-SFHA case, we cannot reject the null hypothesis that price is a strictly exogenous regressor in the case of non-SFHA take-up. Column 2 of Table 3 reports the Hausman (1978) statistic, which is not sufficient to reject the exogeneity assumption on price for average non-SFHA premiums. Column 4 of Table 3 reports the Hausman (1978) statistic in the case of our SFHA two-stage least squares estimate, which is indeed sufficient to reject the exogeneity assumption in line with our finding of substantial bias in the OLS estimate for SFHA prices. Columns 2 and 4 of Table 3 also report the Kleibergen-Paap F-statistics for our non-SFHA and SFHA two-stage least squares estimates, both of which exceed the rule-of-thumb value of 10. These F-statistics together with Figure 5 provide suggestive evidence that our instruments are not weak (Stock and Yogo, 2005).

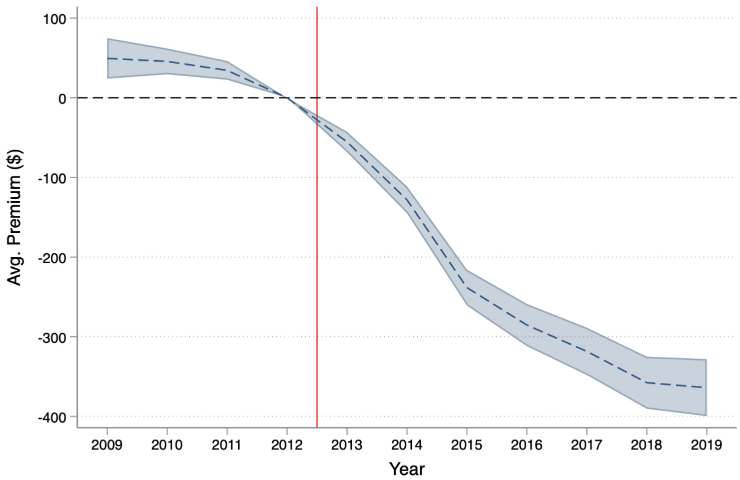
The minor bias apparent in the OLS estimate of the non-SFHA price coefficient is perhaps unsurprising. While the price endogeneity issue is clear in the case of SFHA policies, it is perhaps less apparent in the case of non-SFHA policies. First, as we have noted, there is far less variation in pricing across policies outside of SFHAs relative to those within SFHAs: the standard deviation of premiums is \$686.07 and \$261.22 for SFHA and non-SFHA policies in our sample, respectively, and there is markedly less variation in price per dollar of coverage outside SFHAs. Second, NFIP prices outside SFHAs are essentially not risk-based, dramatically reducing the correlation between risk and price. This is because non-SFHA prices are not based on the elevation of a home and do not reflect variation in the hazard across

Figure 5: Event study plots showing the first-stage used to instrument non-SFHA and SFHA premiums when estimating demand elasticities.

(a) Estimated marginal effect of the number of non-SFHA policies-in-force (PIF) for non-primary homes as a share of total PIF outside the SFHA in a Census tract in 2014 on the average non-SFHA premium over time. The red line indicates the introduction of the Homeowner Flood Insurance Affordability Act (HFIAA) of 2014.



(b) Estimated marginal effect of the number of SFHA PIF for post-FIRM construction homes as a share of total PIF within an SFHA in a Census tract in 2012 on the average SFHA premium over time. The red line indicates the introduction of the first set of relevant reforms, the Biggert-Waters Flood Insurance Reform Act of 2012.



Note: Figure 5(a) is estimated by regressing the average non-SFHA premium on the share of PIF for non-primary homes in a Census tract in 2014 interacted with year indicators; the full set of non-price, time-varying covariates included in columns (3) and (4) of Table 2; and a set of state-by-year fixed effects, tract fixed effects, and county-by-year time trends. Figure 5(b) is estimated by regressing the average SFHA premium on the share of PIF for pre-FIRM construction homes in a Census tract in 2014 interacted with year indicators; the full set of non-price, time-varying covariates included in columns (7) and (8) of Table 2; and a set of state-by-year fixed effects, tract fixed effects, and county-by-year time trends. 95% confidence intervals are constructed using heteroskedasticity-robust standard errors clustered at the Census tract-level.

Table 3: Estimated relationships between price and take-up from baseline results (columns 1 and 3) and from results accounting for price endogeneity (columns 2 and 4).

	nSFHA Take-up		SFHA Take-up	
	(1) OLS	(2) 2SLS	(3) OLS	(4) 2SLS
Avg. Policy Cost nSFHA	-0.013***	-0.015**		
	(0.001)	(0.006)		
Avg. Policy Cost SFHA			-0.003***	-0.077***
			(0.001)	(0.009)
Non-price Controls	✓	✓	✓	✓
Tract FEs	✓	✓	✓	✓
County × Year Trends	✓	✓	✓	✓
State × Year FEs	✓	✓	✓	✓
Hausman (1978) Test		0.058		77.622***
K-P F Stat		250.447		532.856
Elasticity Estimate	-0.260	-0.290	-0.014	-0.327
Observations	612,134	612,134	372,475	372,475

Note: "2SLS" refers to two-stage least squares. p values are calculated using a two-sided t-test and heteroskedasticity-robust standard errors clustered at the Census tract-level; $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Heteroskedasticity-robust standard errors clustered at the Census tract-level are reported in parentheses. Controls include those time-varying covariates included in Table 2.

properties. As such, the core price endogeneity issue should be of less concern for this subset of the market. We still present the instrumented non-SFHA price coefficient as our preferred estimate for this parameter since the two stage least squares estimator also addresses potential errors-in-variable concerns with the average premium variables. Such concerns might arise due to the sample restrictions that we apply to the NFIP policy microdata when constructing our panel or the process of averaging premiums over multi-policy contracts, both of which are described in Section 3.1 and might result in non-trivial measurement error.

We calculate the price elasticities of demand using each of the price coefficients reported in Table 3; however, our preferred elasticity estimates are those based on our two-stage least squares price coefficients reported in columns 2 and 4 of Table 3. These elasticity estimates have important implications for the likely impacts of changes to the NFIP rate-setting structure. Specifically, our results suggest that a 1% increase in the average price of non-SFHA policies results in a 0.290% decrease in the probability that a non-SFHA household purchases insurance and that a 1% increase in the average price of SFHA policies results in a 0.327% decrease in the probability that an SFHA household purchases insurance.

The estimated elasticity of non-SFHA demand of -0.290 is larger in magnitude than that calculated from our OLS estimates. In the case of SFHA take-up, the price elasticity calculated from our two-stage least squares estimate of -0.327 is significantly larger in magnitude from that calculated from our OLS estimates. Importantly, ignoring price endogeneity would result in a vastly different qualitative conclusion: comparing non-SFHA and SFHA price elasticities from our OLS estimates, we would conclude that, while both non-SFHA and SFHA demand appear relatively inelastic, non-SFHA is far more elastic than SFHA demand. Accounting for the potential endogeneity of insurance prices, we see that non-SFHA and SFHA price elasticities are similar in magnitude. Our finding of equally inelastic demand is somewhat surprising: given the presence of the mandatory purchase requirement, we might expect that non-SFHA households would have more elastic demand than SFHA households. This finding may be driven by the fact that not all homes within SFHAs are indeed subject to the mandatory purchase requirement: as of 2019, an estimated 63% of homeowners have a mortgage (Neal, 2019), and prior analysis suggests that the share of households within SFHAs subject to the mandatory purchase requirement may be even lower, closer to 50% (Dixon et al., 2006). Moreover, it is important to note the differences in price and demand levels when comparing estimated demand elasticities across non-SFHA and SFHA households. In general, both elasticity estimates are in line with previous estimates: Wagner (2019) estimates an extensive margin price elasticity of about -0.25 for households within SFHAs.[19]

5.4. Evidence of Asymmetrically Used Information

We test for adverse selection on residual flood risk—that is, flood risk not captured by FEMA FIRMs—by implementing the unused observables test of Finkelstein and Poterba (2014) described in Section 4.4. To implement this test, we use our demand concept of non-SFHA and SFHA tract-level take-up rates and two measures of ex-post risk: the annual probability that a given policy-in-force experiences a claim and the average claim payment per $1000 of coverage, conditional on a claim occurring, both of which we construct separately for non-SFHA and SFHA policies at the Census-tract-level. Our measure of residual risk is the probability that a household within a Census tract will experience any flooding according to FSF-NFM, which we calculate separately for both areas outside and within SFHAs.

We find strong evidence confirming the presence of asymmetrically used information in both segments of the market for flood insurance under the NFIP. The primary results of our test for asymmetrically used information described by Eqs. (4) and (5) are reported in Table 4. Columns (3) and (6) report the coefficient on our measure of residual risk estimating Eq. (4), which align with our results from Table 2. There exists a strong, positive relationship between the FSF-NFM derived estimate of the probability that a household experiences a flood and take-up conditional on prices, non-price controls, and a rich set of fixed effects and time trends, both outside and within SFHAs. We also find strong, positive correlations between our two measures of ex-post cost and our measure of residual risk: both outside and within SFHAs, the probability of a given policy experiencing a claim and the expected claim amount are positively correlated with the FSF-NFM derived probability that a household experiences a flood, again conditional on prices, non-price controls, and a rich set of fixed effects and time trends.

Taken together, Table 4 suggests the potential for informational asymmetries on household-level flood risk

[19]Generally, there are few existing empirical estimates of the price elasticity of demand for flood insurance. Wagner (2019) cites other estimates derived from structural models or panel regression case studies estimate price elasticities between -0.49 and -0.06 (National Research Council, 2015).

