Adapting to Natural Disasters through Better Information: Evidence from the Home Seller Disclosure Requirement

Seunghoon Lee*

2024-02-21

Abstract

While flood damage is determined by both flood intensity and population exposure, the US has predominantly focused on managing the former, with limited success. This paper studies whether a Home Seller Disclosure Requirement can reduce flood exposure and thus flood damage. Leveraging two quasi-experimental variations of the policy, I first show that mandating flood risk disclosure lowers the population living in high-risk areas. Further, using a hydrological measure of flood intensity, I find that it reduces the probability of flood damage by 31 percent. These findings illustrate that easing information frictions can promote voluntary adaptation to natural disasters.

Click Here for the Latest Version.

*Department of Economics, University of Missouri (seunghoon.lee@missouri.edu). I am deeply grateful to Ryan Kellogg, Koichiro Ito, and Dan Black for their guidance, mentorship, and patience, and for facilitating access to data. I am also grateful to Eyal Frank and Amir Jina for their comments and encouragement especially during the early stages of this project. I also benefited tremendeously from numerous discussions with Cory Koedel and Peter Mueser. I thank H. Spencer Banzhaf, Lint Barrage, Zarek Brot-Goldberg, Fiona Burlig, Justin Gallagher, Jonathan Gourley, Michael Greenstone, Solomon Hsiang, Seojeong Lee, George Lefcoe, Christian Leuz, Meagan McCollum, Philip Mulder, Anant Sudarshan, Richard L. Sweeney, Katherine Wagner, Margaret Walls, Austin Wright, Derek Wu, Siqi Zheng, and seminar participants at the AERE Summer Conference, AREUEA National Conference, FHFA Fall Economic Summit, KDI School of Public Policy and Management, Midwest Energyfest, MIT, NBER Summer Institute, OSWEET (Online Summer Workshop in Environment, Energy, and Transportation), University of Chicago, and University of Missouri for their helpful comments. Housing data are provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author and do not reflect the position of Zillow Group.

1 Introduction

Since 1980, floods in the United States have wrought over $1 trillion in damage, making them the costliest type of natural disaster over the last 40 years (NOAA 2020). Climate scientists predict flooding is likely to happen with higher frequency and intensity in the future (Milly et al. 2002, Ghanbari et al. 2019). Thus, effective adaptation, which is an activity to moderate or avoid harm, is increasingly important (IPCC 2014, Aldy and Zeckhauser 2020).

While flood damage is determined by both flood intensity (i.e., physical characteristics) and exposure (i.e., population size in high-risk areas), flood policy in the US has focused primarily on managing the former by adding engineering structures (Changnon et al. 2000, Field et al. 2012, Tarlock 2012, Liao 2014). Unfortunately, such structural approaches, which are not failproof, can inadvertently exacerbate the problem by attracting more people and developments to floodplains (the so-called "levee effect") (Pinter 2005, Kousky et al. 2006, Collenteur et al. 2015).[1] Consequently, governments end up spending billions of dollars for disaster relief and recovery on top of the resources devoted to flood prevention (CBO 2016). Given these limitations, disclosure policies that aim to reduce flood exposure by affecting location choices are getting more attention, but little is known about such a policy's effect and mechanisms.[2]

This paper exploits quasi-experimental variations from a Home Seller Disclosure Requirement (hereafter "the disclosure requirement") to study whether easing information frictions about flood risk can reduce (1) the number of households in high-risk areas and (2) resulting flood damage. Although official flood maps have long been publicly available, earlier research and anecdotal evidence suggest a lack of flood risk awareness among homebuyers. For instance, Chivers and Flores (2002) find only 14 percent of homebuyers in high-risk areas learned about flood risk before closing. Such low awareness makes it unlikely that homebuyers fully internalize the costs of flood risk during real estate transactions. Given that a potential reason for the friction is information acquisition and processing costs (Kunreuther and Pauly 2004), the disclosure requirement could alleviate the problem by efficiently delivering risk information.

[1]Flood prevention structures frequently fail as evidenced by the 1993 Midwest Flood where over 1,000 levees failed (LARSON 1996). A major contributing factore to these failures is the lack of maintenance, with only 1.9% of US levees rated as "Acceptable" (Pinter et al. 2016)

[2]For instance, FEMA has recently proposed an NFIP reform tying a community's flood insurance eligibility to mandatory flood risk disclosure before real estate transactions (U.S. Department of Homeland Security 2022).

The policy mandates that home sellers must disclose known property issues on a wide range of dimensions including land, structure, and ambient environments using a standardized form (Lefcoe 2004). Regarding flood risk, a typical question is if a property is located in a Special Flood Hazard Area (SFHA)—an area with elevated risk defined by the official flood map. Home sellers are generally obliged to fulfill the disclosure requirement before closing (Stern 2005).

The disclosure requirement was rolled out across 26 states in the contiguous US from 1992–2003. The variation in implementation timing is from plausibly exogenous state court rulings on the extent of realtor liability for incomplete disclosure (Roberts 2006), which facilitates a difference-in-differences research design. In addition, the disclosure requirement treats properties located in and out of the SFHA differently, which introduces a third difference to further aid in identification. In exploiting the staggered adoption of the disclosure requirement, I use the stacked approach to overcome potential bias from conventional two-way fixed effect models (Cengiz et al. 2019, Brot-Goldberg et al. 2020, Goodman-Bacon 2021).

I also leverage additional variation stemming from the spatial discontinuity in flood risk information at the flood zone border. That is, homebuyers for two proximate properties located on opposite sides of an SFHA border—over which flood risk is changing continuously—receive starkly different flood risk information, which yields an opportunity to identify the information effect holding true flood risk constant. A potential concern is that being located in the SFHA could invite other treatments such as the mandatory purchase of flood insurance. Thus, I use the difference-in-discontinuity approach to control for time-invariant confounders (Grembi et al. 2016).

I collect multiple datasets to leverage these variation in empirical models designed to recover the causal effect of increased information about flood risk during real estate transactions. I use census-block-level demographic data from the decennial census, and community-level flood insurance policy counts from the National Flood Insurance Program (hereafter “flood insurance” or “NFIP”). To measure flood damage, I use damage records from flood insurance adjuster reports. I also construct historical dataset of community-level flood events based on a hydrological measure of flood intensity (Saharia et al. 2017, England Jr et al. 2019). These data overcome the potential endogeneity of self-reported flood events, such as those from the National Weather Service Storm Events data (Gall et al. 2009). Because the main outcome variables used in the analysis have a mass point at zero with a long right tail, I estimate the extensive and intensive margin effects separately following suggestions of Chen and Roth (2022).

I analyze the data in two parts that correspond to the two research questions posed earlier. First, I investigate homebuyer responses to the disclosure policy. Specifically, building on the insights from Ehrlich and Becker (1972)—households mitigate hazard risk by choosing between self protection and market insurance—I empirically estimate the impact of the disclosure requirement on the geographic distribution of the population and flood insurance take up rates. Second, I use a flood damage function to test if the disclosure policy reduces flood damage.

From the first part, I find that census blocks in an SFHA area (conditional on having a non-zero population) experiences a 7 percent decline in population after the disclosure policy. At the extensive margin, disclosure lowers the probability of a block in the SFHA having any population by 0.01 percentage points, or 1.5 percent from the baseline value of 0.67. I further show that these effects are driven by diverted in-migration (and resulting suppressed development) rather than active out-migration from SFHA areas. In contrast, I find a very small effect of the disclosure policy on flood insurance purchases: the probability of having a positive number of insurance policies in a community increases by 0.003 percentage points (or 0.4 percent) on the extensive margin, while insurance counts per housing unit decreases by 2 percent on the intensive margin. Investigating these two response channels is important because they have different implications for flood damage. That is, while choosing a safer location to live would reduce the probability of flooding, buying flood insurance would simply redistribute income from the "dry state" to the "flooding state" without necessarily affecting the probability distribution (Ehrlich and Becker 1972).

From the second part, I find that the disclosure policy reduces the expected probability of having any flood damage at the community level by 2.3 percentage points (or 31 percent of the baseline mean). To show this, I first estimate a non-parametric flood damage function—a mapping between flood size and damage—using community-level flood history and damage data. Then, I estimate the causal effect of the disclosure policy on the damage function and find that the slope of the function is substantially flatter after the policy. This analysis further reveals that the disclosure effect is disproportionately larger in communities with higher treatment intensities.

This paper contributes to four different bodies of literature. First, it is related to prior studies on factors that mitigate damage from climate change. While earlier studies primarily focus on technology as a driver of adaptation (Miao and Popp 2014, Barreca et al. 2016, Burke and Emerick 2016), I focus on the role information can play in aligning private incentives with socially desirable outcomes. A recent paper by Fairweather et al. (2023), which experimentally show that Redfin users are less likely to purchase a home with flood risk upon receiving flood risk information, is an important exception. I complement Fairweather et al. (2023) with (1) a higher external validity and (2) the ability to estimate changes in the flood damage from information provision.

Second, I contribute to the literature on the role of government in shaping household adaptation behaviors (Kousky et al. 2006, 2018, Gregory 2017, Peralta and Scott 2020, Baylis and Boomhower 2022). Perhaps the closest papers conceptually are Baylis and Boomhower (2021) and Ostriker and Russo (2023), which show how building-code policies can reduce wildfire damage or flood risk exposure, respectively. A key difference is that the policies studied by these papers directly mandate adaptation, whereas disclosure policies studied in this paper encourage voluntary adaptation such as choosing safer places to live.

Third, and more broadly, I build on earlier work on the impacts of flood risk on the housing market (Hallstrom and Smith 2005, Pope 2008, Bin and Landry 2013, Bosker et al. 2019, Muller and Hopkins 2019, Gibson and Mullins 2020, Hino and Burke 2021, Bakkensen and Barrage 2021). While most prior studies focus on understanding how changes to flood risk information or beliefs affect housing prices, I study their impacts on flood damage.[3] Tracing the effect of flood information up to the damage amount is important because while housing price changes, in general, reflect transfers between homebuyers and sellers, a reduction in flood damage enhances social welfare.

Finally, I contribute methodologically by constructing a novel measure of flood exposure, which is a critical step in identifying climate change effects (Hsiang 2016). My approach leverages hydrological measures of flood intensity, allowing me to document flood events objectively for a wide range of causes including rainfall, snow melt, or storm surge. My measure complements existing ones that specialize in capturing the impacts of rainfall or hurricanes (Strobl 2011, Davenport et al. 2021).

The paper proceeds as follows. Section 2 provides background on the Home Seller Disclosure Requirement and the Special Flood Hazard Area. Section 3 details the data sources and provides some summary statistics. Section 4 presents estimation results on household responses while Section 5 shows the disclosure policy effect on flood damage. Section 6 concludes.

[3]For example, Hino and Burke (2021) use flood map updates as the main source of information shock and test if the housing market efficiently prices flood risk. While they focus on the price effect, I study household responses to the information shock—which provide an explanation for price adjustments—and resulting change in flood damage.

2 Background

2.1 Home Seller Disclosure Requirement

Publicly available Flood Insurance Rate Maps contain the information homebuyers need to determine whether a property is located in an SFHA. Furthermore, the Flood Insurance Reform Act of 1994 requires the purchase of flood insurance as a condition for federally-backed mortgage approval for properties in SFHAs, which should inform homebuyers of the associated flood risk. However, compliance with the flood insurance mandate is far from perfect (Tobin and Calfee 2005, Michel-Kerjan 2010, GAO 2021, Wagner 2022) and prior work shows homebuyers do not have a good understanding of the flood risk they face (Chivers and Flores 2002, Pope 2008, Bin and Landry 2013). A disclosure on flood risk could be a useful apparatus to address this information gap.

Disclosure content. A statutory disclosure requirement mandates that home sellers provide buyers with a detailed account of known material defects in the listed property by filling out a standardized form. Importantly, the disclosure requirement is not exclusively about flood risk. As Appendix Figure D.1 illustrates, a typical form covers a wide range of property conditions including structural issues (e.g., problems with walls) and surroundings (e.g., natural hazards such as flood risk).[4]

The exact language of disclosure on flood risk varies slightly from state to state, but some combination of the following three questions usually appears: whether a property is in the SFHA; whether a property had flood damage history; and whether a property has flood insurance.[5] Because properties on the SFHA are more susceptible to flooding, these questions are highly correlated. Indeed, flood insurance policy and claims data that I acquired through FOIA show that 71 percent (75 percent) of the claims (flood insurance policies) are from properties in the SFHA. Therefore, irrespective of the language, disclosure is likely to raise homebuyers' flood risk awareness for properties in SFHAs relative to those outside.

Disclosure background and determinants of policy adoption. Traditionally, homebuyers were ex-

[4]Since the disclosure delivers a bundle of information, discerning treatment mechanism can be challenging especially when there is positive correlation between flood risk and other property defects. In Appendix Table D.1, I demonstrate that properties in tracts with SFHAs are notably newer compared to those in tracts without SFHAs. As property defects typically emerge over time, this table suggests that SFHA properties are less prone to issues, implying that the disclosure policy's impact on flood-related outcomes stems from flood risk information.