Table 4: Results from the unused observables test for asymmetrically used information. Estimated relationships are reported between the probability of a household having a First Street Foundation Flood Factor greater than 1 on claim probabilities (columns 1 and 4), average claim amounts per unit of coverage (columns 2 and 5), and take-up rates (columns 3 and 6).

	nSFHA			SFHA		
	(1) Pr(Claim)	(2) E[Claim/Coverage]	(3) Take-up	(4) Pr(Claim)	(5) E[Claim/Coverage]	(6) Take-up
nSFHA Pr(FF>1)	0.001**	5.651***	0.089***			
	(0.001)	(0.575)	(0.004)			
SFHA Pr(FF>1)				0.008***	8.078***	0.136***
				(0.001)	(0.543)	(0.005)
Controls	✓	✓	✓	✓	✓	✓
County FEs	✓	✓	✓	✓	✓	✓
County × Year Trends	✓	✓	✓	✓	✓	✓
State × Year FEs	✓	✓	✓	✓	✓	✓
Observations	609,889	470,448	609,889	370,497	326,575	370,497

Note: p values are calculated using a two-sided t-test and heteroskedasticity-robust standard errors clustered at the Census tract-level; $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Heteroskedasticity-robust standard errors clustered at the Census tract-level are reported in parentheses. Controls include those time-varying covariates included in Table 2.

both outside and within SFHAs: ceteris paribus, homeowners who face a greater risk of flooding based on the more comprehensive First Street Foundation model are both more likely to purchase insurance and more likely to experience a claim after doing so. The findings in Table 4 imply that a 1 percentage point increase in the probability that a household is at risk of flooding according to FSF-NFM increases the probability of purchasing insurance by 369% and the average claim amount by 60% outside of SFHAs and increases the probability of purchasing insurance by 48% and the average claim amount by 58% inside of SFHAs. This aligns with the qualitative predictions of many of the classic models of equilibrium in insurance markets with asymmetrically used information: due either to adverse selection or moral hazard, those who buy more insurance should be more likely to experience an insured loss (Chiappori and Salanie, 2000).

While the unused observables test allows us to reject the null hypothesis of symmetric information, it does not allow us to determine the relative impacts of moral hazard or adverse selection in driving the observed asymmetries (Finkelstein and Poterba, 2014). However, there are reasons why we believe that our results indicating the presence of asymmetrically used information are likely driven by underlying selection, not moral hazard. Moral hazard occurs when the insured take excessive risk since they are covered for potential losses through their insurance. There is little evidence, even anecdotally, that take-up of flood insurance crowds out private flood risk adaptation investments. While a substantial literature documents that both individuals and municipalities under-invest in flood adaptation investments, there is no indication that this is correlated with insurance purchase (Bakkensen and Mendelsohn, 2016; Mendelsohn et al., 2020; Poussin et al., 2013, 2014). Furthermore, there is growing evidence that homeowners treat flood insurance and individual risk reduction investments as complements, not substitutes (Atreya et al., 2015; Botzen et al., 2019). This suggests that the purchasing of flood insurance is unlikely to result in substantial moral hazard; however, further analysis of this question is certainly warranted.

We also explore heterogeneity in the results reported in Table 4 along several sociodemographic variables, which we argue provides further evidence in favor of our belief that the observed asymmetrically used information results from adverse selection. A small but growing literature documents that differential access to information about environmental risks and disamenities along economic and demographic lines can lead to differences in exposure to these risks (e.g., Hausman and Stolper (2020)). Bakkensen and Ma (2020) model heterogeneity in preferences for flood risk in the presence of limited risk information, finding that low income and minority residents are more likely to move into high-risk areas and that greater public provision of flood risk information reduces socioeconomic differences in exposure to flood risk. In light of these findings, if households condition their insurance purchase decisions on information about their inundation risk not reflected in FEMA maps and our measure of residual risk serves as a reasonable proxy for this information, then we might expect to observe differences in the impact of this proxy on insurance demand across factors such as household income. Furthermore, it is possible that factors such as education affect a households' ability to identify and access sources of private information about their inundation risk, as has been found in the case of other environmental hazards such as air pollution (Flatø, 2020; Liu et al., 2016; Ramírez et al., 2019).

We estimate Eqs. (4) and (5) allowing for heterogeneity in the effect of the probability that a household within a Census tract will experience any flooding according to FSF-NFM by tract-level median household income deciles and tract-level college educated deciles. In addition to estimating coefficients on the interaction between the FSF-NFM derived flood risk measure and income and education decile indicators, we also include un-interacted indicators for these deciles to account for mean differences in take-up across income and education levels. This is particularly important in the case of income deciles since we would expect differences in take-up based on the different budget constraints that individuals face within each of these deciles: lower income individuals are less likely to have disposable income with which to purchase insurance.

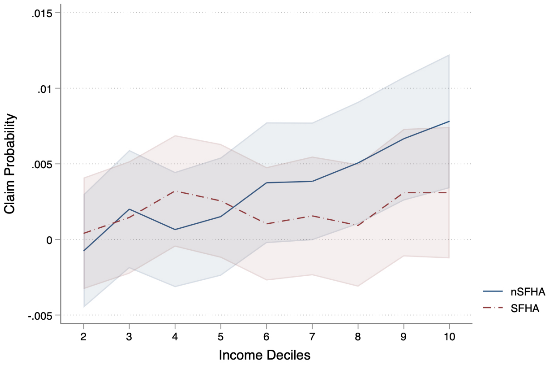
Figure 6 shows the results relative to the first decile of both the income and education variables. We find a striking positive relationship between the income and education levels of a Census tract and the effect of the probability that a household within a Census tract will experience any flooding according to FSF-NFM on take-up; as shown in Figures 6(b) and 6(d), this positive relationship is strong for both non-SFHA and SFHA take-up, though noticeably larger in magnitude in the case of the former. We also find a positive relationship between tract-level income and education levels and the effect of the FSF-NFM risk measure on ex-post claim probabilities for non-SFHA policies; however, there does not appear to be a significant relationship in the case of SFHA policies.

We argue that the estimated heterogeneity in the effect of the probability that a household within a Census tract will experience any flooding according to FSF-NFM on insurance demand along income and education levels is suggestive of active adverse selection.[20] If higher income and higher educated populations are better able to access

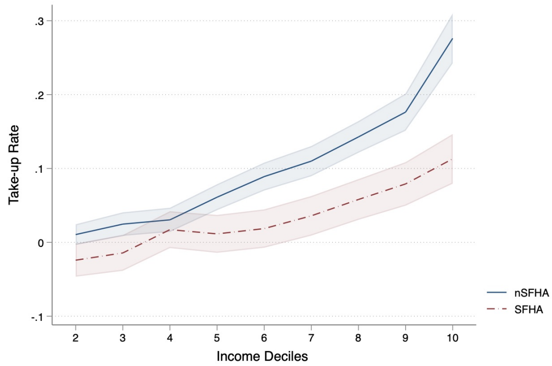
[20]While income has been shown to correlate with preferences for risk (e.g., Cohen and Einav (2007)), this does not invalidate our conclusions regarding adverse selection. Even if the observed heterogeneity in the effect of the FSF-NFM risk measure on take-up across income deciles is driven by underlying differences in risk preferences rather than differences in access to private risk information, this finding is still suggestive of

Figure 6: Effect of the probability of a household having a First Street Foundation Flood Factor greater than 1 on claim probabilities and take-up rates by income (panels a and b) and college educated (panels c and d) deciles.

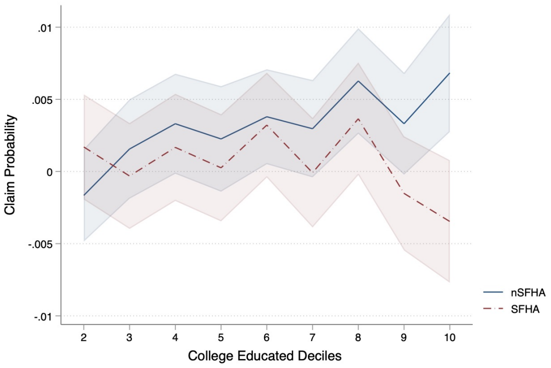
(a) Effect of unmapped risk on claim probabilities by income decile.



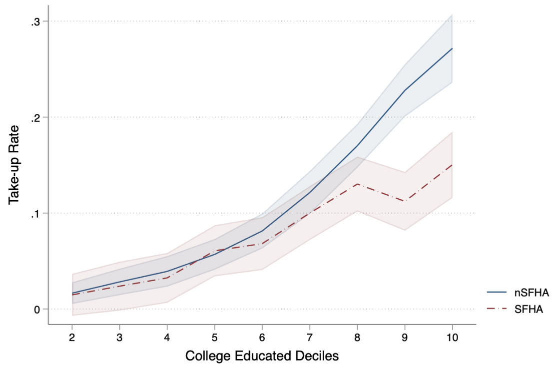
(b) Effect of unmapped risk on take-up rates by income decile.



(c) Effect of unmapped risk on claim probabilities by college educated decile.



(d) Effect of unmapped risk on take-up rates by college educated decile.



Note: Relationships between deciles of Census tract-wide median household income and Census tract-wide share of college educated individuals are reported relative to the first decile. Relationships are estimated by regressing claim probabilities and take-up rates on the full set of interactions between binary variables indicating whether a Census tract falls within a given median household income or college educated population share decile and the probability that a household has a First Street Foundation Flood Factor greater than 1 as well as the full set of time-varying and time-invariant controls included in Table 2, and a set of county fixed effects, state-by-year fixed effects, and county-by-year time trends for both non-SFHA and SFHA claim probabilities and take-up. 95% confidence intervals are constructed using heteroskedasticity-robust standard errors clustered at the Census tract-level.

private risk information, then we would expect these household types to be more likely to condition their insurance purchase decisions on this private information, selecting into or out of flood insurance based on their risk type. Indeed, this is what we observe in Figure 6.

adverse selection, albeit selection based on underlying preferences: higher income households would still need to be accessing private information about their inundation risk in order to select into insurance based on their risk preferences.

We explore heterogeneity in the effect of the probability that a household within a Census tract will experience any flooding according to FSF-NFM on insurance demand over time in our sample. Appendix C reports these results: overall, we do find that the effect of the probability that a household within a Census tract will experience any flooding according to FSF-NFM on insurance demand increases in the later years in our sample, affirming the notion that household awareness of residual risk increases over the sample since we are using a 2020 measure of flood risk. This phenomenon of increased attention to inundation risks is documented in the capitalization of risks in several asset markets—most notably housing markets—particularly in areas exposed to greater inundation risks due to sea level rise (e.g., Bernstein et al. (2019); Keys and Mulder (2020)).

We also explore potential heterogeneity in the presence of asymmetrically used information over space. Appendix D reports these results: overall we do find that the effect of the probability that a household within a Census tract will experience any flooding according to FSF-NFM on both claim probabilities and take-up rates varies over space. New York and New Jersey as well as several other coastal states and states in the Midwest and Great Lakes regions appear to be characterized by a relatively high degree of asymmetrically used information. While the results reported in Appendix D are informative, there are no clear patterns that suggest or confirm potential mechanisms driving the observed heterogeneity over space. This is perhaps a result of the high-level of aggregate geography over which we allow for heterogeneity in the degree of asymmetrically used information. Examination of the mechanisms through which the observed spatial heterogeneity operates is an interesting and important avenue for future research.

6. Conclusion

This paper conducts a holistic analysis of the market for publicly provided flood insurance in the U.S., focusing on not only high-risk areas subject to an incomplete mandate requiring the purchase of insurance, but also lower risk areas where no such mandate exists. In so doing, we are able to better understand determinants of demand for insurance in a setting where purchases are voluntary, and thus provide a more complete analysis of the market for flood insurance in the U.S. than previous work. In addition to exploring correlates of demand for flood insurance, this paper provides estimates of households' willingness-to-pay for flood insurance and finds strong evidence to suggest the presence of adverse selection in this setting. We offer three main takeaways from our analysis.

First, we find that households appear to be responsive to prices when opting into flood insurance under the NFIP. While we estimate relatively inelastic demand for both non-SFHA and SFHA policies, our estimates of these key policy parameters are both negative and statistically distinguishable from zero. Though estimates of the price elasticity of demand for households in SFHAs exist in the literature, we provide what we believe to be the first estimates of the price elasticity of demand for non-SFHA households, which account for a non-trivial—and growing—share of total policies-in-force and claims payments under the NFIP in recent years as indicated by Figure 1. The finding that households respond to changes in price has important policy implications. As new rates are rolled out, take-up of flood insurance might increase in places where prices fall; however, in places where prices go up, fewer households may purchase insurance. Given the importance of flood insurance to households' financial resilience post-flood, policymakers must carefully balance the need to address the NFIP's financial solvency issues with the potential welfare losses due to plausibly inefficiently low insurance take-up.

This leads us to our second main takeaway: the market for flood insurance under the NFIP is characterized by some level of asymmetrically used information on flood risk. While we cannot reject the potential for moral hazard to explain our findings regarding asymmetrically used information, we believe that the strong correlations we observe between insurance demand, ex-post risk experiences, and a measure of un-mapped residual risk are most likely attributable to adverse selection. The fact that we observe a greater impact of our measure which proxies private information on inundation risks in higher income and higher educated communities—which quite likely have a better ability to access private inundation risk information—coupled with existing evidence regarding the substitutability of flood insurance and private adaptation investments suggests that the observed information asymmetry is likely indicative of underlying adverse selection. Work has shown that socioeconomic attributes such as income also correlate with risk aversion (e.g., Cohen and Einav (2007)); the fact that we observe higher income communities—who likely have greater aggregate risk aversion—responding more to the FSF-NFM derived measure of risk provides further justification for our interpretation that the observed information asymmetry is indeed due to some form of adverse selection. Adverse selection into the NFIP has a number of important policy implications which are likely to become increasingly salient in the coming decades: flood risk is expected to increase in many parts of the country due to climate change and the NFIP will need to price for this higher risk to ensure the fiscal solvency of the program.

A third and related takeaway is that there appears to be substantial heterogeneity in insurance take-up across certain observable household attributes. We observe positive relationships between take-up both outside and within SFHAs and education, household income, and age. As our exploration of heterogeneity in our asymmetrically used information results indicates, this may be driven by differences in access to private information or even preferences for risk. Regardless, these findings provide an important roadmap for policymakers interested in increasing the financial well-being of households exposed to natural hazard risks, indicating those communities in which take-up rates are particularly low and where targeted efforts may be most needed. The correlations between take-up rates and key socioeconomic attributes indicate that there may be a need for targeted means-tested assistance for lower-income families in need of flood insurance.

Risk Rating 2.0 will likely mitigate the asymmetric use of information in the NFIP, more closely aligning prices with risk and thus potentially reducing adverse selection in the program. The extent to which these reforms address adverse selection in the program depends on the joint distribution of price changes and flood risk. Our results suggest that higher cost individuals have greater demand for flood insurance conditional on prices; if rate increases focus on the highest risk areas, leaving premiums for lower risk households unchanged (or lower), then it is likely that rate changes can improve the fiscal sustainability of the program and possibly increase overall insurance penetration. Fully quantifying the extent of Risk Rating 2.0's impacts remains a first order empirical question that is outside the scope of this analysis.

FEMA has expressly stated that closing the insurance gap—which refers to the general low take-up of flood insurance—is a motivation for pursuing Risk Rating 2.0.[21] Our findings underscore the importance of understanding the causes and correlates of the insurance gap, particularly in the event that planned reforms do not address—or even compound—this issue. Our hope is that future work will continue to prioritize better understanding the decision to purchase flood insurance with the ultimate goal of improving the overall design of the NFIP and thereby enhancing the financial stability of households at-risk of flooding throughout the U.S.

[21] FEMA lists this as a primary explanation of "why FEMA is undertaking Risk Rating 2.0" in informational materials: https://www.fema.gov/flood-insurance/work-with-nfip/risk-rating.

Acknowledgements

The authors would like to thank the First Street Foundation for the use of their data and their support. We are grateful to Yanjun (Penny) Liao, Philip Mulder, David Alexander, Elizabeth Asche, Niki Crewes, and Tyler Corson-Rikert; seminar participants at the First Street Foundation Flood Lab and Harvard; and Nicholas Rivers and two anonymous referees for their invaluable input on this paper. Remaining errors are our own.

Funding Statement

This paper is based on work supported by the Science and Technology Directorate of the US Department of Homeland Security for the Flood Apex Program under Contract HSHQDC-17-C-B0032. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the US Department of Homeland Security.

References

Association of State Floodplain Managers, 2020. Flood Mapping for the Nation: A Cost Analysis for Completing and Maintaining the Nation's NFIP Flood Map Inventory. Technical Report. Association of State Floodplain Managers.

Atreya, A., Ferreira, S., Michel-Kerjan, E., 2015. What drives households to buy flood insurance? New evidence from Georgia. Ecological Economics 117, 153–161. URL: https://www.sciencedirect.com/science/article/pii/S0921800915002876, doi:10.1016/j.ecolecon.2015.06.024.

Bakkensen, L.A., Ma, L., 2020. Sorting over flood risk and implications for policy reform. Journal of Environmental Economics and Management 104, 102362. URL: https://linkinghub.elsevier.com/retrieve/pii/S0095069620300851, doi:10.1016/j.jeem.2020.102362.

Bakkensen, L.A., Mendelsohn, R.O., 2016. Risk and Adaptation: Evidence from Global Hurricane Damages and Fatalities. Journal of the Association of Environmental and Resource Economists 3, 555–587. URL: https://www.journals.uchicago.edu/doi/abs/10.1086/685908, doi:10.1086/685908. publisher: The University of Chicago Press.

Bates, P.D., Quinn, N., Sampson, C., Smith, A., Wing, O., Sosa, J., Savage, J., Olcese, G., Neal, J., Schumann, G., Giustarini, L., Coxon, G., Porter, J.R., Amodeo, M.F., Chu, Z., Lewis-Gruss, S., Freeman, N.B., Houser, T., Delgado, M., Hamidi, A., Bolliger, I., McCusker, K.E., Emanuel, K., Ferreira, C.M., Khalid, A., Haigh, I.D., Couasnon, A., Kopp, R.E., Hsiang, S., Krajewski, W.F., 2021. Combined Modeling of US Fluvial, Pluvial, and Coastal Flood Hazard Under Current and Future Climates. Water Resources Research 57, e2020WR028673. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020WR028673, doi:https://doi.org/10.1029/2020WR028673. _eprint: https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2020WR028673.

Bernstein, A., Gustafson, M.T., Lewis, R., 2019. Disaster on the horizon: The price effect of sea level rise. Journal of Financial Economics 134, 253–272. URL: https://www.sciencedirect.com/science/article/pii/S0304405X19300807, doi:10.1016/j.jfineco.2019.03.013.

Botzen, W.J.W., Kunreuther, H., Michel-Kerjan, E., 2019. Protecting against disaster risks: Why insurance and prevention may be complements. Journal of Risk and Uncertainty 59, 151–169. URL: http://link.springer.com/10.1007/s11166-019-09312-6, doi:10.1007/s11166-019-09312-6.

Brody, S.D., Highfield, W.E., Wilson, M., Lindell, M.K., Blessing, R., 2017. Understanding the motivations of coastal residents to voluntarily purchase federal flood insurance. Journal of Risk Research 20, 760–775. URL: http://ezp-prod1.hul.harvard.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=121611602&site=ehost-live&scope=site, doi:10.1080/13669877.2015.1119179. publisher: Routledge.

Browne, M.J., Hoyt, R.E., 2000. The Demand for Flood Insurance: Empirical Evidence. Journal of Risk and Uncertainty 20, 291–306. URL: https://doi.org/10.1023/A:1007823631497, doi:10.1023/A:1007823631497.

Chiappori, P., Salanie, B., 2000. Testing for Asymmetric Information in Insurance Markets. Journal of Political Economy 108, 56–78. URL: http://www.jstor.org/stable/10.1086/262111, doi:10.1086/262111. publisher: The University of Chicago Press.

City of New York, 2013. A Stronger, More Resilient New York. Technical Report. Mayor's Office of Long Term PLanning and Sustainability. New York City, NY. URL: https://www1.nyc.gov/site/sirr/report/report.page.