[5]As of 2021, 5 states ask just the first question about the SFHA status, 15 states ask about SFHA status and past flood experience, and 4 states ask all three questions. MI and TN ask about the latter two only.

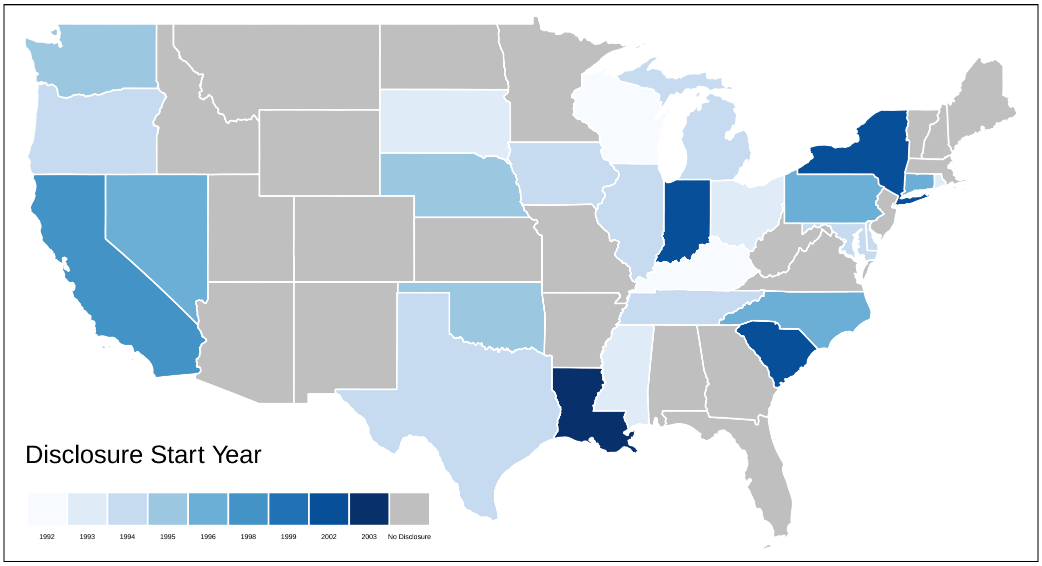


Figure 2.1: The Disclosure Requirement Implementation over Time

pected to practice caution regarding property defects ("*caveat emptor*" or "let the buyer beware" doctrine). However, due to increasing consumer protectionism and public awareness of environmental and health concerns, state courts began holding listing agents accountable for incomplete disclosures (Weinberger 1996, Lefcoe 2004). In response, the National Association of Realtors issued a resolution in 1991 urging state associations to develop and support legislation regarding the statutory disclosure requirement in an effort to deflect potential liability to sellers (Tyszka 1995, Washburn 1995).

Consequently, between 1992 and 2003, 26 states in the contiguous US (excluding DC) adopted a disclosure requirement with an explicit question on flood risk while the remaining 22 states never adopted such a requirement (Figure 2.1). In Appendix B, I show that (1) the 22 never-adopted states are different on demographic, economic, and political characteristics from the 26 ever-adopted states and (2) such a difference does not appear in the early vs. late adopting states comparison. Given these observations, I leverages differences in implementation timing, which reflects the timing of the change in the state court's ruling (Roberts 2006), when employing panel regression models.

It is also worth pointing out that five of the 22 non-disclosure states adopted a variant of a home seller disclosure mandate, although it does not have a question on flood risk.[6] These "placebo" states are useful for checking the robustness of the main results.

[6]For details regarding the extent of disclosure in these states, see the following. Idaho: 1994 Ida. 55-2508 (1994), Maine: Title 33 Section 173 (1999), Minnesota: CHAPTER 306-—S.F.No. 2697 (2003), New Hampshire: NH. Rev. Stat. Ann. § 477:4-c (1994), and Virginia: VA. CODE ANN. §§ 55.1-704 (2005).

Does the disclosure requirement matter? Despite potential penalties for imperfect compliance, the disclosure requirement's effectiveness in raising homebuyers' flood risk awareness remains uncertain due to possible inaccuracies in information provided by sellers or buyers' failure to process the new information. While I cannot observe homebuyers' perceptions of flood risk directly, in Appendix C, I find disclosure reduces housing prices in SHFA's by 4.5 percent. The magnitude coincides with existing estimates of the effect of flood risk information on housing prices (e.g., see Hino and Burke (2021)), and it suggests that at least some information is being conveyed to buyers through the disclosure requirement. The extent to which this information translates to risk-mitigating behaviors is uncertain and the focus on my empirical work below.

2.2 Flood Map and Special Flood Hazard Area (SFHA)

The SFHA, designated by an official flood map for potential inundation by a 100-year flood, holds significant importance as it frequently serves as a reference point for flood risk information in disclosure requirements (FEMA 2011). The SFHA boundary is determined by comparing water surface elevation with the ground elevation under a 100-year flood scenario (FEMA 2005), which implies that flood risk, which is a function of land contour, is continuously changing even at the SFHA border (Noonan et al. 2022). This gives rise to the spatial discontinuity design at the SFHA border because the disclosure form treats flood risk discontinuously for two areas on different sides of the border with possibly very similar true flood risk.

It is also worth noting that these maps are updated occasionally, albeit much less frequently than legally mandated (DHS Office of Inspector General 2017), which could potentially confound the disclosure effect. Throughout empirical exercises, I show that my results are robust to the map updates.

The jurisdiction of each flood map is a NFIP "community". These communities are local political entities comparable to a US census place. Appendix Figure D.2 shows a sample flood map from a part of the Borough of Stonington, CT, and similar to this place, a typical entity contains both SFHA and non-SFHA areas within it. Indeed, Appendix Figure D.3 illustrates that there is substantial variation in the fraction of area covered by an SFHA for communities in the sample, which suggests that the intensity of disclosure treatment varies across communities.

3 Data Sources and Descriptive Statistics

Demography and flood insurance. I collect census-block-level population and occupancy data from the 1990, 2000, 2010, and 2020 decennial censuses. To account for changing block boundaries and resulting one-to-many matches across different decennial census years,[7] I calculate the weighted sum of count variables using interpolation weights from the NHGIS block-to-block crosswalk (Manson et al. 2022).[8]. This creates a geographically standardized time series. Data on the number of flood insurance policies at the NFIP community level come from FEMA for 1978–2008.[9]

Flood damage. I use damage records from the flood insurance adjuster's report. The damage amount is defined by the actual cash value—a replacement value net of depreciation (FEMA 2014). I observe individual property level damage with loss date, community ID, and building type. I restrict the sample to damage records from single-family houses that has sustained the largest flood event for a given community-year. Then I collapse the data to the community by year level to match it with the annual maximum flood event data.

Flood history. The measurement of climate exposure is a critical methodological step in identifying climate effects on economic outcomes (Hsiang 2016). In the domain of floods, two different measures have been widely used. The first approach measures flood intensity using outcome variables such as economic cost, which suffers an endogeneity problem by construction (for a review, see Felbermayr and Gröschl (2014)). The second approach uses a meteorological measure, but only for a subset of events such as a hurricane or rainfall (Strobl 2011, Hsiang and Jina 2014, Deryugina 2017, Davenport et al. 2021). Given that rainfall alone explains just one-third of cumulative flood damage in the US (Davenport et al. 2021), such an approach cannot capture the entire scope of floods.

To overcome these limitations, I construct hydrology-based flood history dataset at the community level using daily water volume records from over 3,000 USGS and NOAA stations located within my 26-state sample (Milly et al. 2002, Mallakpour and Villarini 2015, Slater and Villarini 2016). Using this approach, flood size is described by the recurrence interval (Task Committee on Hydrology

[7]For instance, block G06000104003003006 in 2000 is matched to five different blocks in 2010 ending in 3010, 3011, 3017, 3020, and 3028.

[8]Interpolation weights represent the expected proportion of the source block's counts (e.g., population or housing units) located in each target block (Manson et al. 2022)

[9]I thank Justin Gallagher for graciously sharing these data.

Handbook of Management Group D of ASCE 1996): the expected number of years for a flood of the same magnitude to come back for a given site. For instance, a 10-year flood is the size of a flood that would happen on average once every 10 years, which would be less severe than a 100-year flood that is large enough to happen only once in 100 years on average. An idea behind this approach is similar to model extreme temperatures as deviations from local mean temperatures.

Practically speaking, I construct the data in four steps. First, I estimate a gauge-specific flood frequency distribution by fitting the Log-Pearson III distribution using the annual peak flow records of each gauge. Second, I convert the daily maximum discharge volume at each gauge into quantiles of the fitted distribution from step 1. Third, I translate the quantiles into recurrence intervals and take the maximum recurrence interval for each year. [10] Finally, I match each community to the three nearest gauges and calculate community-year-level flood size by taking the inverse-distance weighted average of three closest gauges' recurrence intervals. More details on the flood data construction procedure and summary statistics are in the Appendix A.1.

Other data sources. To determine the SFHA status of geographic units such as census block, I use the Q3 map—the first generation of a digitized flood map—that captures flood risk as of the mid-1990s for over 1,300 counties (FEMA 1996). Also, the primary data source to track the disclosure requirement legislative history is the *Nexisuni* database. I cross-validate this database with prior studies on the disclosure requirement (Washburn 1995, Pancak et al. 1996, Lefcoe 2004) and reports from the National Association of Realtors (National Association of Realtors 2019).

Summary statistics. Table 3.1 shows summary statistics for key independent variable: flood size; and dependent variables: population, flood insurance policy counts per housing unit, and flood damage per housing unit. Population figures are for the census blocks within the optimal bandwidth (more detail in Section 4). The last three values are for the NFIP communities in my sample.

A notable aspect of the data is the high prevalence of zeros among the dependent variables. For instance, 27 percent of observations for the block population and counts and 17 percent of the observations of flood insurance policy counts are zeros. For the community-level flood damage per housing unit variable, 95 percent of observations are zero. The high prevalence of zeroes for these variables is

[10] The recurrence interval for quantile q is $\frac{1}{1-q}$. For instance, a discharge volume of the 90th percentile, which means it is the 90th highest among 100 yearly maximum observations, corresponds to a 10-year flood.

Table 3.1: Summary Statistics for Key Variables

Variables	Min.	Q25	Median	Mean	Q75	Max.	N
Census Block Population	0	0	10	34.6	40	7,597	1,484,709
NFIP Policies Per Housing Unit	0	0.001	0.005	0.029	0.018	6.53	400,919
Flood Damage Per Housing Unit	0	0	0	5.8	0	23,991	505,383
N of 10-Year Floods (For 20 Years)	0	1	2	2.18	3	15	8,194

consistent with external sources (details are in Appendix A.2).

In addition, these variables also exhibit substantial skewness (long and thin right tails), as the difference between median and mean values suggests. To account for this, I follow Chen and Roth (2022) and estimate extensive and intensive margin effects separately for each dependent variable. This approach resonates with a hurdle or two-part model, which is used extensively in modeling health expenditures characterized by a similar distribution (Mullahy and Norton 2022).

Finally, the last row of Table 3.1 indicates that an average community experiences 2.18 10-year flood events over a 20-year period. This is close to the expected value of 2.0.

4 Responses to the Disclosure Requirement

In this section, I investigate how homebuyers respond to flood risk information by estimating the impact of the disclosure policy on geographic distribution of the population and flood insurance take up. Investigating both responses is important because they have starkly different implications for flood damage—responses along the first margin can reduce total flood risk whereas responses along the second margin simply redistribute income from the "dry state" to the "flooding state" without necessarily affecting exposure (Ehrlich and Becker 1972).[11]

4.1 Estimation Framework

Spatial Discontinuity. Yes-or-no check box questions on disclosure forms create a spatial discontinuity in flood risk information, which allows me to disentangle the information effect from the true flood risk effect. However, a potential concern is that other policies such as flood insurance requirements

[11] Ehrlich and Becker (1972) suggests that, when self-protection is financially rewarded, self-protection and market insurance are complements. However, as Kousky (2019) points out, the NFIP premium is heavily subsidized and the NFIP premium structure is too coarse to account for all self-protection measures. Wagner (2022) also finds that substitution between self-protection (property elevation) and flood insurance is prevalent in the flood insurance market.

also change at the border, which could confound the disclosure effect. To account for this, I leverage a difference-in-discontinuity approach as equation (1) following Grembi et al. (2016).

$$\begin{aligned} Y_{bst} =& \delta_0 + \delta_1 X_{bs} + \delta_2 D_{bs} + \delta_3 X_{bs} * D_{bs} + \\ & T_{st}[\delta_4 + \delta_5 X_{bs} + \delta_6 D_{bs} + \delta_7 X_{bs} * D_{bs}] + \epsilon_{bst} \end{aligned} \tag{1}$$

Y_{bst} is an outcome variable such as the probability of having any population, log of population conditional on having non-zero population, or the vacancy rate in block b in state s in time t.[12] X_{bs} is the distance between a block border and the closest SFHA border in meters (negative if in a non-SFHA area),[13] $D_{bs} = 1$ if $X_{bs} > 0$ is a treatment group indicator variable, and $T_{st} = 1$ if $t > T^*_s$ is a post-period indicator variable, where T^*_s is the policy change date for s. δ_6 captures the impact of the disclosure policy for blocks located in close proximity to the SFHA border. To estimate δ_6, I first estimate the optimal bandwidth for each outcome variable. Then, I estimate equation (1) using blocks within the optimal bandwidth (Calonico et al. 2014, Cattaneo et al. 2019).[14] For states that have implemented disclosure policies between 1990-1999 (2000-2009), I use the 1990, 2000, and 2010 (2000, 2010, and 2020) decennial census.