Cohen, A., Einav, L., 2007. Estimating Risk Preferences from Deductible Choice. American Economic Review 97, 745–788. URL: https://www.aeaweb.org/articles?id=10.1257/aer.97.3.745, doi:10.1257/aer.97.3.745.

Collier, B., Ellis, C., 2020. Lending as recovery policy: Evidence from household applications to the U.S. Federal Disaster Loan Program. Working Paper .

Dixon, L., Clancy, N., Seabury, S.A., Overton, A., 2006. The National Flood Insurance Program's Market Penetration Rate: Estimates and Policy Implications. Technical Report. The RAND Corporation.

Federal Emergency Management Agency, 2017. Group Flood Insurance: How It Works | FEMA.gov. URL: https://www.fema.gov/press-release/20210318/group-flood-insurance-how-it-works.

Federal Reserve Board of Governors, 2020. Report on the Economic Well-Being of U.S. Households (SHED). Technical Report. Consumer and Community Research Section, Federal Reserve Board's Division of Consumer and Community Affairs. Washington, DC. URL: https://www.federalreserve.gov/publications/2020-economic-well-being-of-us-households-in-2019-executive-summary.htm.

Finkelstein, A., Poterba, J., 2014. Testing for Asymmetric Information Using "Unused Observables" in Insurance Markets: Evidence From the U.K. Annuity Market. The Journal of Risk and Insurance 81, 709–734. URL: http://www.jstor.org/stable/24548137. publisher: [American Risk and Insurance Association, Wiley].

Flatø, H., 2020. Socioeconomic status, air pollution and desire for local environmental protection in China: insights from national survey data.

Journal of Environmental Planning and Management 63, 49–66. URL: https://doi.org/10.1080/09640568.2019.1630373, doi:10.1080/09640568.2019.1630373. publisher: Routledge _eprint: https://doi.org/10.1080/09640568.2019.1630373.

Gallagher, J., 2014. Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States. American Economic Journal: Applied Economics 6, 206–233. URL: http://www.jstor.org/stable/43189495. publisher: American Economic Association.

Hausman, C., Stolper, S., 2020. Inequality, Information Failures, and Air Pollution. Technical Report w26682. National Bureau of Economic Research. Cambridge, MA. URL: http://www.nber.org/papers/w26682.pdf, doi:10.3386/w26682.

Hausman, J.A., 1978. Specification Tests in Econometrics. Econometrica 46, 1251–1271. URL: http://www.jstor.org/stable/1913827, doi:10.2307/1913827. publisher: [Wiley, Econometric Society].

Honoré, B.E., Tamer, E., 2006. Bounds on Parameters in Panel Dynamic Discrete Choice Models. Econometrica 74, 611–629. URL: http://www.jstor.org/stable/4123096. publisher: [Wiley, Econometric Society].

Horn, D.P., 2021. National Flood Insurance Program: The Current Rating Structure and Risk Rating 2.0. Technical Report R45999. Congressional Research Service.

Horn, D.P., Webel, B., 2021. Introduction to the National Flood Insurance Program (NFIP). R44593. Congressional Research Service.

Keys, B.J., Mulder, P., 2020. Neglected No More: Housing Markets, Mortgage Lending, and Sea Level Rise. Technical Report w27930. National Bureau of Economic Research. URL: http://www.nber.org/papers/w27930, doi:10.3386/w27930.

Kousky, C., 2011. Understanding the Demand for Flood Insurance. Natural Hazards Review 12, 96–110. URL: https://ascelibrary.org/doi/abs/10.1061/%28ASCE%29NH.1527-6996.0000025, doi:10.1061/(ASCE)NH.1527-6996.0000025. publisher: American Society of Civil Engineers.

Kousky, C., 2017. Disasters as Learning Experiences or Disasters as Policy Opportunities? Examining Flood Insurance Purchases after Hurricanes. Risk Analysis 37, 517–530. URL: http://onlinelibrary.wiley.com/doi/abs/10.1111/risa.12646, doi:https://doi.org/10.1111/risa.12646. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/risa.12646.

Kousky, C., 2018. Financing Flood Losses: A Discussion of the National Flood Insurance Program. Risk Management and Insurance Review 21, 11–32. URL: http://onlinelibrary.wiley.com/doi/abs/10.1111/rmir.12090, doi:https://doi.org/10.1111/rmir.12090. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/rmir.12090.

Kousky, C., 2019. The Role of Natural Disaster Insurance in Recovery and Risk Reduction. Annual Review of Resource Economics 11, 399–418. URL: http://www.annualreviews.org/doi/10.1146/annurev-resource-100518-094028, doi:10.1146/annurev-resource-100518-094028. publisher: Annual Reviews.

Kousky, C., Kunreuther, H., Lingle, B., Shabman, L., 2018a. The Emerging Private Residential Flood Insurance Market in the United States. Technical Report. Wharton Risk Management and Decision Processes Center. Philadelphia, PA.

Kousky, C., Lingle, B., Shabman, L., 2017. The Pricing of Flood Insurance. Journal of Extreme Events 04, 1750001. URL: https://www.worldscientific.com/doi/10.1142/S2345737617500014, doi:10.1142/S2345737617500014. publisher: World Scientific Publishing Co.

Kousky, C., Michel-Kerjan, E.O., Raschky, P.A., 2018b. Does federal disaster assistance crowd out flood insurance? Journal of Environmental Economics and Management 87, 150–164. URL: https://www.sciencedirect.com/science/article/pii/S0095069617303479, doi:10.1016/j.jeem.2017.05.010.

Kriesel, W., Landry, C., 2004. Participation in the National Flood Insurance Program: An Empirical Analysis for Coastal Properties. Journal of Risk and Insurance 71, 405–420. URL: http://search.proquest.com/docview/226934059/abstract/44140018014F4B85PQ/1. num Pages: 16 Place: Malvern, United Kingdom Publisher: Blackwell Publishing Ltd.

Landry, C.E., Jahan-Parvar, M.R., 2011. Flood Insurance Coverage in the Coastal Zone. Journal of Risk and Insurance 78, 361–388. URL: http://search.proquest.com/docview/871107199/abstract/C6788F30B0B740A4PQ/1. num Pages: 28 Place: Malvern, United Kingdom Publisher: Blackwell Publishing Ltd.

Liu, X., Zhu, H., Hu, Y., Feng, S., Chu, Y., Wu, Y., Wang, C., Zhang, Y., Yuan, Z., Lu, Y., 2016. Public's Health Risk Awareness on Urban Air Pollution in Chinese Megacities: The Cases of Shanghai, Wuhan and Nanchang. International Journal of Environmental Research and Public Health 13, 845. URL: https://www.mdpi.com/1660-4601/13/9/845, doi:10.3390/ijerph13090845. number: 9 Publisher: Multidisciplinary Digital Publishing Institute.

Marsooli, R., Lin, N., Emanuel, K., Feng, K., 2019. Climate change exacerbates hurricane flood hazards along US Atlantic and Gulf Coasts in spatially varying patterns. Nature Communications 10, 3785. URL: https://www.nature.com/articles/s41467-019-11755-z, doi:10.1038/s41467-019-11755-z. number: 1 Publisher: Nature Publishing Group.

McKnight, A., 2020. Financial resilience among EU households - Publications Office of the EU. Technical Report. Directorate-General for Employment, Social Affairs and Inclusion (European Commission). URL: https://op.europa.eu/en/publication-detail/-/publication/9a2c5116-3296-11ea-ba6e-01aa75ed71a1.

Mendelsohn, R.O., IV, J.G.S., Felson, A., 2020. Are American Coasts Under-Protected? Coastal Management 48, 23–37. URL: https://doi.org/10.1080/08920753.2020.1691482, doi:10.1080/08920753.2020.1691482. publisher: Taylor & Francis _eprint: https://doi.org/10.1080/08920753.2020.1691482.

Mundlak, Y., 1978. On the Pooling of Time Series and Cross Section Data. Econometrica 46, 69–85. URL: http://www.jstor.org/stable/1913646, doi:10.2307/1913646. publisher: [Wiley, Econometric Society].

National Research Council, 2015. Affordability of National Flood Insurance Program Premiums: Report 1. Technical Report. National Research Council. URL: https://www.nap.edu/catalog/21709/affordability-of-national-flood-insurance-program-premiums-report-1, doi:10.17226/21709.

Neal, M., 2019. Mortgage Debt Has Peaked. Why Has the Share of Homeowners with a Mortgage Fallen to a 13-Year Low? URL: https://www.urban.org/urban-wire/mortgage-debt-has-peaked-why-has-share-homeowners-mortgage-fallen-13-year-low.

Neumann, J.E., Emanuel, K., Ravela, S., Ludwig, L., Kirshen, P., Bosma, K., Martinich, J., 2015. Joint effects of storm surge and sea-level rise on US Coasts: new economic estimates of impacts, adaptation, and benefits of mitigation policy. Climatic Change 129, 337–349. URL: https://doi.org/10.1007/s10584-014-1304-z, doi:10.1007/s10584-014-1304-z.

NOAA National Centers for Environmental Information, 2021. U.S. Billion-Dollar Weather and Climate Disasters. URL: https://www.ncdc.noaa.gov/billions/. last Modified: 2021-02-01.

Papke, L.E., Wooldridge, J.M., 1996. Econometric Methods for Fractional Response Variables With an Application to 401 (K) Plan Participation Rates. Journal of Applied Econometrics 11, 619–632. URL: http://www.jstor.org/stable/2285155. publisher: Wiley.

Papke, L.E., Wooldridge, J.M., 2008. Panel data methods for fractional response variables with an application to test pass rates. Journal of Econometrics 145, 121–133. URL: https://www.sciencedirect.com/science/article/pii/S030440760800050X, doi:10.1016/j.jeconom.2008.05.009.

Poussin, J.K., Botzen, W.J.W., Aerts, J.C.J.H., 2013. Stimulating flood damage mitigation through insurance: an assessment of the French CatNat system. Environmental Hazards 12, 258–277. URL: https://doi.org/10.1080/17477891.2013.832650, doi:10.1080/17477891.2013.832650. publisher: Taylor & Francis _eprint: https://doi.org/10.1080/17477891.2013.832650.