A potential concern of using a geographic area such as a census block (namely, a polygon) for a spatial discontinuity design is that the distance from a block to an SFHA border may not well defined, especially when a block contains an SFHA border within it. While this might be a serious problem for larger geographical units such as tracts, it will be less of a problem for blocks, which is the smallest census geographic unit (see Appendix Figure D.4): the median size of blocks in my sample is just 0.009 square miles, and 83 percent of blocks are entirely within or outside an SFHA. I remove the 17 percent of blocks that contain SFHA borders from the analysis.

Staggered Adoption. Flood insurance policy counts are observed at the community level. As a typical community contains both SFHA and non-SFHA areas (Appendix Figure D.2 and D.3), the distance to the nearest SFHA border is not defined. Thus, I employ a triple difference design using

[12] A property is considered vacant if no one is residing in the unit at the time of enumeration unless its occupants are only temporarily absent (US Census Bureau 2000).

[13] X_{bs} is approximated by taking the difference of (1) the distance between block centroids and the closest SFHA border and (2) a block diameter.

[14] I estimate the mean squared error optimal bandwidth for 2000 and 2010 respectively and take the average following Grembi et al. (2016). I ignore 1990 and 2020 because these years have only a subset of the states in the sample.

equation (2). Here, Y_{mstd} denotes outcome variables on NFIP for community m in state s in year t in stack d. H_{md} is an indicator variable equal to one if a community has an above-median fraction of the area covered by an SFHA, which proxies for the treatment intensity. I_{std} is a post disclosure indicator and D_{std} is a treatment group indicator. α_1 captures the disclosure effect.[15]

$$log(Y_{mstd}) = \alpha_0 H_{std} I_{std} + \alpha_1 H_{md} D_{std} I_{std} + \omega_{std} + \psi_{md} + \epsilon_{mstd} \tag{2}$$

As the subscript d indicates, I use the stacked approach to estimate the policy impact using clean controls (Cengiz et al. 2019, Brot-Goldberg et al. 2020). This approach alleviates concerns about potential biases in the staggered adoption design (Goodman-Bacon 2021). As I exploit the timing of the disclosure requirement for identification, not-yet-treated states form the control group.

To construct the data stack, I first keep each state's flood insurance data for seven years before and after the policy change to prevent composition changes.[16] Each stack consists of communities in the treated states, which adopted the disclosure policy in year t^*, and communities in the control states, which adopted the policy in $\tilde{t} > t^*$.[17] I drop observations from the control states for $t >= \tilde{t}$ because they are no longer "not-yet-treated". Equation (2) also include ω_{std}, the state $\times$ time $\times$ stack fixed effect to account for year-specific state level shocks and a community $\times$ stack fixed effect ψ_{md}, which captures unobserved community characteristics. Including these fixed effects ensures that the comparisons are made within each stack. Throughout the analysis in Section 4, standard errors are clustered at the state level, which corresponds to the level of disclosure treatment.

4.2 Findings

Self-protection. Table 4.1 reports the impact of the disclosure policy on the population over a period up to 20 years after the policy change. In column (1), I find that the disclosure reduces the probability of having any population in an SFHA block by 0.01 percentage points relative to a non-SFHA block (or 1.5 percent of the baseline value 0.68). In column (2), I limit the sample to blocks with non-zero population and find that the disclosure reduces population in a populated SFHA block by 7

[15] Other terms in a standard triple difference model is subsumed by the fixed effects.

[16] The data from 1978–2008 are sufficient to cover a 15-year window for policy changes in all states except Louisiana, which implemented its policy in 2003, leaving just six post-policy years for analysis.

[17] Stack refers to data that is created for a specific treatment year (or a cohort year). A state belongs to either the treatment or control group depending on the stack. For instance, PA and CT, which changed their policy in 1996 are in the "treatment group" in a stack for $t^* = 1996$. The two states belong to the "control group" when $t^* < 1996$.

Table 4.1: Effect of Discosure Requirement on Net Population Flow

	Prob. of Any Population	Log Population		Vacancy Rate
	(1)	(2)	(3)	(4)
SFHA × Post	$-.011^{***}$	$-.073^{**}$		$.014^{***}$
	(.003)	(.030)		(.004)
High SFHA × Disclosure × Post			$-.009$	
			(.006)	
Avg D.V.	0.675			0.095
State × Year × Stack FE			X	
Community × Stack FE			X	
Bandwidth	138	301		262
Num. obs.	1484709	1918077	499075	1701999

Note: Columns (1)–(2) and (4) are estimated based on equation (1) using the decennial census block-level data in 1990, 2000, 2010, and 2020. Columns (3) is estimated based on equation (2) using community-level population data. Standard errors are clustered at the state level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

percent relative to a non-SFHA block. Taking these extensive and intensive margin effects together, the policy discourages both in-migration into SFHA blocks with existing population and new developments in previously uninhabited SFHA blocks.

Using equation (2), I show in column (3) that the effect of the disclosure policy on the community level population is only –0.9 percent—nearly an order of magnitude smaller than in column (2).[18] This is plausible because, as Appendix Figure D.3 shows, a typical community has large non-SFHA land areas and thus flood risk can be easily avoided by within-community adjustments. Such local adjustments are consistent with prior research (Noonan and Sadiq 2019).

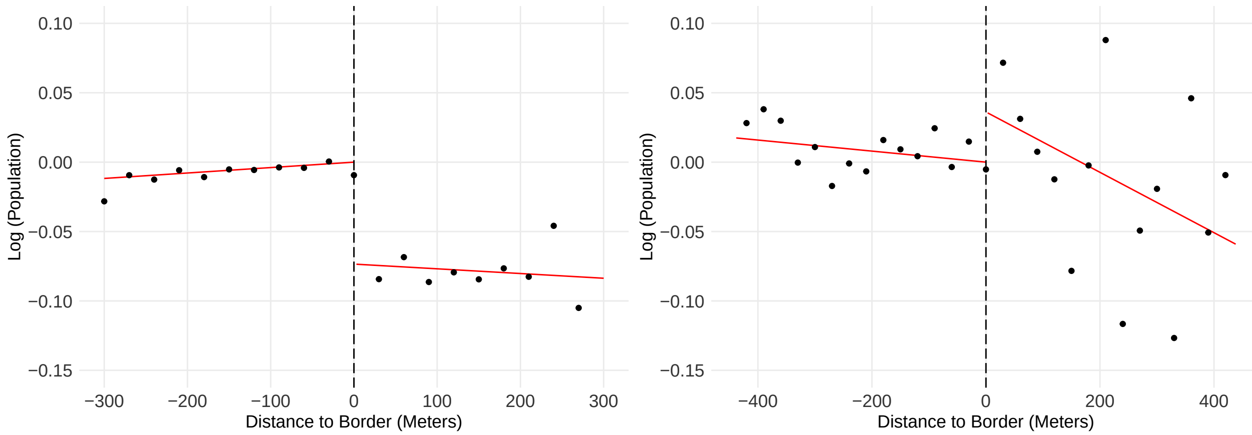
In column (4), I report that the disclosure increases the vacancy rate for the blocks in an SFHA from 0.095 to 0.109. This finding suggests that after the disclosure, selling properties in the SFHA becomes harder (or takes longer) and a larger share of them are vacant at any given time.[19] This finding is consistent with evidence that people migrate away from negative environmental conditions, although the extent of migration here is localized (Banzhaf and Walsh 2008, Boustan et al. 2012, Hornbeck 2012, Hornbeck and Naidu 2014).[20]

Figure 4.1 (a) graphically illustrates the effect in column (2) of Table 4.1. The horizontal axis is

[18] I use the decennial census to linearly interpolate annual community population.

[19] Indeed, New Orleans, one of the highest flood-risk areas in the nation, has the highest vacancy rate among the 75 largest MSAs in the US (Fudge and Wellburn 2014).

[20] When the housing supply is fixed, the disclosure will not affect the population distribution (market will clear via price adjustment alone). However, with an upward-sloping housing supply curve, which is likely to be the case given the time frame in Table 4.1, the disclosure will impact both the price and population distribution.



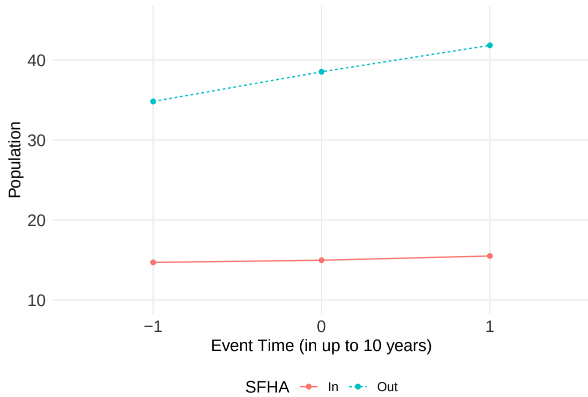
(a) Disclosed States (b) Placebo States

Figure 4.1: The Effect of Disclosure on Population. These figures illustrate difference-in-discontinuity estimates for the log of block population for the (a) disclosed and (b) placebo states. The discontinuity at the threshold (dashed vertical line) corresponds to the δ_6 term in equation (1). Dependent variables come from the decennial census block-level data in 1990, 2000, 2010, and 2020. The running variable is defined by the distance between a census block and the nearest SFHA border.

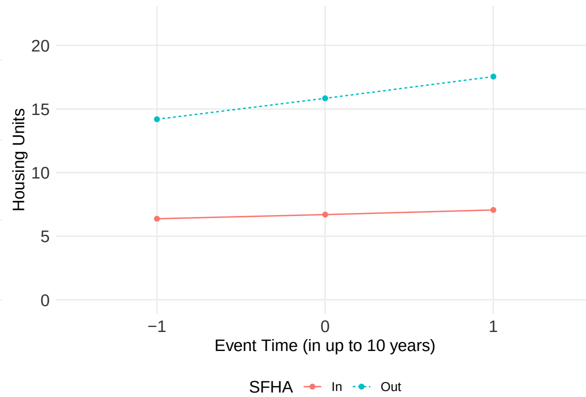
the distance between a block and the nearest SFHA border. The blocks within an SFHA are presented on the right-hand side of the border (the vertical line at 0), and the blocks outside of an SFHA are presented on the left-hand side. The solid lines represent the regression fit from equation (1) and the change in the logged population between the pre and post disclosure periods for the non-SFHA blocks is normalized to 0. I also overlay a scatterplot, which shows the difference in log population between pre and post treatment periods for each distance bin.

The figure indicates there is a sharp drop in the log population for SFHA blocks relative to the log population of non-SFHA blocks at the SFHA boundary. Visually, the discontinuous jump is approximately 0.07 log points, matching the estimate in column (2) of Table 4.1. Note, the regression line fits the scatter plot tightly, which suggests that the choice of functional form for the running variable is unlikely to have a substantial impact on the estimates.

To investigate the mechanism behind population adjustments, in Figure 4.2 (a), I plot the average population for census blocks in event time by the SFHA status. Here, event time is defined as -1 for pre-treatment periods, 0 for post-treatment periods up to 9 years, and 1 for post-treatment period of 10–19 years. The figure illustrates that the relative population of SFHA blocks is decreasing primarily due to an increasing population of non-SFHA blocks (blue dotted lines) rather than a shrinking population of SFHA blocks (red solid lines). Similarly, in panel (b), I find a rapid expansion of hous-



(a) Population



(b) Housing Units

Figure 4.2: Population and Housing Unit Trends in Event Time. These figures plot the (a) average population and (b) average number of housing units for SFHA and non-SFHA blocks within the optimal bandwidth in event time. Event time is defined as -1 for pre-treatment periods, 0 for post-treatment periods up to 9 years after the policy change, and 1 for post-treatment period of 10–19 years after the policy change.

ing units in non-SFHA blocks (blue dotted lines) and a stagnation for SFHA blocks (red solid lines). These empirical patterns suggest that the population adjustments reflect diverted in-migration (and resulting suppressed development) rather than active out-migration from SFHA areas, which is plausible given that the disclosure requirement provides new information to homebuyers rather than to homesellers or existing homeowners.

Two potential concerns regarding the validity of the results merit attention. First, as previously noted, location adjustments in large part appear local, suggesting a potential strategic sorting of households at the border. This not only challenges an underlying assumption in regression discontinuity design that individuals lack control over treatment status (Bosker et al. 2019), but also raises doubts about the deterrence effect of the disclosure policy. If diverted buyers opt for properties just outside the SFHA border, their exposure to flood risk essentially remains unchanged.

Second, while my difference-in-discontinuity design controls for time-invariant confounders, concurrent policy changes are still a threat to identification. Although, as discussed in Section 2.1, compliance with the flood insurance mandate was far from perfect, especially during the prime sample period of this study, one might worry that there have been changes in the enforcement of the flood insurance purchase requirement over time. Additionally, updates to flood maps could have served as a competing source of informational shock (Weill 2021, Hino and Burke 2021).

Per the first issue, I present three sets of results. First, I repeat my analysis using a doughnut difference-in-discontinuity approach that excludes blocks very close to the border. The idea is that if there is endogenous sorting near the border, the treatment effect may change when those observations are excluded (Cattaneo and Titiunik 2022). In Appendix Table D.3, I show that the estimates are similar even if I remove blocks that are within 20 or 40 meters from the border.[21] Second, in Appendix Figure D.6, I show that the policy effect does not diminish even if I expand the bandwidth. Finally, in Appendix Figure D.7, I repeat my analysis using a progressively farther away control blocks while holding treated blocks fixed (to those within the optimal bandwidth). In particular, I estimate equation (1) using control blocks that are within the distance of $(r-1) \times$ *optimal bandwidth* and $r \times$ *optimal bandwidth* for $r \in \{1,2,3,4,5\}$. Again, Appendix Figure D.7 shows that disclosure policy reduces population and increases vacancy rate. These findings suggest strategic sorting is not prevalent and deterred homebuyers choose locations that meaningfully lower flood risk.[22]

To address the second concern, I conduct three robustness checks. First, I allow time-varying discontinuity at the SFHA border to more directly control for confounding policy changes.[23] In Appendix Figure D.8, I show that allowing for time varying discontinuity does not change the previous conclusion. For a wide range of bandwidth choices, the effect size is much larger in magnitude (although with less precision due to a large number of parameters estimated) than the preferred specification in equation (1). Second, I use the five placebo states that have implemented disclosure policies without a question about flood risk. If my findings are driven by concurrent policy changes rather than the disclosure of flood risk, I would expect to find similar results in the placebo states. In Appendix Table D.2, I repeat Table 4.1 for placebo states and find no evidence of a reduction in population or an increase in the vacancy rate in the placebo states (Figure 4.1 (b) provides corresponding visual evidence). Similarly, Appendix Figure D.9 shows that my estimates are either null (for the vacancy rate) or positive (for population) for a wide range of bandwidth choices. Third, in Appendix Table D.4, I reproduce Table 4.1 after removing geographic units that have experienced

[21] Consistent with this, Appendix Figure D.5 shows that the distribution of running variable is generally smooth. A discrete change at the border may arise from other flood policies but such a level difference will cancel out by the difference-in-discontinuity design. Note, the number of observation monotonically decreases as the distance to the border increases because only blocks with an overlap with flood maps are in the sample.

[22] These findings also rules out the Stable Unit Treatment Value Assumption (SUTVA) violation, which could overestimate the effect size in Table 4.1.

[23] For this, I estimate $Y_{bst} = \delta_0 + \delta_1 X_{bs} + \delta_2 D_{bs} + \delta_3 X_{bs} * D_{bs} + \sum_{t=\{2000,2010,2020\}} G_t * [\delta_0^t + \delta_1^t X_{bs} + \delta_2^t D_{bs} + \delta_3^t X_{bs} * D_{bs}] + T_{st}[\delta_4 + \delta_5 X_{bs} + \delta_6 D_{bs} + \delta_7 X_{bs} * D_{bs}] + \epsilon_{bst}$. Here G_t is an indicator that takes 1 if time period is in t. Importantly, $G_t * \delta_2^t$ for $t = \{2000, 2010, 2020\}$ (1990 is omitted as baseline) allows period-specific discontinuities. The rest of notations follow equation (1) and coefficient of interest is δ_6 as before.

flood map updates over the sample period and find that the results barely change.[24]

While choosing a safe location represents an extensive margin self-protection strategy, previous studies consider property elevation as a potentially important intensive margin responses (Mobley et al. 2020). Although data limitations do not allow me to analyze this possibility directly, property elevation is unlikely to be a widely adopted self-protection measure because of its high cost. For instance, the median cost of elevations through the FEMA mitigation program (between 2008 and 2013) is $166,000 (National Research Council 2015), which is over 50 percent of the average property value in the SFHA area ($327,171). In addition, elevation takes at least several months to complete, which means that the foregone use value is also substantial. This assessment is consistent with Montgomery and Kunreuther (2018), which finds that property elevation in general is rarely cost effective.

Market Insurance. In Table 4.2 column (1), I show the disclosure policy increases the probability of having at least one flood insurance policy in high-risk communities relative to low-risk communities by 0.003 percentage points (or 0.4 percent from the baseline of 0.82). Column (2) indicates the intensive margin effect of the disclosure policy on the number of insurance policies per housing unit is also small at –2 percent. Given the point estimates, flood insurance does not seem to be the primary margin homebuyers respond to the disclosure policy.

In Appendix Figure D.10, I plot the differential impact of disclosure policy on the probability of having flood insurance for high-risk communities in event time using an event study version of equation (2). The estimated coefficients show no pre-trend and a small increase in the probability of flood insurance take up after the policy change. Simple average of estimated coefficients in the pre vs. post treatment event time suggests that the magnitude of policy effect is 0.007, which is larger than column (1) in Table 4.2 yet still small. Further, in Appendix Table D.5, I reproduce Table 4.2 after removing communities that have experienced map update during the sample period and find that the result does not change.

Why do homebuyers engage in self-protection despite the option to buy flood insurance? One possibility is that the cost of location adjustment is substantially lower for many homebuyers especially compared to households not intending to move. Indeed, Zumpano et al. (2003) documents that

[24] To generate the list of communities that have experienced map updates, I use the "L_Comm_Revis" layer from the National Flood Hazard Layer from FEMA (FEMA 2019). 7% of blocks in my sample are located in communities that had any map update during the sample period.

Table 4.2: Effect of Discosure Requirement on Flood Insurance Take-Up

	Prob. of Any Insurance	Log Insurance Per Housing Unit
	(1)	(2)
High SFHA × Disclosure × Post	.003	−.024
	(.007)	(.030)
Avg D.V.	0.823	
State × Year × Stack FE	X	X
Community × Stack FE	X	X
Num. obs.	400919	329863

Note: This table is produced from equation (2) using community-level NFIP data. Standard errors are clustered at the state level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

homebuyers actively search across alternatives, and an average buyer physically visits 17 properties before closing. Moreover, the benefit of flood insurance could be insufficient. For instance, the NFIP offers incomplete insurance with coverage capped at $250,000. Further, a flood could negatively affect an individual's health or employment status and disrupt daily life, all of which are not covered by flood insurance (Kahn 2005, Deryugina 2017, Lee et al. 2023).

5 The Effect of the Disclosure Requirement on Flood Damage

5.1 Estimation Framework

For a flood of a given size, how does flood damage change after the disclosure requirement? To answer this question, I estimate a damage function, which is a mapping between flood size and damage, and show how the functional relationship changes due to the policy. Damage functions have been widely used in the economics literature to understand the relationship between heat and economic outcomes.[25] Surprisingly, there has been limited attention directed towards damage functions specific to floods, despite the substantial disruptions they cause. This is partly because objective measurement of flood size is difficult. I overcome this challenge by constructing a hydrology-based flood history dataset described in detail in Appendix A.1.

$$\text{Per Housing Unit Damage} = \sum_k [\alpha_1^k F^k + \alpha_2^k F^k D] \tag{3}$$

[25] For a review, see Dell et al. (2014), Carleton and Hsiang (2016), and Auffhammer (2018).

Before describing the estimation procedure, it is useful to conceptualize the damage function. Consider equation (3), where the dependent variable is flood damage per housing unit, D is an indicator variable for the treated (i.e., disclosed) group assignment and F^k is an indicator variable equal to 1 when the annual maximum flood size is in flood size bin k where $k \in \{$2-10, 10-20, 20-30, 30-40, 40-50$\}$.[26] Equation (3) follows a non-parametric approach of Barreca et al. (2016), which lets the data rather than the functional form assumption, determine the shape of the damage function. In my models, flood sizes between 1–2 serve as the baseline omitted category. Thus, α_1^k is the additional damage per housing unit when a community in the control group experiences a flood of size k as opposed to a flood size between 1–2. α_2^k allows a different slope for the treated group for flood size k.

There are three points to discuss regarding F^k. First, I use the annual maximum flood size as a proxy for flood exposure for a given community-year. While this means smaller floods in the same community and year are ignored, this is unlikely to be a practical concern because the majority of the community-years in the dataset had just one flood (Appendix Figure A.3 (c)). Moreover, after restricting attention to floods of size over 10 or larger, which cause disproportionately large damage, over 90 percent of community-years have only one incident (Appendix Figure A.3 (d)).

Second, I focus on flood sizes between 1 and 50 because larger floods are frequently accompanied by interrelated perils, which cause substantial measurement errors (Kron et al. 2012). Further, as shown in Appendix Figure A.3 (b), the frequency of flood events reduces exponentially as flood size increases, making it difficult to non-parametrically identify statistical relationships for very large floods. Appendix Table A.2 shows that the threshold for flood sizes 10 and 50 are closely matched to the threshold for "moderate" and "major" floods defined by the National Weather Service, indicating that the chosen flood sizes cover a wide enough band to capture floods of different severity. Third, the assumption behind binning is that the damage per housing unit is identical within each k. While flood sizes of 41 and 49, for example, might have a different effect in reality, I choose a bin size of 10 to strike a balance between flexibility and precision (Barreca et al. 2016).

$$\text{Per Housing Unit Damage} = \sum_k [\beta_1^k F^k + \beta_2^k F^k I + \beta_3^k F^k D + \beta_4^k F^k ID] \tag{4}$$

Equation (4), which mirrors a canonical difference-in-differences model, shows how equation (3)

[26] For instance, $k = 40 - 50$ means the flood size is of the magnitude that would be expected every 40–50 years. For more details see Section 3 and Appendix A.1.

changes when the post disclosure indicator I is introduced. The coefficient for the interaction term (β_4^k) captures the treatment effect.

$$Y_{mtd} = \sum_k [\beta_1^k F_{mtd}^k + \beta_2^k F_{mtd}^k I_{mtd} + \beta_3^k F_{mtd}^k D_{mtd} + \beta_4^k F_{mtd}^k I_{mtd} D_{mtd}] + \theta_{md} + \omega_{td} + \epsilon_{mtd} \tag{5}$$

For estimation, I use equation (5). Y_{mtd} is either an indicator variable for positive flood damage in community m (extensive margin), or log(Per Housing Unit Damage) conditional on having positive damage for community m (intensive margin), in year t for data stack d.[27] While I report both the extensive and intensive margin effects, an emphasis is given to the former due to greater generalizability—only a small fraction of communities experience repeated damage—and higher statistical power. I also includes year × stack (ω_{td}) and community × stack (θ_{md}) fixed effects, to control for overall time trend and unobserved community characteristics. Similar to Section 4.2, I exploit the timing of the disclosure requirement for identification and thus not-yet-treated states form the control group. Also, equation (5) is a stacked difference-in-differences model (i.e., ignoring differences in treatment intensity), a choice made for tractability, although the differences in treatment intensity margin (i.e., fraction of the area covered by an SFHA) is exploited in estimating heterogeneous treatment effects. I use 20 years of observation for each state around the disclosure policy change year.

Because the impact of natural disasters is not confined by administrative units, previous studies on cyclone damage function have used spatial-HAC standard errors (Hsiang 2010). Following this, I allow spatial correlation of up to 500 miles for inference (Newey and West 1987, Conley 1999), but I also show that state-level clustering produces similar results.[28]

Before proceeding further, it is worth briefly discussing the difference between the damage function of this paper and "depth-damage functions" from earlier engineering studies. As its name suggests, the measure of flood size in these engineering studies is the water depth for an individual property (Meyer et al. 2013). While useful for predicting property-level flood damage, this approach has two limitations for estimating aggregate flood damage. First, by focusing on an individual property, it does not directly take into account that a larger flood increases the number of affected properties. A

[27] Data stacks are constructed in a similar procedure described in Section 4.1.