Poussin, J.K., Botzen, W.W., Aerts, J.C., 2014. Factors of influence on flood damage mitigation behaviour by households. Environmental Science & Policy 40, 69–77. URL: https://linkinghub.elsevier.com/retrieve/pii/S1462901114000264, doi:10.1016/j.envsci.2014.01.013.

Prein, A.F., Liu, C., Ikeda, K., Trier, S.B., Rasmussen, R.M., Holland, G.J., Clark, M.P., 2017. Increased rainfall volume from future convective storms in the US. Nature Climate Change 7, 880–884. URL: http://www.nature.com/articles/s41558-017-0007-7, doi:10.1038/s41558-017-0007-7. number: 12 Publisher: Nature Publishing Group.

Ramírez, A.S., Ramondt, S., Bogart, K.V., Perez-Zuniga, R., 2019. Public Awareness of Air Pollution and Health Threats: Challenges and Opportunities for Communication Strategies To Improve Environmental Health Literacy. Journal of Health Communication 24, 75–83. URL: https://doi.org/10.1080/10810730.2019.1574320, doi:10.1080/10810730.2019.1574320. publisher: Taylor & Francis _eprint: https://doi.org/10.1080/10810730.2019.1574320.

Stock, J., Yogo, M., 2005. Testing for Weak Instruments in Linear IV Regression, in: Identification and Inference for Econometric Models. Cambridge University Press, New York, NY, pp. 80–108.

Strauss, B., Kopp, R., Sweet, W., Bittermann, K., 2016. Unnatural Coastal Floods: Sea level rise and the human fingerprint on U.S. floods since 1950. Technical Report. Climate Central Research Report,.

Wagner, K., 2019. Adaptation and Adverse Selection in Markets for Natural Disaster Insurance. SSRN Electronic Journal URL: https://www.ssrn.com/abstract=3467329, doi:10.2139/ssrn.3467329.

Webster, E.M., 2019. FEMA Individual Assistance Programs: An Overview. Technical Report R46014. Congressional Research Service.

Wing, O.E.J., Bates, P.D., Sampson, C.C., Smith, A.M., Johnson, K.A., Erickson, T.A., 2017. Validation of a 30 m resolution flood hazard model of the conterminous United States. Water Resources Research 53, 7968–7986. URL: http://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2017WR020917, doi:https://doi.org/10.1002/2017WR020917. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/2017WR020917.

Wing, O.E.J., Bates, P.D., Smith, A.M., Sampson, C.C., Johnson, K.A., Fargione, J., Morefield, P., 2018. Estimates of present and future flood risk in the conterminous United States. Environmental Research Letters 13, 034023. URL: https://doi.org/10.1088/1748-9326/aaac65, doi:10.1088/1748-9326/aaac65. publisher: IOP Publishing.

Yu, R., 2017. Less than 20% Harvey victims have flood insurance as FEMA braces for claims. USA Today URL: https://www.usatoday.com/story/money/2017/08/29/insurance-woes-await-flood-victims-under-covered-houston-area/613239001/.

Appendix A. Fractional Probit Results

Appendix A.1. Fractional Probit Model

Given that our dependent variables are bounded between zero and one, we also estimate a non-linear version of the model described by Eq. (2) using a fractional probit. One potential issue with estimating Eq. (2) by ordinary least squares (OLS) is that the assumption that the effect of any covariate is constant over its observed values is perhaps tenuous, unless the range of values over which each covariate is observed is very limited (Papke and Wooldridge, 1996). Since the outcomes in which we are interested are not defined outside the interval between zero and one, the assumption of constant relationships between the covariates and the outcome variables will potentially result in predictions outside the support of the dependent variables. It may, therefore, be preferable to make a distributional assumption on take-up, y^z_{jt}, given the covariates, and then estimate the parameters of the conditional distribution by maximum likelihood. To test the robustness of our results to our chosen functional form, we explore results under an alternative model where we assume that the conditional distribution of take-up, y^z_{jt}, follows the standard normal distribution, resulting in what the literature refers to as the fractional probit estimator. We then compare the resulting estimated average marginal effects with the main results of our paper. Specifically, we assume that the conditional expectation of the dependent variable y^z_{jt} is given by:

$$\mathbb{E}[y^z_{jt}|\boldsymbol{X^z_{jt}}, \theta^z_j, \delta^z_{st}, (\lambda^z_c \times t)] = \Phi\left(\boldsymbol{X^{z\prime}_{jt}\beta^z} + \theta^z_j + \delta^z_{st} + (\lambda^z_c \times t)\right) \tag{A.1}$$

where $\Phi\cdot$ is the cumulative density function of the standard normal distribution, y^z_{jt} is the take-up rate in zone $z \in \{nSFHA, SFHA\}$ in Census tract j in year t, $\boldsymbol{X^z_{jt}}$ is a vector of observable characteristics, θ^z_j is a tract fixed effect, δ^z_{st} is a state-by-year fixed effect, and $(\lambda^z_c \times t)$ is a county-year linear trend. Eq. (A.1) is the non-linear probit analog of Eq. (2).

Just as with our primary analysis, we estimate versions of Eq. (A.1) which control for different levels of unobserved unit-level heterogeneity given that several of our observable characteristics of interest, $\boldsymbol{X^z_{jt}}$, are time-invariant. Importantly, we cannot estimate the fractional probit described by Eq. (A.1) by simply adding Census tract fixed effects due to the incidental parameters problem associated with dynamic nonlinear panel data models (Honoré and Tamer, 2006). Papke and Wooldridge (2008) provide an alternative method for estimating Eq. (A.1) which allows for unit-specific effects, θ^z_j, to be potentially correlated with $\boldsymbol{X^z_{jt}}$. Specifically, assuming θ^z_j is normal and follows a distribution with mean equal to $\boldsymbol{\bar{X}^{z\prime}_{jt}\psi^z}$, estimates that control for tract fixed effects can be obtained by estimating

$$\mathbb{E}[y^z_{jt}|\boldsymbol{X^z_{jt}}, \theta^z_j, \delta^z_{st}, (\lambda^z_c \times t)] = \Phi\left(\boldsymbol{X^{z\prime}_{jt}\beta^z} + \boldsymbol{\bar{X}^{z\prime}_{jt}\psi^z} + \delta^z_{st} + (\lambda^z_c \times t)\right) \tag{A.2}$$

where $\boldsymbol{\bar{X}^z_{jt}}$ are Census tract-specific means. Conceptually, this is analogous to the Mundlak (1978) estimator, where we can derive the conventional fixed effects estimator in a linear model by assuming that random effects have expected values that are linear functions of the unit-specific means (Papke and Wooldridge, 2008).

We estimate several versions of Eq. (A.2), controlling for unobserved unit effects at the Census tract and county levels separately, by maximum likelihood. The log-likelihood function which we maximize in the estimation procedure is as follows:

$$\ln \mathcal{L} = \sum_{j=1}^{N} \sum_{t=1}^{T} \left[y_{jt}^{z} \ln(G_{jt}^{z}) + (1 - y_{jt}^{z}) \ln(1 - G_{jt}^{z}) \right] \tag{A.3}$$

where $G_{jt}^{z} = \Phi\left(\boldsymbol{X}_{jt}^{z\prime}\boldsymbol{\beta}^{z} + \bar{\boldsymbol{X}}_{jt}^{z\prime}\boldsymbol{\psi}^{z} + \delta_{st}^{z} + (\lambda_{c}^{z} \times t)\right)$; N is the total number of Census tracts; and T is the total number of time periods. Note that the non-linear model described by Eq. (A.2) requires a balanced panel, which requires us to drop a small number of Census tracts from our main estimation sample.

Partial effects, which are directly comparable to parameters in the linear model, can be estimated using the parameters in Eq. (A.2). For an arbitrary continuous covariate, say x, the estimated marginal effect based on Eq. (A.2) is:

$$\frac{\partial \mathbb{E}[y_{jt}^{z} | \boldsymbol{X}_{jt}^{z}, \theta_{j}^{z}, \delta_{st}^{z}, (\lambda_{c}^{z} \times t)]}{\partial x} = \hat{\beta}_{x}^{z} \phi\left(\boldsymbol{X}_{jt}^{z\prime}\hat{\boldsymbol{\beta}}^{z} + \bar{\boldsymbol{X}}_{jt}^{z\prime}\hat{\boldsymbol{\psi}}^{z} + \hat{\delta}_{st}^{z} + (\hat{\lambda}_{c}^{z} \times t)\right) \tag{A.4}$$

where $\phi(\cdot)$ is the probability density function of the standard normal distribution. For some arbitrary discrete covariate, the marginal effect based on Eq. (A.2) is estimated as:

$$\Phi\left(\boldsymbol{X}_{jt}^{z\,(1)\prime}\hat{\boldsymbol{\beta}}^{z} + \bar{\boldsymbol{X}}_{jt}^{z\prime}\hat{\boldsymbol{\psi}}^{z} + \hat{\delta}_{st}^{z} + (\hat{\lambda}_{c}^{z} \times t)\right) - \Phi\left(\boldsymbol{X}_{jt}^{z\,(0)\prime}\hat{\boldsymbol{\beta}}^{z} + \bar{\boldsymbol{X}}_{jt}^{z\prime}\hat{\boldsymbol{\psi}}^{z} + \hat{\delta}_{st}^{z} + (\hat{\lambda}_{c}^{z} \times t)\right) \tag{A.5}$$

where $\boldsymbol{X}_{jt}^{z\,(1)}$ and $\boldsymbol{X}_{jt}^{z\,(0)}$ are two different vectors, each with a different value of the discrete covariate.

Appendix A.2. Fractional Probit Model

Table A.1 reports the results estimating two versions of Eq. (A.2), one accounting for county-level unobserved heterogeneity and the other accounting for tract-level unobserved heterogeneity, for both non-SFHA and SFHA take-up. We report the estimated average marginal effects in columns 3 and 4 for non-SFHA take-up and columns 7 and 8 for SFHA take-up. To ease comparison, we report the estimated marginal effects alongside our main estimates from our primary linear model.