[28] Weights in this matrix are uniform up to that cutoff distance. When the variance-covariance matrix is not positive-semidefinite, I use eigendecomposition of the estimated variance matrix and convert any negative eigenvalue(s) to zero following Cameron et al. (2011).

detailed hydraulic study can overcome this issue, but most communities either lack access to them or are reliant on outdated versions because of cost issues (FEMA 2005, Bakkensen and Ma 2020, Weill 2021). Second, and more importantly, it typically lacks the capability to account for adaptations at the property level, which is likely to cause bias in the estimated damage function.[29]

I take a "reduced-form" approach to overcome these issues. By directly relating flood size to flood damage at the community level, my approach can be applied even in areas without up-to-date hydraulic studies. Moreover, the community level damage metric factors in the number of properties damaged and embeds the impact of any pre-existing adaptation measures.

5.2 Findings

In Figure 5.1, I plot the damage functions for the (a) control and (b) treatment groups using the estimated coefficients from equation (5).[30] For instance, $\hat{\beta}_1^k$ and $\hat{\beta}_1^k + \hat{\beta}_3^k$ for each k are used to plot the pre-treatment period damage functions for panel (a) and (b), respectively. Because the dependent variable in Figure 5.1 is the probability of any damage, the estimated coefficients indicate the additional probability of damage when the baseline flood (k=1–2) is replaced by a flood of size k.

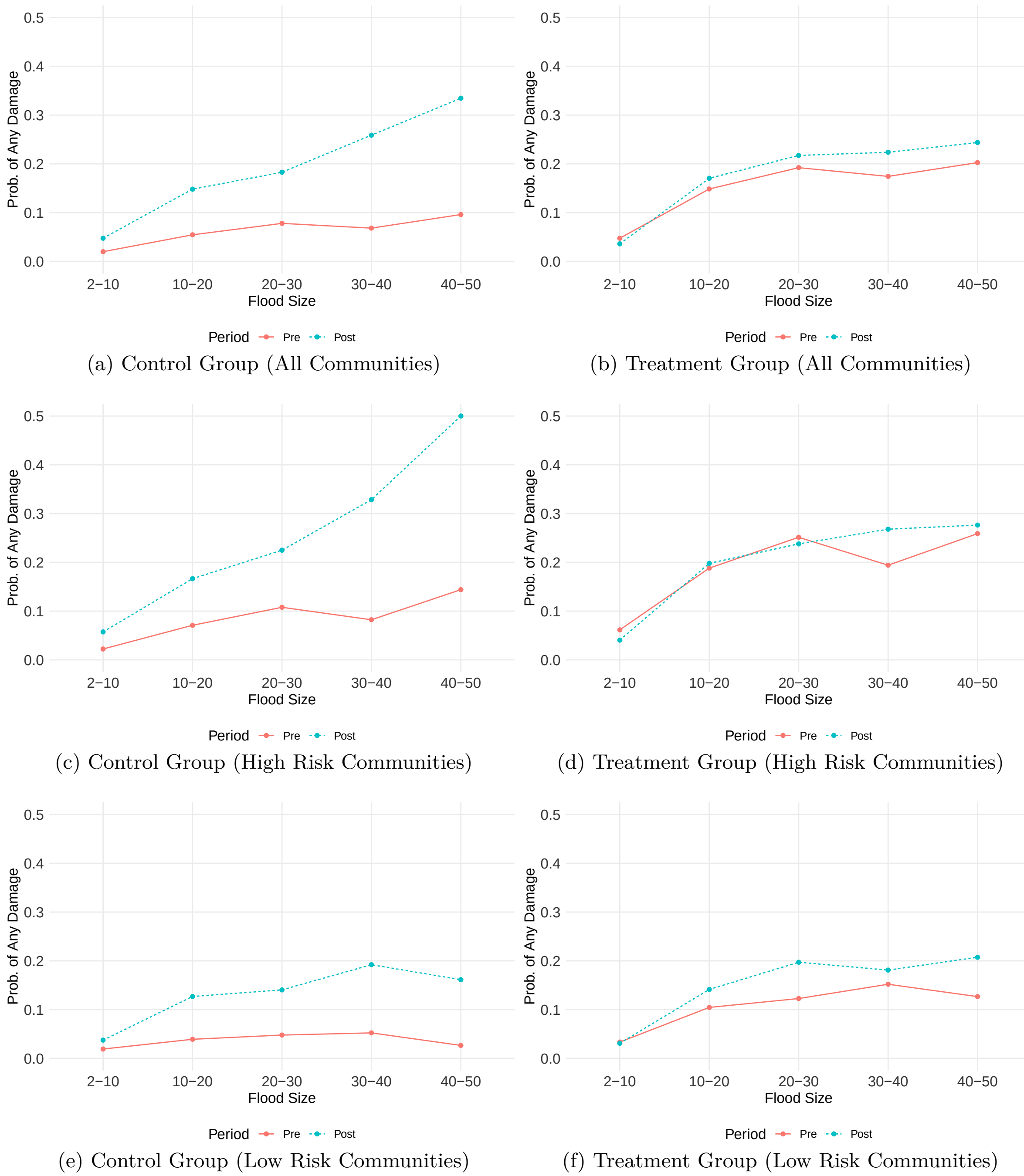
Figure 5.1 allows visual inspection of the validity of the estimated damage function. To begin, I first focus on the slope of this function, which reveals a monotonically increasing relationship between flood size and the probability of any flood damage. In subsequent panels (c)–(f), I test heterogeneity in the damage function for further assessment. That is, even when faced with floods of the same size (as defined by community-specific recurrence intervals, which, heuristically, can be considered as deviations from local averages), communities with higher risk should exhibit higher levels of damage.[31] Indeed, I find that high risk communities in panels (c)–(d) (an above-median fraction of the area covered by an SFHA) have much higher vertical levels and steeper slopes in comparison to the low risk communities in panels (e)–(f).

Table 5.1 highlights the impact of the disclosure requirement on flood damage. For brevity, I only report $\hat{\beta}_4^k$ from equation (5), but the full sets of coefficients are in Appendix Table D.6. In column (1),

[29] In theory, this issue can be addressed by (1) modeling how various defensive measures such as property elevation or the use of waterproof building materials impact damage, and (2) collecting property level data on these defensive measures. However, this approach is impractical due to limitations in modeling techniques and data availability.

[30] Appendix Figure D.11 reproduces Figure 5.1 with a 95 percent confidence interval.

[31] To illustrate this, consider two communities, A and B with starkly different risk profiles: A is entirely situated within the SFHA while B lies outside the SFHA. In the event of a 100-year flood, the entire property in A would be expected to be underwater (by the definition of SFHA), whereas B would remain unaffected.



(a) Control Group (All Communities)
(b) Treatment Group (All Communities)
(c) Control Group (High Risk Communities)
(d) Treatment Group (High Risk Communities)
(e) Control Group (Low Risk Communities)
(f) Treatment Group (Low Risk Communities)

Figure 5.1: The Effect of Disclosure on the Damage Function. These plots illustrate estimated damage functions (dep.var: probability of any damage) from equation (5). Panels (a)–(b) are damage functions for all communities. Panels (c)–(d) and (e)–(f) illustrate the damage functions for high (above–median SFHA ratio) and low (below–median SFHA ratio) flood risk communities, respectively. Appendix Figure D.11 reproduces Figure 5.1 with 95% confidence intervals.

Table 5.1: Effect of Disclosure Requirement on Flood Damage

	Prob. of Any Damage Per Housing Unit			Log Damage Per Housing Unit
	(1)	(2)	(3)	(4)
Post × Disclosure (Size 2-10)	$-.039^{*}$	$-.056^{*}$	$-.021$	$-.063$
	(.023)	(.030)	(.015)	(.325)
Post × Disclosure (Size 10-20)	$-.072^{*}$	$-.086^{*}$	$-.051^{*}$	.189
	(.039)	(.050)	(.029)	(.189)
Post × Disclosure (Size 20-30)	$-.080^{***}$	$-.131^{***}$	$-.018$	.170
	(.029)	(.038)	(.031)	(.562)
Post × Disclosure (Size 30-40)	$-.141^{*}$	$-.172^{**}$	$-.111$	$-.360$
	(.073)	(.072)	(.082)	(.442)
Post × Disclosure (Size 40-50)	$-.197^{***}$	$-.339^{***}$	$-.054$	$-.425$
	(.055)	(.061)	(.068)	(.540)
Annual Effect	$-.023$	$-.034$	$-.012$	$-.012$
	(0.009)	(0.011)	(0.008)	(0.063)
Sample	All	High SFHA	Low SFHA	Damage > 0
Year × Stack FE	X	X	X	X
Community × Stack FE	X	X	X	X
Num. obs.	505383	242458	262925	22100

Note: The dependent variable in columns (1) to (3) is the probability of having any flood damage per housing unit. Column (1) is from the entire set of communities while in columns (2) and (3), I repeat (1) using the subsample of communities with different levels of risk exposure. Dependent variables in columns (4) is log transformed per housing unit damage. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

I report the policy effect using all communities in my sample. The results show that the disclosure requirement reduces the probability of having any flood damage per housing unit by 4–20 percentage points for different values of k for the communities in the disclosed states relative to the ones in the not-yet-disclosed states.[32] The damage reduction effect can be verified visually as well: Figure 5.1 shows that in panel (a) (control), flood probability has substantially increased over time, whereas in panel (b) (treated), it remains nearly identical.

Using equation (6), I summarize the coefficients in Table 5.1 into probability-weighted average treatment effects. Note, because $Pr(K = k)$ is the likelihood of flood occurrence for bin size k each year and β_4^k is the change in probability of having damage, equation (6) can be interpreted as the reduction in annualized loss probability due to the disclosure policy.[33]

$$\sum_k Pr(K = k) \times \beta_4^k \tag{6}$$

[32]For per housing unit damage, I divide community-year level damage using the housing stock in 1990.

[33]Since the flood size is defined by recurrence interval, the inverse of the size corresponds to $Pr(K = k)$. For instance, the probability of having a flood of size 40–50 in a given year is $\frac{1}{45}$ (45 is the median value of the bin).

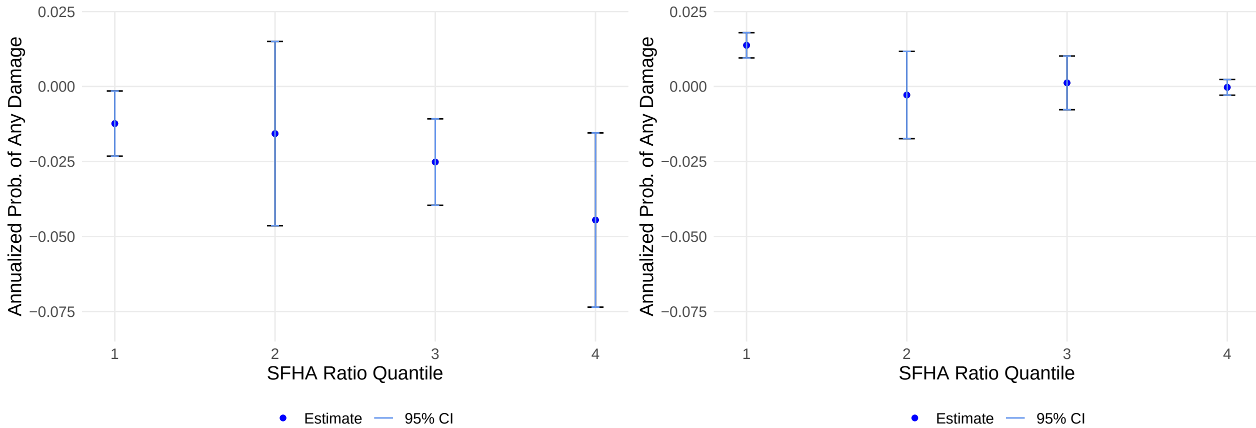
In Table 5.1 column (1), I report that the reduction in the annualized loss probability is 2.3 percentage points. When I compare this with the baseline of 7.4—average probability of having any damage conditional on exposure to a flood of size 2 or larger—the effect size is a 31 percent reduction.

In columns (2) and (3), I split the sample into communities with above- and below-median fractions of the area covered by an SFHA to investigate the heterogeneous treatment effects. Because the disclosure policy targets properties in an SFHA, the policy effect should be driven by the high-SFHA communities. Indeed, the reduction in the annualized loss probability is three times larger for high-SFHA communities than low-SFHA communities. Figure 5.2 (a) further splits columns (2) and (3) and presents the change in the annualized probability of having any damage for four groups of communities with differential SFHA ratios, which clearly shows monotonically increasing magnitude in the SFHA ratio.

Column (4) reports the intensive margin effect, where the dependent variable is the log of damage per housing unit. Because the sample for this exercise is restricted to community-years with positive damage, the model is under powered. Still, I find suggestive evidence that the disclosure policy reduces damage for communities with repetitive flood events.

Taken together, the disclosure requirement reduces flood damage in treated communities compared to controls, primarily due to a higher damage increase in the control group. Appendix Figure D.12 provides a potential explanation for this result by showing how population and housing unit within the SFHA have changed over years by the treatment status. For this, using decennial census block data from Section 4.2, I regress outcomes on decennial census year dummies interacted with early treatment (disclosure between 1990 and 2000) status with block fixed effects. The difference between the two groups in 1990 is normalized to zero. In panel (a), the point estimate for the year 2000 indicates a lower likelihood of development in previously uninhabited SFHAs in early-disclosed states. Interestingly, the difference between two groups disappear by 2010, coinciding with the full implementation of disclosure policies including late-disclosed states. A similar pattern is observed for the probability of having any housing unit in Panel (b). These suggestive evidence implies that the damage increase in the control group is likely attributable to a faster population and housing units growth in high-risk areas in the absence of flood risk disclosure.

While the estimated impact of a simple disclosure policy is non-trivial, this number is likely to underestimate the true benefit because the analysis excludes flood sizes over 50, which incur dispro-



(a) Disclosed States (b) Placebo States

Figure 5.2: Change in Annualized Loss Probability by the SFHA Ratio. These figures plot the annualized damage reduction effect from the disclosure requirement by SFHA ratios for (a) disclosed and (b) placebo states. I estimated equation (5) for communities in different quantile of SFHA ratio, and aggregated coefficients using equation (6).

portionately large damage. Besides, I abstract away from a potential gain due to a better matching in flood risk preferences between properties and homebuyers (Bakkensen and Ma 2020).

Robustness check. I test the robustness of my findings by conducting a placebo test using the five states that had implemented disclosure policies without a question on the flood risk. In Appendix Table D.7, I estimate a version of equation (5) with coarser flood bins.[34] In columns (1) to (3), the coefficients suggest that disclosure without flood risk information does not reduce the probability of flood damage at all. The estimates are statistically insignificant and economically small, which is consistent with Figure 5.2 (b): the effect is zero for all four groups of communities with varying SFHA exposure.

Another robustness check comes from an event study plot in Appendix Figure D.13, which illustrates the marginal effects of disclosure policy on the probability of flood damage for larger ($k = 30 - 50$) floods. Similar to Appendix Table D.7, I use coarser flood bins to increase power. I also impose an endpoint restriction at -5 and 4. The estimated coefficients show no pre-trend and sharp and persistent reduction in the probability of flood damage after the policy change. Simple average of estimated coefficients in the pre vs. post treatment event time suggests that the magnitude of policy effect is $-.19$. This on par with the average policy effect ($-.17$) of two flood categories $k = 30 - 40$ and

[34]For statistical power, I group flood events into baseline ($k = 1 - 2$), small ($k = 2 - 30$) and large ($k = 30 - 50$).

$k = 40 - 50$ in column (1) in Table 5.1.

Appendix Table D.8 shows that excluding communities with map revision produces essentially the same results. Finally, Appendix Table D.9 shows that clustering standard errors at the state level reaches similar conclusion as Table 5.1 especially for the annualized effects.

6 Conclusion

Floods are the costliest natural disaster in the US and are expected to become more frequent and severe in the future. Thus, curbing economic loss from these events is of first-order importance. The primary policy prescription in the US to flood risk is engineering-based—i.e., using physical structures and other building-based responses to reduce damage. However, this approach can attract more people to areas with flood risk by distorting location choices.

In this paper, I study whether alleviating information frictions regarding flood risk in the housing market can be an effective way to foster adaptation. By exploiting plausibly exogenous variations created by flood-risk disclosure requirements, I explore if and how home buyers respond to required flood-risk disclosure and investigate its implications for flood damage. I find the disclosure requirement reduces the population and increases the vacancy rate in high-risk areas. With fewer people exposed to flood risk, the probability of having any flood damage decreases by 2.3 percentage points (or 31 percent from the baseline probability). The findings of this paper shows that the disclosure requirement can facilitate voluntary adaptation by helping homebuyers make more informed choices.

References

Aldy, J. E., and R. Zeckhauser. 2020. Three Prongs for Prudent Climate Policy. Southern Economic Journal 87:3–29.

Amlani, S., and C. Algara. 2021. Partisanship & nationalization in American elections: Evidence from presidential, senatorial, & gubernatorial elections in the U.S. Counties, 1872–2020. Electoral Studies 73:102387.

Bakkensen, L. A., and L. Barrage. 2021. Going Underwater? Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics. The Review of Financial Studies:hhab122.

Bakkensen, L. A., and L. Ma. 2020. Sorting over flood risk and implications for policy reform. Journal of Environmental Economics and Management 104:102362.

Banzhaf, H. S., and R. P. Walsh. 2008. Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism. American Economic Review 98:843–863.

Barreca, A., K. Clay, O. Deschenes, M. Greenstone, and J. S. Shapiro. 2016. Adapting to Climate Change: The Remarkable Decline in the US Temperature-Mortality Relationship over the Twentieth Century. Journal of Political Economy 124:105–159.

Baylis, P. W., and J. Boomhower. 2021. Mandated Vs. Voluntary Adaptation to Natural Disasters: The Case of U.S. Wildfires. NBER Working Paper.

Baylis, P., and J. Boomhower. 2022. The Economic Incidence of Wildfire Suppression in the United States. American Economic Journal: Applied Economics:51.

Bin, O., and C. E. Landry. 2013. Changes in implicit flood risk premiums: Empirical evidence from the housing market. Journal of Environmental Economics and Management 65:361–376.

Bosker, M., H. Garretsen, G. Marlet, and C. Van Woerkens. 2019. Nether Lands: Evidence on the Price and Perception of Rare Natural Disasters. Journal of the European Economic Association 17:413–453.

Boustan, L. P., M. E. Kahn, and P. W. Rhode. 2012. Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century. American Economic Review 102:238–244.

Brot-Goldberg, Z., T. Layton, B. Vabson, and A. Y. Wang. 2020. The Behavioral Foundations of Default Effects: Theory and Evidence from Medicare Part D. Working Paper:87.

Bureau of the Census. 1994. Geographic Areas Reference Manual.

Burke, M., and K. Emerick. 2016. Adaptation to Climate Change: Evidence from US Agriculture. American Economic Journal: Economic Policy 8:106–140.

Burke, M., S. M. Hsiang, and E. Miguel. 2015. Global non-linear effect of temperature on economic production. Nature 527:235–239.

Calonico, S., M. D. Cattaneo, and R. Titiunik. 2014. Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs: Robust Nonparametric Confidence Intervals. Econometrica 82:2295–2326.

Cameron, A. C., J. B. Gelbach, and D. L. Miller. 2011. Robust Inference With Multiway Clustering. Journal of Business & Economic Statistics 29:238–249.

Carleton, T. A., and S. M. Hsiang. 2016. Social and economic impacts of climate. Science 353:aad9837–aad9837.

Cattaneo, M. D., N. Idrobo, and R. Titiunik. 2019. A Practical Introduction to Regression Discontinuity Designs: Foundations. arXiv:1911.09511 [econ, stat].

Cattaneo, M. D., and R. Titiunik. 2022. Regression Discontinuity Designs. Annual Review of Economics 14:821–851.

CBO. 2016. Potential Increases in Hurricane Damage in the United States: Implications for the Federal Budget. Congressional Budget Office:46.

Cengiz, D., A. Dube, A. Lindner, and B. Zipperer. 2019. The Effect of Minimum Wages on Low-Wage Jobs. The Quarterly Journal of Economics 134:1405–1454.

Changnon, S. A., R. A. Pielke Jr, D. Changnon, R. T. Sylves, and R. Pulwarty. 2000. Human Factors Explain the Increased Losses from Weather and Climate Extremes. Bulletin of the American Meteorological Society 81:437–442.

Chen, J., and J. Roth. 2022. Log-like? Identified ATEs Defined with Zero-valued Outcomes are (Arbitrarily) Scale-dependent. Working Paper.

Chivers, J., and N. E. Flores. 2002. Market Failure in Information: The National Flood Insurance Program. Land Economics 78:515–521.

Cicco, L. A. D., D. Lorenz, R. M. Hirsch, and W. Watkins. 2018. dataRetrieval: R packages for discovering and retrieving water data available from U.S. Federal hydrologic web services. U.S. Geological Survey, Reston, VA.

Collenteur, R. A., H. de Moel, B. Jongman, and G. Di Baldassarre. 2015. The failed-levee effect: Do societies learn from flood disasters? Natural Hazards 76:373–388.

Conley, T. G. 1999. GMM estimation with cross sectional dependence. Journal of Econometrics 92:1–45.

Davenport, F. V., M. Burke, and N. S. Diffenbaugh. 2021. Contribution of historical precipitation change to US flood damages. Proceedings of the National Academy of Sciences 118:e2017524118.

Deryugina, T. 2017. The Fiscal Cost of Hurricanes: Disaster Aid versus Social Insurance. American Economic Journal: Economic Policy 9:168–198.

DHS Office of Inspector General. 2017. FEMA Needs to Improve Management of Its Flood Mapping Programs. Department of Homeland Security.

Edmund, H., S. Chamberlain, and K. Ram. 2014. Rnoaa: NOAA climate data from R.

Ehrlich, I., and G. S. Becker. 1972. Market Insurance, Self-Insurance, and Self-Protection. Journal of Political Economy 80:623–648.

England Jr, J. F., T. A. Cohn, B. A. Faber, J. R. Stedinger, W. O. Thomas Jr, A. G. Veilleux, J. E. Kiang, and R. R. Mason Jr. 2019. Guidelines for determining flood flow frequency—Bulletin 17C. US Geological Survey.

Fairweather, D., M. E. Kahn, R. D. Metcalfe, and S. Sandoval-Olascoaga. 2023. PRELIMINARY: The Impact of Climate Risk Disclosure on Housing Search and Buying Dynamics: Evidence from a Nationwide Field Experiment with Redfin. Working Paper.

Felbermayr, G., and J. Gröschl. 2014. Naturally negative: The growth effects of natural disasters. Journal of Development Economics 111:92–106.

FEMA. 1996. Q3 Flood Data Users Guide, Draft. FEMA.

FEMA. 2005. National Flood Insurance Program (NFIP) Floodplain Management Requirements: A Study Guide and Desk Reference for Local Officials. FEMA.

FEMA. 2011. Answers to Questions About the NFIP. FEMA.

FEMA. 2014. Transaction Record Reporting and Processing (TRRP) Plan. FEMA.

FEMA. 2019. Flood Insurance Rate Map (FIRM) Database Technical Reference.

Field, C. B., V. Barros, T. F. Stocker, and Q. Dahe (Eds.). 2012. Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation: Special Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Fudge, K., and R. Wellburn. 2014. Vacant and abandoned properties: Turning liabilities into assets. Evidence Matters.

Fuller, W. E. 1913. Flood Flows. American Society of Civil Engineers.

Gall, M., K. A. Borden, and S. L. Cutter. 2009. When Do Losses Count?: Six Fallacies of Natural Hazards Loss Data. Bulletin of the American Meteorological Society 90:799–810.

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.

GAO. 2021. NATIONAL FLOOD INSURANCE PROGRAM: Congress Should Consider Updating the Mandatory Purchase Requirement.

Ghanbari, M., M. Arabi, J. Obeysekera, and W. Sweet. 2019. A Coherent Statistical Model for Coastal Flood Frequency Analysis Under Nonstationary Sea Level Conditions. Earth's Future 7:162–177.

Gibson, M., and J. T. Mullins. 2020. Climate Risk and Beliefs in New York Floodplains. Journal of the Association of Environmental and Resource Economists 7:1069–1111.

Goodman-Bacon, A. 2021. Difference in Differences with Variation in Treatment Timing. Journal of Econometrics 225:254–277.

Gourley, J. J., Y. Hong, Zach ary L. Flamig, Ami Arthur, Robert Clark, Martin Calianno, Isabelle Ruin, Terry Ortel, Mich ael E. Wieczorek, and Kirstetter. 2013. A Unified Flash Flood Database over the US. American Meteorological Society:799–805.

Gregory, J. 2017. The Impact of Post-Katrina Rebuilding Grants on the Resettlement Choices of New Orleans Homeowners:53.

Grembi, V., T. Nannicini, and U. Troiano. 2016. Do Fiscal Rules Matter? American Economic Journal: Applied Economics 8:1–30.

Hallstrom, D. G., and V. K. Smith. 2005. Market responses to hurricanes. Journal of Environmental Economics and Management 50:541–561.

Hino, M., and M. Burke. 2021. The effect of information about climate risk on property values. Proceedings of the National Academy of Sciences 118:e2003374118.

Hornbeck, R. 2012. The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe. American Economic Review 102:1477–1507.

Hornbeck, R., and S. Naidu. 2014. When the Levee Breaks: Black Migration and Economic Development in the American South. American Economic Review 104:963–990.

Hsiang, S. 2016. Climate Econometrics. Annual Review of Resource Economics 8:43–75.

Hsiang, S. M. 2010. Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. Proceedings of the National Academy of Sciences 107:15367–15372.

Hsiang, S., and A. Jina. 2014. The Causal Effect of Environmental Catastrophe on Long-Run Economic Growth: Evidence From 6,700 Cyclones. National Bureau of Economic Research, Cambridge, MA.