Overall, our primary qualitative findings appear to be robust to an alternative assumption on our model of the conditional expectation of non-SFHA and SFHA take-up rates. As before, we find that take-up rates are positively related to income, education, and age. We also find similar relationships between home value, the median home construction age, and share of households with a mortgage across the linear and probit models. The estimated relationships between the number of federal disaster declarations, IA grant counts, and take-up do appear to be somewhat variable across the different models; however, given the potential for heterogeneous impacts of disaster declarations and the authorization of IA grants over time as indicated in Figure 4, the results in Table A.1 may mask important heterogeneity. Importantly for our tests for asymmetrically used information, we find consistent relationships across the linear and fractional probit models between the probability that a home is at risk of flooding according to FSF-NFM and take-up, for both non-SFHA and SFHA homes.

Table A.1: Results estimating the conditional expectation of the Census tract-level outside SFHA take-up rate (columns 1-4) and the Census tract-level within SFHA take-up rate (columns 5-8) on a set of covariates of interest using the linear (columns 1-2 and 5-6) and fractional probit (columns 3-4 and 7-8) models.

	nSFHA Take-up				SFHA Take-up			
	Linear		Probit		Linear		Probit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Homeowner Attributes								
Pct. of Pop. with College Degree	0.028***	0.008***	0.019***	0.008***	0.143***	0.025***	0.098***	0.013
	(0.003)	(0.002)	(0.003)	(0.002)	(0.013)	(0.009)	(0.014)	(0.010)
Pct. of Pop. under 35	-0.046***	-0.016***	-0.036***	-0.014***	0.031*	-0.034***	0.020	-0.037***
	(0.004)	(0.003)	(0.003)	(0.002)	(0.017)	(0.010)	(0.018)	(0.011)
Unemployment Rate	0.010*	0.000	-0.004	-0.006***	-0.081***	-0.002	-0.079***	0.016
	(0.006)	(0.002)	(0.005)	(0.002)	(0.025)	(0.010)	(0.027)	(0.012)
Minority Pct. of Pop.	-0.012***	0.002	-0.012***	-0.000	-0.026***	0.005	-0.033***	-0.007
	(0.002)	(0.001)	(0.002)	(0.001)	(0.009)	(0.008)	(0.009)	(0.008)
Log(Total Population)	-0.002**	0.015***	-0.001*	0.014***	-0.021***	0.030***	-0.021***	0.024***
	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)	(0.005)	(0.003)	(0.006)
Household Attributes								
Log(Median HH Income)	0.000	0.001	0.002**	0.001	0.012**	0.004	0.016***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.003)	(0.006)	(0.003)
Log(Median Home Value)	0.016***	0.003***	0.008***	0.003***	0.036***	0.014***	0.040***	0.022***
	(0.002)	(0.001)	(0.001)	(0.001)	(0.005)	(0.003)	(0.005)	(0.003)
Median Home Construction Age	-0.000***	-0.000***	-0.000***	-0.000	0.002***	-0.000**	0.002***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Pct. of HH with a Mortgage	0.005	0.000	0.003	-0.002*	0.155***	0.021***	0.143***	0.027***
	(0.003)	(0.001)	(0.003)	(0.001)	(0.012)	(0.005)	(0.013)	(0.006)
Geography Attributes								
Number of High Precipitation Days	-0.003*	-0.003***	-0.001	0.001	0.010	-0.004	0.011	0.011*
	(0.001)	(0.000)	(0.001)	(0.001)	(0.009)	(0.003)	(0.011)	(0.007)
DD: All Rel. Cumul. 5-year Lag	0.000***	0.001***	0.000***	0.001***	-0.000	-0.000	0.001***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)
Total IA Count Cumul. 5-year Lag	0.004***	0.005***	-0.000*	0.003***	-0.004***	-0.002***	-0.009***	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Coastal Tract	0.023***		0.010***		0.081***		0.063***	
	(0.002)		(0.001)		(0.006)		(0.005)	
Total Tract Area Share: Water	0.055***		0.032***		0.145***		0.102***	
	(0.009)		(0.005)		(0.024)		(0.026)	
Soil Permeability	0.002***		0.001***		0.009***		0.008***	
	(0.000)		(0.000)		(0.001)		(0.001)	
nSFHA Pr(FF>1)	0.089***		0.037***					
	(0.004)		(0.002)					
SFHA Pr(FF>1)					0.136***		0.129***	
					(0.005)		(0.005)	
NFIP Policy Attributes								
Avg. CRS Discount nSFHA	0.321***	-0.059***	0.218***	-0.030***				
	(0.043)	(0.009)	(0.021)	(0.006)				
Avg. Policy Cost nSFHA	0.014***	-0.013***	0.012***	-0.011***				
	(0.002)	(0.001)	(0.001)	(0.001)				
Avg. CRS Discount SFHA					0.194***	-0.039**	0.171***	-0.005
					(0.026)	(0.016)	(0.024)	(0.015)
Avg. Policy Cost SFHA					-0.018***	-0.003***	-0.019***	-0.003***
					(0.001)	(0.001)	(0.002)	(0.001)
County FEs	✓		✓		✓		✓	
Tract FEs		✓				✓		
Tract Means				✓				✓
County × Year Trends	✓	✓	✓	✓	✓	✓	✓	✓
State × Year FEs	✓	✓	✓	✓	✓	✓	✓	✓
Observations	609,889	612,134	577,768	581,490	370,497	372,475	360,014	362,948

Note: p values are calculated using a two-sided t-test and heteroskedasticity-robust standard errors clustered at the Census tract-level; $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. Heteroskedasticity-robust standard errors clustered at the Census tract-level are reported in parentheses.

Appendix B. Policies-in-Force Results

We note several limitations to our approach to estimating the non-SFHA and SFHA housing unit counts at the Census tract-level in Section 3.1. First, the assumption that the ratio between population and housing units is constant is strong. In practice there is likely significant heterogeneity in the relationship between population and the number of housing units over space: certain areas are more likely to have fewer individuals per housing unit than others. Furthermore, while it is possible to account for variation in the total number of housing units at the Census tract-level using annual estimates from the ACS, this approach does not account for variation over time in the underlying estimated spatial population distribution given that we construct static estimates of this distribution based on data from 2010.

The static nature of several elements which enter our calculation of these housing unit counts and the potential for measurement error introduced at each step of the calculation—for instance, remote sensing data used to categorize land use contain non-trivial margins of error—motivate an additional robustness check. We re-estimate our primary results using annual tract-level non-SFHA and SFHA policies-in-force (PIF) counts to determine how robust our qualitative findings are to the way in which we calculate take-up rates. Specifically, we model PIF using a linear model that is analogous to that described in Eq. (2):

$$\tilde{y}^z_{jt} = \boldsymbol{X}^{z\prime}_{jt}\boldsymbol{\beta}^z + \theta^z_j + \delta^z_{st} + (\lambda^z_c \times t) + \mu^z_{jt} \tag{B.1}$$

where $\tilde{y}^z_{jt}$ is the number of PIF in zone $z \in \{nSFHA, SFHA\}$ in Census tract j in year t, X^z_{jt} is a vector of observable characteristics, δ^z_{st} is a tract fixed effect, δ^z_{st} is a state-by-year fixed effect, and $(\lambda^z_c \times t)$ is a county-year linear trend. As indicated by the use of superscripts, we again estimate models of non-SFHA and SFHA take-up separately across Census tracts. We cluster standard errors at the Census tract-level throughout our analysis to allow for arbitrary dependence of μ^z_{jt} over time within tracts.

To avoid having our results estimating Eq. (B.1) by ordinary least squares be biased by particularly large Census tracts, we index the number of annual non-SFHA and SFHA PIF for each Census-tract to the value in 2009. The results are reported in Table B.1. Overall, we find that most of the relationships that we estimate using our estimated non-SFHA and SFHA take-up rates hold qualitatively when analyzing normalized PIF, though several correlations are not as strong as in our primary results. We find positive relationships between tract education, income, and age levels, though in the case of income and age these relationships are estimated with less precision across specifications. We also estimate mixed results on the share of tract households with a mortgage and the estimated coefficients on the FSF-NFM variables are positive but not statistically significant. We also estimate a negative coefficient on the lagged IA grant count in the case of non-SFHA PIF; however, as shown in Figure 4 it is possible that this masks important heterogeneity over time.

Table B.1: Results regressing the Census tract-level outside SFHA normalized policies-in-force (columns 1-2) and the Census tract-level within SFHA normalized policies-in-force (columns 3-4) on a set of covariates of interest.