IPCC. 2014. Climate Change 2014 Mitigation of Climate Change: Working Group III Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge.

Jackson, L. E. 2013. Frequency and Magnitude of Events. Pages 359–363 *in* P. T. Bobrowsky, editor. Encyclopedia of Natural Hazards. Springer, Dordrecht.

Kahn, M. E. 2005. The Death Toll from Natural Disasters: The Role of Income, Geography, and Institutions. Review of Economics and Statistics 87:271–284.

Kousky, C. 2019. The Role of Natural Disaster Insurance in Recovery and Risk Reduction. Annual Review of Resource Economics 11:399–418.

Kousky, C., E. F. P. Luttmer, and R. J. Zeckhauser. 2006. Private investment and government protection. Journal of Risk and Uncertainty 33:73–100.

Kousky, C., and E. Michel-Kerjan. 2015. Examining Flood Insurance Claims in the United States: Six Key Findings: Examining Flood Insurance Claims in the United States. Journal of Risk and Insurance 84:819–850.

Kousky, C., E. O. Michel-Kerjan, and P. A. Raschky. 2018. Does federal disaster assistance crowd out flood insurance? Journal of Environmental Economics and Management 87:150–164.

Kron, W., M. Steuer, P. Löw, and A. Wirtz. 2012. How to deal properly with a natural catastrophe database – analysis of flood losses. Natural Hazards and Earth System Sciences 12:535–550.

Kunreuther, H., and M. Pauly. 2004. Neglecting Disaster: Why Don't People Insure Against Large Losses? Journal of Risk and Uncertainty 28:5–21.

LARSON, L. W. 1996. The Great USA Flood of 1993. Presented at IAHS Conference, Anaheim, CA.

Lee, S., X. Wan, and S. Zheng. 2023. Estimating the Inconvenience Cost of Floods: Evidence from High-Tide Flooding. Working Paper.

Lefcoe, G. 2004. Property Condition Disclosure Forms: How the Real Estate Industry Eased the Transition from Caveat Emptor to "Seller Tell All". Real Property, Probate and Trust Journal 39:193–250.

Liao, K.-H. 2014. From flood control to flood adaptation: A case study on the Lower Green River Valley and the City of Kent in King County, Washington. Natural Hazards 71:723–750.

Maimone, M., and T. Adams. 2023. A Practical Method for Estimating Climate-Related Changes to Riverine Flood Elevation and Frequency. Journal of Water and Climate Change 14:748–763.

Mallakpour, I., and G. Villarini. 2015. The changing nature of flooding across the central United States. Nature Climate Change 5:250–254.

Manson, S. M., J. Schroeder, D. V. Riper, T. Kugler, and S. Ruggles. 2022. IPUMS national historical geographic information system: Version 17.0. Minneapolis, MN.

Meyer, V., N. Becker, V. Markantonis, R. Schwarze, J. C. J. M. van den Bergh, L. M. Bouwer, P. Bubeck, P. Ciavola, E. Genovese, C. Green, S. Hallegatte, H. Kreibich, Q. Lequeux, I. Logar, E. Papyrakis, C. Pfurtscheller, J. Poussin, V. Przyluski, A. H. Thieken, and C. Viavattene. 2013. Review article: Assessing the costs of natural hazards – state of the art and knowledge gaps. Natural Hazards and Earth System Sciences 13:1351–1373.

Miao, Q., and D. Popp. 2014. Necessity as the Mother of Invention: Innovative Responses to Natural Disasters. Journal of Environmental Economics and Management 68:280–295.

Michel-Kerjan, E. O. 2010. Catastrophe Economics: The National Flood Insurance Program. Journal of Economic Perspectives 24:165–186.

Michel-Kerjan, E. O., and C. Kousky. 2010. Come Rain or Shine: Evidence on Flood Insurance Purchases in Florida. Journal of Risk and Insurance 77:369–397.

Milly, P. C. D., R. T. Wetherald, K. A. Dunne, and T. L. Delworth. 2002. Increasing risk of great floods in a changing climate. Nature 415:514–517.

Mobley, W., K. O. Atoba, and W. E. Highfield. 2020. Uncertainty in Flood Mitigation Practices: Assessing the Economic Benefits of Property Acquisition and Elevation in Flood-Prone Communities. Sustainability 12:2098.

Montgomery, M., and H. Kunreuther. 2018. Pricing Storm Surge Risks in Florida: Implications for Determining Flood Insurance Premiums and Evaluating Mitigation Measures: Pricing Storm Surge Risks in Florida. Risk Analysis 38:2275–2299.

Mullahy, J., and E. C. Norton. 2022. Why Transform Y? A Critical Assessment of Dependent-Variable Transformations in Regression Models for Skewed and Sometimes-Zero Outcomes. NBER Working Paper.

Mullenix, L. S. 2019. Regulatory and judicial consumer protection in the United States of America: An assessment. Stellenbosch Law Review 30:33–60.

Muller, N., and C. Hopkins. 2019. Hurricane Katrina Floods New Jersey: The Role of Information in the Market Response to Flood Risk. Page w25984. National Bureau of Economic Research, Cambridge, MA.

National Association of Realtors. 2019. State Flood Hazard Disclosures Survey.
National Research Council. 2015. Affordability of National Flood Insurance Program Premiums: Report 1. National Academies Press, Washington, D.C.
National Weather Service. 2019. Natinal Weather Service Manual 10-950: Definitions and General Terminology. National Weather Service.
Newey, W. K., and K. D. West. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelationconsistent covariance matrix. Econometrica 55:703–708.
NOAA. 2020. US Billion-Dollar Weather & Climate Disasters 1980-2020.
Noonan, D. S., and A.-A. Sadiq. 2019. Community-scale Flood Risk Management: Effects of a Voluntary National Program on Migration and Development. Ecological Economics 157:92–99.
Noonan, D., L. Richardson, and P. Sun. 2022. Flood Zoning Policies and Residential Housing Characteristics in Texas. SSRN Electronic Journal.
Ostriker, A., and A. Russo. 2023. The Effects of Floodplain Regulation on Housing Markets. Working Paper.
Pancak, K. A., T. J. Miceli, and C. F. Sirmans. 1996. Residential Disclosure Laws: The Further Demise of Caveat Emptor. Real Estate Law Journal 24:291–332.
Peralta, A., and J. B. Scott. 2020. Does Subsidized Flood Insurance Alter Location Incentives? Evidence from the National Flood Insurance Program. Working Paper.
Pinter, N. 2005. One Step Forward, Two Steps Back on U.S. Floodplains. Science 308:207–208.
Pinter, N., F. Huthoff, J. Dierauer, J. W. F. Remo, and A. Damptz. 2016. Modeling Residual Flood Risk Behind Levees, Upper Mississippi River, USA. Environmental Science & Policy 58:131–140.
Pope, J. C. 2008. Do Seller Disclosures Affect Property Values? Buyer Information and the Hedonic Model. Land Economics 84:551–572.
Reeves, A. 2011. Political Disaster: Unilateral Powers, Electoral Incentives, and Presidential Disaster Declarations. The Journal of Politics 73:1142–1151.
Roberts, F. Y. 2006. Off-Site Conditions and Disclosure Duties: Drawing the Line at the Property Line. Brigham Young University Law Review 957:39.
Saharia, M., P.-E. Kirstetter, H. Vergara, J. J. Gourley, Y. Hong, and M. Giroud. 2017. Mapping Flash Flood Severity in the United States. Journal of Hydrometeorology 18:397–411.
Sangal, B. P. 1983. Practical Method of Estimating Peak Flow. Journal of Hydraulic Engineering 109:549–563.
Slater, L. J., and G. Villarini. 2016. Recent trends in U.S. Flood risk: Recent Trends in U.S. Flood Risk. Geophysical Research Letters 43:12, 428–12, 436.
Stern, S. 2005. Temporal Dynamics of Disclosure: The Example of Residential Real Estate Conveyancing. Utah L. Rev.:57–95.
Strobl, E. 2011. The Economic Growth Impact of Hurricanes: Evidence from U.S. Coastal Counties. Review of Economics and Statistics 93:575–589.
Tarlock, A. D. 2012. United States Flood Control Policy: The Incomplete Transition From the Illusion of Total Protection to Risk Management. Duke Environmental Law & Policy Forum 23:151–183.
Task Committee on Hydrology Handbook of Management Group D of ASCE. 1996. Floods. Hydrology Handbook. American Society of Civil Engineers Publications, Reston, VA.
Tobin, R. J., and C. Calfee. 2005, March. The National Flood Insurance Program's Mandatory Purchase Requirement: Policies, Processes, and Stakeholders. American Institutes for Research.
Tyszka, S. C. 1995. Remnants of the Doctrine of Caveat Emptor May Remain Despite Enactment of Michigan's Seller Disclosure Act. Wayne Law Review 41:1497–1530.
U.S. Department of Homeland Security. 2022. Summary of Proposed Reforms.
US Census Bureau. 2000. 2000 Census of Population and Housing Technical Documentation.

Wagner, K. R. H. 2022. Adaptation and Adverse Selection in Markets for Natural Disaster Insurance. American Economic Journal: Economic Policy 14:380–421.

Washburn, R. M. 1995. Residential Real Estate Condition Disclosure Legislation. DePaul L. Rev. 44:381–459.

Weill, J. A. 2021. Perilous Flood Risk Assessments:74.

Weinberger, A. M. 1996. Let the Buyer Be Well Informed? - Doubting the Demise of Caveat Emptor. Maryland Law Review 55:387–424.

Zervas, C. 2013. Extreme Water Levels of the United States 1893-2010. National Oceanic; Atmospheric Administration, Silver Spring, Maryland.

Zumpano, L. V., K. H. Johnson, and R. I. Anderson. 2003. Internet use and real estate brokerage market intermediation. Journal of Housing Economics 12:134–150.

A Appendix A: Data Appendix

A.1 Flood History Data

Background

A key input to flood damage function is flood size data. An ideal measure of flood size should satisfy the following four conditions. First, it should be a continuous measure that allows a non-linear relationship between flood size and damage (Burke et al. 2015, Hsiang 2016).

Second, it should be objective. For instance, the widely used EM-DAT measures flood size using economic cost or death tolls, which are directly correlated with outcome variables of interest (Felbermayr and Gröschl 2014). Another example of potentially endogenous measure is the occurrence of the Presidential Disaster Declaration (PDD) floods (Gallagher 2014), which depends on the discretion of the president and thus could reflect political interests (Reeves 2011).

Third, it should comprehensive. A few existing studies have leveraged meteorological measures to objectively measure disasters, but most of them focus on a subset of events. For instance, Deryugina (2017), Hsiang and Jina (2014), and Strobl (2011) have used physical measures of hurricane intensity while Davenport et al. (2021) leveraged precipitation data. Despite objectivity, such an approach has limits in coverage—for instance, precipitation changes alone can explain only one-third of cumulative flood damages (Davenport et al. 2021).

Lastly, since I measure flood damage at the community by year level, flood size should be measured at the same level. This is not trivial because most climate data are collected to answer physical science questions, and thus are not readily mapped into an administrative unit such as community (Carleton and Hsiang 2016).

To the best of my knowledge, no existing dataset satisfies all of these properties. In this paper, I construct an objective measure of past flood events by applying a hydrologic method to the USGS/NOAA water gauge records. This approach does not distinguish the cause of floods—hurricane, rainfall, snow melt, etc, as long as it is reflected in the water gauge level. Flood size is defined and recorded by a recurrence interval, which represents the expected number of years for a flood of given size to come back, and thus is continuous by construction. Also, by matching gauge stations to a community, I can measure flood size at the community level.

Procedure

Following the USGS guideline (England Jr et al. 2019), I implemented the following steps using USGS/NOAA water levels data from 3,505 gauge stations distributed in the 26 ever-disclosed states in the contiguous US (Appendix Figure A.1).[35]

First, I construct a site-specific flood size distribution. For this, I retrieved annual peak flow records using the R package "dataRetrieval" and fit the Log-Pearson III distribution to estimate gauge-specific parameters (Cicco et al. 2018). Importantly, as I use annual peak discharge data to fit the distribution, the quantile of the distribution has an intuitive interpretation. For instance, if a certain water level is the 95th percentile of the distribution, it means that such an event would happen with a 5 percent probability in a given year. Equivalently, such an event is called a 20-year ($\frac{1}{0.05} = 20$) flood. I keep stations with at least 10 or more annual peak observations following the USGS guideline. Also, I use annual peak data until 1990 to fix flood thresholds and make flood size comparable across different years.

Second, I convert daily water level into the recurrence interval using the fitted flood size distri-

[35] I randomly sampled 1000 sites in Appendix Figure A.1 for visibility.