	nSFHA Take-up		SFHA Take-up	
	(1)	(2)	(3)	(4)
Homeowner Attributes				
Pct. of Pop. with College Degree	0.512**	0.123	1.009**	0.896
	(0.217)	(0.378)	(0.456)	(0.662)
Pct. of Pop. under 35	1.865***	-1.215**	1.166	-2.504***
	(0.352)	(0.525)	(0.844)	(0.724)
Unemployment Rate	0.276	0.093	-1.350	-0.981
	(0.323)	(0.375)	(0.990)	(0.812)
Minority Pct. of Pop.	0.243	-0.019	-0.387	0.260
	(0.152)	(0.223)	(0.294)	(0.393)
Log(Total Population)	0.004	1.473***	0.348	1.475**
	(0.046)	(0.327)	(0.260)	(0.701)
Household Attributes				
Log(Median HH Income)	0.239**	0.095	-0.302**	0.104
	(0.104)	(0.114)	(0.132)	(0.210)
Log(Median Home Value)	0.092	0.115	-0.447*	0.111
	(0.073)	(0.208)	(0.256)	(0.163)
Median Home Construction Age	-0.011***	-0.011	-0.005**	-0.006
	(0.002)	(0.008)	(0.003)	(0.004)
Pct. of HH with a Mortgage	0.339	-0.225	1.453**	-0.191
	(0.208)	(0.287)	(0.645)	(0.263)
Geography Attributes				
Number of High Precipitation Days	-0.542***	-0.607***	-0.505**	-0.334**
	(0.115)	(0.081)	(0.238)	(0.134)
DD: All Rel. Cumul. 5-year Lag	0.050***	0.036***	0.088***	0.075***
	(0.010)	(0.010)	(0.027)	(0.025)
Total IA Count Cumul. 5-year Lag	-0.029	-0.061**	-0.055*	-0.112**
	(0.022)	(0.025)	(0.032)	(0.048)
Coastal Tract	-0.110		-0.732*	
	(0.107)		(0.384)	
Total Tract Area Share: Water	0.680*		-0.052	
	(0.356)		(0.403)	
Soil Permeability	0.090***		-0.007	
	(0.017)		(0.026)	
nSFHA Pr(FF>1)	0.290*			
	(0.174)			
SFHA Pr(FF>1)			0.292	
			(0.242)	
NFIP Policy Attributes				
Avg. CRS Discount nSFHA	3.162	-3.791*		
	(2.458)	(2.031)		
Avg. Policy Cost nSFHA	-0.780***	-2.227***		
	(0.096)	(0.262)		
Avg. CRS Discount SFHA			-1.733***	-3.162***
			(0.632)	(1.159)
Avg. Policy Cost SFHA			-0.155***	-0.220***
			(0.038)	(0.048)
County FEs	✓		✓	
Tract FEs		✓		✓
County × Year Trends	✓	✓	✓	✓
State × Year FEs	✓	✓	✓	✓
Observations	561,504	568,152	342,585	371,531

Note: p values are calculated using a two-sided t-test and heteroskedasticity-robust standard errors clustered at the Census tract-level; $*p < 0.1$, $**p < 0.05$, $***p < 0.01$. Heteroskedasticity-robust standard errors clustered at the Census tract-level are reported in parentheses.

Appendix C. Effect of Private Risk Information over Time

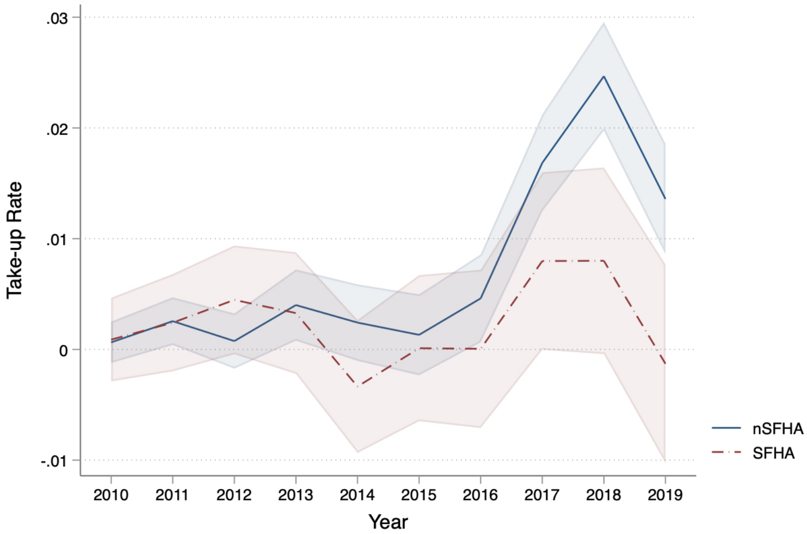
As discussed in Section 3.2, there is an important limitation in the measure that we use to proxy private inundation risk information. Specifically, there are important timing considerations that may affect our interpretation of this measure: given that FSF-NFM-derived Flood Factors are made available to the broader public starting in 2020 and that the model uses hydrological and climate observations through the end of our sample period, there certainly is an element of these measures which cannot be in the information set of households during the early part of our sample period. We still view this as a strong proxy for inundation risk not measured by FEMA maps—and therefore unpriced in NFIP rates—since if it correlates strongly with this residual inundation risk in the later years in our sample, it should also correlate to residual risk in the earlier years of our sample, albeit a weaker correlation. To further explore the relationship between our measure of private risk information and take-up over time, we estimate a form of our test for asymmetrically used information in which we allow for heterogeneity in the correlation between take-up and residual inundation risk over time.

Specifically, following the setup of the unused observables test described in Section 4.4, we partition our vector of observable characteristics, $\boldsymbol{X}^{z\prime}_{jt} = \begin{bmatrix} \boldsymbol{D}^{z\prime}_{jt} & \boldsymbol{W}^{z\prime}_{jt} \end{bmatrix}$, letting $\boldsymbol{D}^{z}_{jt}$ be a set of attributes used by the NFIP to classify buyer risk types and letting $p^z_{jt} = f(\boldsymbol{D}^z_{jt})$, where $f(\cdot)$ is the NFIP's pricing function. Further, we define $\boldsymbol{W}^z_{jt}$ to be a set of attributes which are not used by the NFIP in setting premiums. In practice we use our FSF-NFM derived measure of inundation risk as the sole element of $\boldsymbol{W}^z_{jt}$. To examine the evolution of the effect of this candidate unused observable on demand over time in our sample, we implement a version of Eq. (4) of the unused observables test using the following estimating equation:

$$y^z_{jt} = \alpha^z p^z_{jt} + \boldsymbol{D}^{z\prime}_{jt}\boldsymbol{\eta}^z + \sum_{\tau=2009}^{\tau=2019} \left(1[t=\tau] \times \boldsymbol{W}^{z\prime}_{jt}\boldsymbol{\gamma}^z_t \right) + \theta^z_j + \delta^z_{st} + (\lambda^z_c \times t) + \epsilon^z_{jt} \tag{C.1}$$

Figure C.1 plots the estimated values of $\boldsymbol{\gamma}^z_t$. Overall, we do see heterogeneity in the effect of the probability that a household within a Census tract is at risk of experiencing any flooding according to FSF-NFM on insurance demand over time in our sample. We observe an increase in the effect of this FSF-NFM derived inundation risk measure over our sample period, which we view as affirming our interpretation of our primary results from the unused observables test reported in Table 4 that this indeed represents asymmetrically used information. Given the fact that that FSF-NFM derived Flood Factors are made available to the broader public starting in 2020 and that the model uses hydrological and climate observations through the end of our sample period, if it does proxy private information about inundation risk, the FSF-NFM derived inundation risk measure's effect should be larger in the later years of our sample as these are the observations for which this information most reflects the state of knowledge about flood risk. This is the result that we find in Figure C.1.

Figure C.1: Effect of FSF-NFM derived probability of a household experiencing flooding on insurance demand over time.



Note: The relationship between FSF-NFM derived probability of a household experiencing flooding on insurance demand over time is reported relative to the omitted base year of 2009. Figure C.1 is estimated by regressing take-up rates on the probability that a household is at risk of experiencing a flood according to FSF-NFM interacted with year indicators; the full set of time-varying and time-invariant covariates included in Table 2; and a set of county fixed effects, state-by-year fixed effects, and county-by-year time trends. 95% confidence intervals are constructed using heteroskedasticity-robust standard errors clustered at the Census tract-level.

Appendix D. Spatial Heterogeneity of Asymmetrically Used Information

We explore potential heterogeneity in the presence of asymmetrically used information over space by implementing the unused observables test described in Section 4.4 allowing for different relationships between our measure of residual flood risk and take-up and ex-post claim probabilities across different geographies. Spatial heterogeneity in the degree of asymmetrically used information might arise for several reasons. First, the underlying population of households may be better able to access private information about their flood risk; as Figure 6 indicates, the effect of the probability of a household having a First Street Foundation Flood Factor greater than one on claim probabilities and take-up rates increases with income and education. It is also likely that the degree to which FEMA flood maps and the NFIP rate schedule misprice flood risk varies across geographies. While we are not able to distinguish between these mechanisms, we still believe that exploring potential heterogeneity in the degree of asymmetrically used information over space is a worthwhile exercise: the results of such an exercise provides policymakers with valuable insight into where the pool of insured households under the NFIP may be more adversely selected.

Specifically, following the setup of the unused observables test described in Section 4.4, we partition our vector of observable characteristics, $\boldsymbol{X}^{z\prime}_{jt} = \begin{bmatrix} \boldsymbol{D}^{z\prime}_{jt} & \boldsymbol{W}^{z\prime}_{jt} \end{bmatrix}$, letting $\boldsymbol{D}^{z}_{jt}$ be a set of attributes used by the NFIP to classify buyer risk types and letting $p^z_{jt} = f(\boldsymbol{D}^z_{jt})$, where $f(\cdot)$ is the NFIP's pricing function. Further, we define $\boldsymbol{W}^z_{jt}$ to be a set of attributes which are not used by the NFIP in setting premiums. In practice we use our FSF-NFM derived measure of inundation risk as the sole element of $\boldsymbol{W}^z_{jt}$. To examine the evolution of the effect of this candidate unused observable on demand over space in our sample, we implement a version of Eqs. (4) and (5) of the unused observables test using the following estimating equations:

$$y^z_{jt} = \alpha^z p^z_{jt} + \boldsymbol{D}^{z\prime}_{jt}\boldsymbol{\eta}^z + \sum_{\kappa\in\mathcal{G}}\left(1[g_j = \kappa] \times \boldsymbol{W}^{z\prime}_{jt}\boldsymbol{\gamma}^z_g\right) + \theta^z_j + \delta^z_{st} + (\lambda^z_c \times t) + \epsilon^z_{jt} \tag{D.1}$$

$$c^z_{jt} = \sigma^z p^z_{jt} + \boldsymbol{D}^{z\prime}_{jt}\boldsymbol{\pi}^z + \sum_{\kappa\in\mathcal{G}}\left(1[g_j = \kappa] \times \boldsymbol{W}^{z\prime}_{jt}\boldsymbol{\xi}^z_g\right) + \theta^z_j + \delta^z_{st} + (\lambda^z_c \times t) + \varepsilon^z_{jt} \tag{D.2}$$

where g_j is a variable encoding Census tract j's parent geography among the set of possible aggregate geographies, $\mathcal{G}$. We estimate Eqs. (D.1) and (D.2) for two different sets of aggregate geographies (for each of which a Census tract is a complete subset): the 10 FEMA Regions (each of which is a collection of states) and the states within the contiguous U.S. plus the District of Columbia.

Figure D.1a plots the estimated values of $\boldsymbol{\xi}^z_g$ and Figure D.1b plots the estimated values of $\boldsymbol{\gamma}^z_g$ where we allow for FEMA Region-specific coefficients on our measure of residual risk. The FEMA Regions are defined as follows for the contiguous U.S.:

- Region 1: Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont
- Region 2: New Jersey and New York

- Region 3: Delaware, District of Columbia, Maryland, Pennsylvania, Virginia, and West Virginia
- Region 4: Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, Tennessee
- Region 5: Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin
- Region 6: Arkansas, Louisiana, New Mexico, Oklahoma, and Texas
- Region 7: Iowa, Kansas, Missouri, Nebraska
- Region 8: Colorado, Montana, North Dakota, South Dakota, Utah, Wyoming
- Region 9: Arizona, California, and Nevada
- Region 10: Idaho, Oregon, and Washington

A Region having non-zero estimated relationships between the probability of a household having a First Street Foundation Flood Factor greater than 1 on both claim probabilities and take-up rates rejects the null hypothesis of symmetrically used information in that Region. Moreover, for reasons discussed in Section 5.4, we argue that positive relationships between this measure of risk and both claim probabilities and take-up rates suggests an adversely selected pool of insured households.

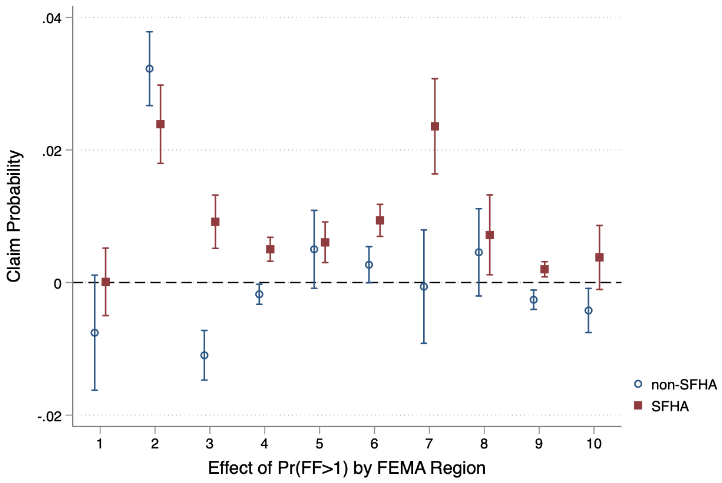
As shown in Figure D.1, we reject the null hypothesis of symmetrically used information among both non-SFHA and SFHA households in FEMA Regions 2, 3, 4, 6, 9, and 10. For all but Region 1, we reject the null hypothesis of symmetrically used information among SFHA households, with positive estimated correlations between our risk measure and both claim probabilities and take-up rates. In Regions 2 and 6, we estimate positive relationships between our risk measure and claim probabilities and take-up rates for both non-SFHA and SFHA households, with Region 2 having the largest estimated relationships between the risk measure and claim probabilities.

Figure D.2 further explores spatial heterogeneity in the degree of asymmetrically used information by mapping the estimated values of ξ_g^z and γ_g^z where we allow for state-specific coefficients on our measure of residual risk. The results in Figure D.2 are consistent with those in Figure D.1, particularly the suggestive result of a high degree of asymmetrically used information in FEMA region 2, with large, positive coefficients estimated for non-SFHA and SFHA households in both New York and New Jersey. Figure D.2 also suggests the presence of asymmetrically used information in certain additional coastal states, including Texas and North Carolina. Nebraska and several other states in the Midwest and Great Lakes regions also appear to be characterized by some degree of asymmetrically used information.

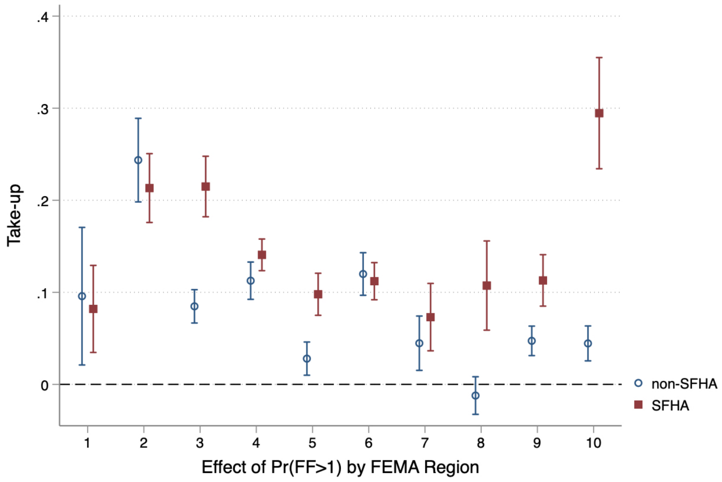
While these FEMA Region- and state-specific results are informative, there are no clear patterns that suggest potential mechanisms driving the observed heterogeneity over space. This is perhaps a result of the high-level of aggregate geography over which we allow for heterogeneity in the degree of asymmetrically used information. Examination of the mechanisms through which the observed spatial heterogeneity operates is an interesting avenue for future research.

Figure D.1: Effect of the probability of a household having a First Street Foundation Flood Factor greater than 1 on (a) claim probabilities and (b) take-up rates by FEMA region (only includes the contiguous U.S.).

(a) Effect of unmapped risk on claim probabilities by FEMA region.



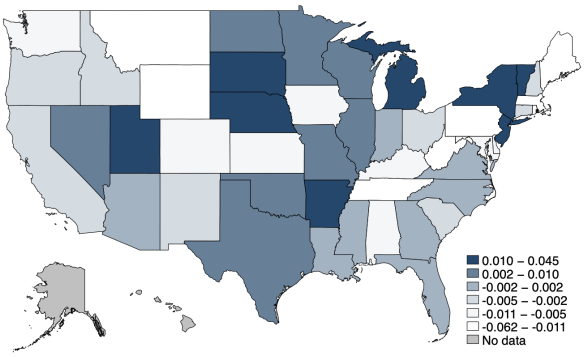
(b) Effect of unmapped risk on take-up rates by FEMA region.



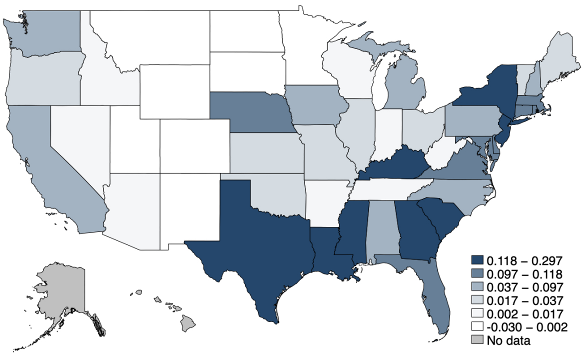
Note: Relationships are estimated by regressing non-SFHA and SFHA claim probabilities and take-up rates on the full set of interactions between binary variables indicating whether a Census tract falls within a given FEMA region and the probability that a household has a First Street Foundation Flood Factor greater than 1. Regressions also include the full set of time-varying and time-invariant controls included in Table 2 and a set of state-by-year fixed effects, county fixed effects, and county-by-year time trends for both non-SFHA and SFHA claim probabilities and take-up. 95% confidence intervals are constructed using heteroskedasticity-robust standard errors clustered at the Census tract-level.

Figure D.2: Effect of the probability of a household having a First Street Foundation Flood Factor greater than 1 on (a) non-SFHA claim probabilities, (b) non-SFHA take-up rates, (c) SFHA claim probabilities, and (d) SFHA take-up rates by state.

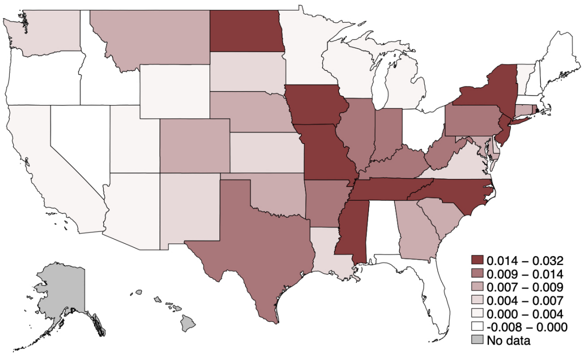
(a) Effect of unmapped risk on non-SFHA claim probabilities by state.



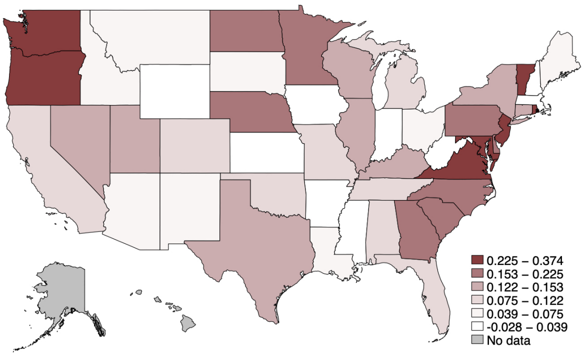
(b) Effect of unmapped risk on non-SFHA take-up rates by state.



(c) Effect of unmapped risk on SFHA claim probabilities by state.



(d) Effect of unmapped risk on SFHA take-up rates by state.



Note: Relationships are estimated by regressing non-SFHA and SFHA claim probabilities and take-up rates on the full set of interactions between binary variables encoding a Census tract's state and the probability that a household has a First Street Foundation Flood Factor greater than 1. Regression also include the full set of time-varying and time-invariant controls included in Table 2 and a set of state-by-year fixed effects, county fixed effects, and county-by-year time trends for both non-SFHA and SFHA claim probabilities and take-up. 95% confidence intervals are constructed using heteroskedasticity-robust standard errors clustered at the Census tract-level.