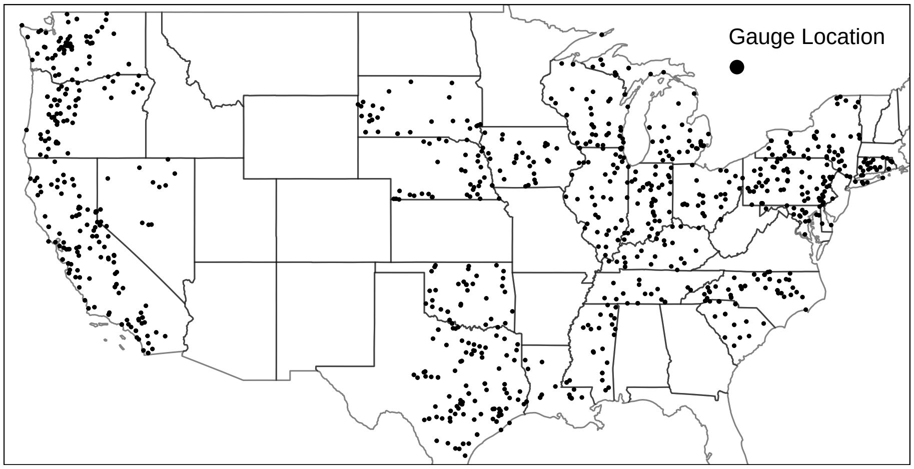


Figure A.1: The Distribution of a Sample of USGS/NOAA Gauges

Table A.1: Number of MDF Stations vs. IPF Stations in Iowa

Name	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998
N Gauges (Mean Daily Flow)	112	112	105	107	109	109	105	109	112	111
N Gauges (Maximum Daily Flow)	3	8	40	72	34	31	29	34	59	95

bution from step 1. For this, I need an instantaneous flow, because flood exposure is determined by the maximum, rather than mean, water level. The problem is that for most of the stations, the maximum daily flow (or more precisely the instantaneous peak flow which enables calculating maximum daily flow) data have too many missing values. This is problematic because, with many missing observations, flood events will be significantly under-recorded. To solve this problem, I estimate a projected instantaneous peak flow from the *mean* daily flow. In Appendix Table A.1, I report the number of water gauge stations in Iowa that have daily water level records for at least 80 percent of the days (i.e., 292 days or more) for a given year. It can be easily seen that there can be an order of magnitude difference in the number of stations that have mean versus maximum daily water records.

To estimate the daily maximum water level from daily mean water level, I use the Fuller method (Fuller 1913). Specifically, for a given geographical unit, I estimate Fuller coefficients by regressing instantaneous peak flow (Q_{it}^{IPF}) for site i in time t on mean daily flow (Q_{it}^{MDF}) and the size of the drainage area (A) as equation (7) (Fuller 1913).[36] I use three different levels of geographic units, namely state, HUC4, and HUC2 and separately estimate Fuller coefficients.[37] Using the estimated coefficients, I calculate projected instantaneous peak flow, and compare that with the actual instantaneous peak flow to pick the geographic unit that minimizes the prediction error.[38]

[36] I also did conversion following Sangal (1983), but the error between actual and the estimated IPF was much smaller with Fuller (1913).

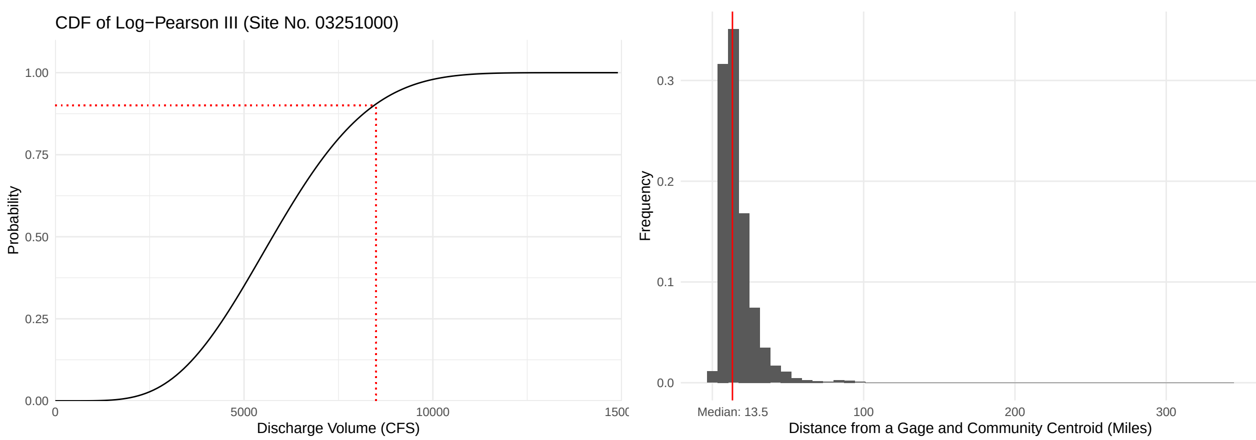
[37] A watershed is uniquely identified by a hydrologic unit code (HUC). There are six levels in the hierarchy, and HUC2 (regions) and HUC4 (sub-regions) are two highest levels. There are total of 18 and 202 HUC2s and HUC4s in the contiguous US (Maimone and Adams 2023).

[38] Practically, I apply the following hierarchy among state, HUC4, and HUC2 models: (1) When a site has the best match (meaning that a site has both daily mean flow and instantaneous flow records), I use it. (2) If a site does not have site-specific match (meaning that this site did not have instantaneous flow records), I use prioritized HUC4, HUC2, and State, because HUC4 had the least overall prediction error. I also remove the cases where a site does not have drain area (and thus Fuller coefficients cannot be estimated).

$$Q_{it}^{IPF} = Q_{it}^{MDF}(1 + \alpha A^{\beta}) \tag{7}$$

Now, by converting the estimated instantaneous peak flow to the quantile of the estimated Log-Pearson III CDF from step 1, I identify each day's flood size.

Finally, to translate gauge-level flood events to the community-level floods, I match each community to the three nearest gauges based on the distance between a centroid of community and gauge station. Then, I calculate the average flood size for a community using the inverse distance as a weight. Appendix Figure A.2 (b) presents the distribution of the average distance between gauges and community centroid. Over 90 percent of them are within 20 miles with a median distance of 13.5 miles.



(a) Flood Frequency Analysis

(b) Distance between Gauge and Community Centroids

Figure A.2: Flood Frequency Analysis and Gauge Matching. Plot (a) is an example of flood frequency analysis. The black solid line represents the CDF of the fitted Log-Pearson III distribution from the USGS site 03251000. If a daily discharge volume is 8,500 CFS, it corresponds to the 90th quantile or a 10-year flood. Plot (b) presents the distribution of the average distance between a gauge and community centroid. Over 90% of them are within 20 miles with the median distance 13.5 miles.

Appendix Figure A.2 (a) illustrates steps 1 and 2 described above. The black solid line is the fitted Log-Pearson III CDF from the USGS site 03251000. To fit the distribution, I use the annual peak flow data from 1947 to 1990 to calculate the mean, standard deviation, and skewness parameters. Now suppose that on a given date, the daily discharge volume is 8,500 CFS. As it corresponds to the 90th percentile of the CDF, it can be concluded that there was a 10-year flood on that day.

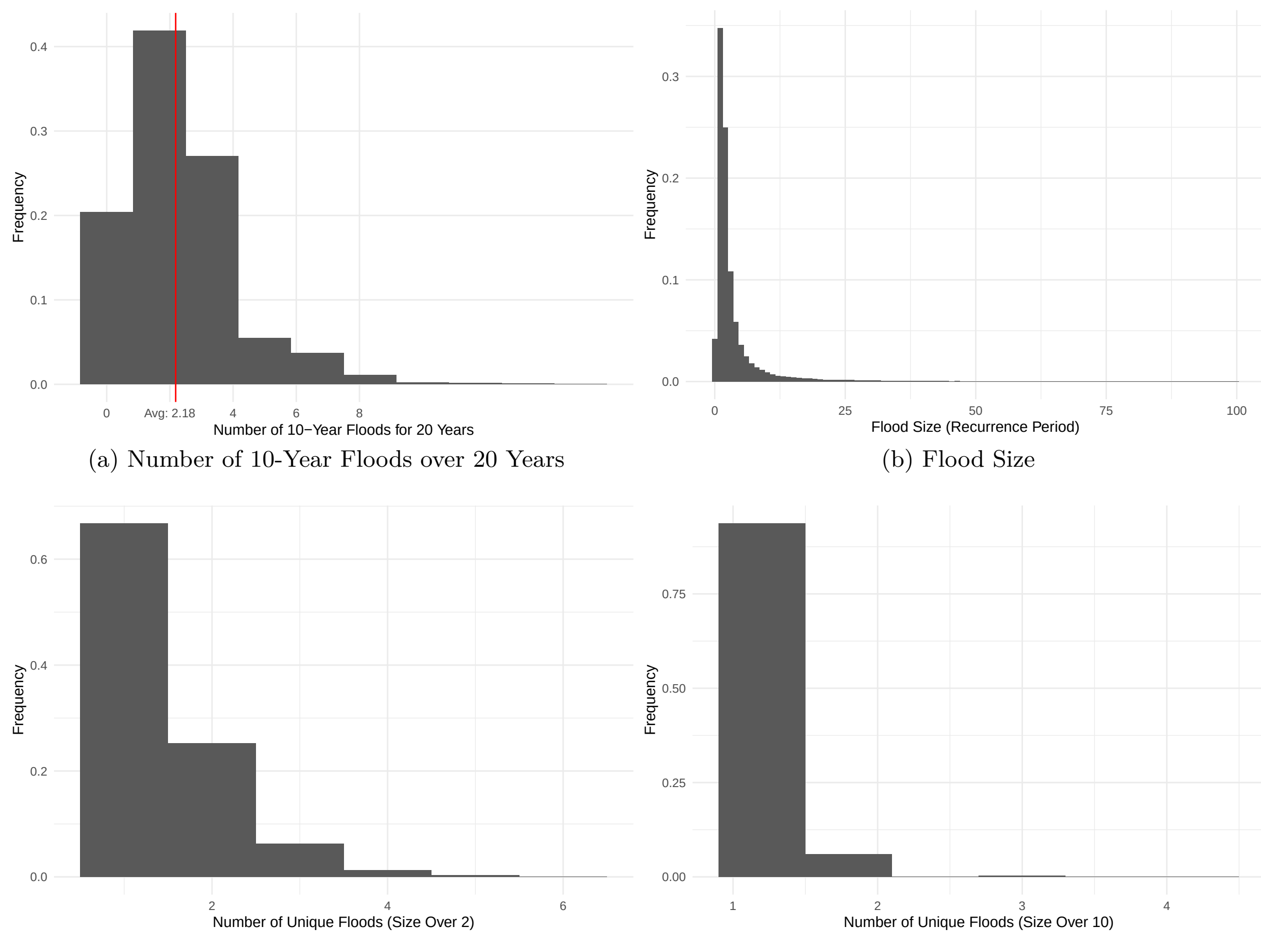
Note, because the USGS gauge stations rarely cover coastal areas, I add 45 additional NOAA sites to the gauge station data. Zervas (2013) documents the flood threshold for the all NOAA sites by fitting GEV distribution, so I adopt them directly. NOAA water level data are retrieved using the R package "Rnoaa" (Edmund et al. 2014).

Unified Flash Flood Database

The Unified Flash Flood Database (Gourley et al. 2013) is USGS-gauge record based dataset constructed following a similar procedure outlined. It is a comprehensive and objective measure of flood events that can present the overall trend of flood events for the contiguous US, which overcomes many limitations of the existing data. However, I decided not to use this database because the data are constructed based on the instantaneous peak flow. As Appendix Table A.1 shows, relying solely on the instantaneous peak flow can substantially under report flood events due to missing water level records.

Validation and Summary Statistics

To validate the flood history data, I check the number of the average 10-year flood events over a 20-year period for the 8,194 communities. These communities are from the 26 ever-disclosed states that are on the Q3 map. By definition, a 10-year flood happens twice in a 20-year period on average. Figure A.3 (a) shows that most communities had 1 or 2 10-year floods over the 20 years whereas the average number of 10-year floods is 2.18. While this is slightly higher than 2, it is plausible given that I use the annual peak flow data until 1990. Fixing flood threshold is necessary to compare floods across different times (namely, a 10-year flood should have the same magnitude whether it is in 1990 or 2000). Although this approach can be problematic as the period in consideration gets longer, it should not be a major problem for this paper as the longest sample period is 20 years.



(a) Number of 10-Year Floods over 20 Years (b) Flood Size

(c) N of Unique Floods (size over 2) by Community-year (d) N of Unique Floods (size over 10) by Community-year

Figure A.3: Flood Data Summary Information. Panel (a) shows that most communities had 1 or 2 10-year floods over the 20 years and the average number of 10-year flood is 2.18. Panel (b) shows the distribution of flood event size (i.e., recurrence interval), where flood size is truncated at 100 for readability. Panel (c) illustrates the number of unique floods (size over 2) for community-year. Panel (d) repeats panel (c) for floods with size over 10.

Figure A.3 (b) shows the distribution of flood size (i.e., recurrence interval), where flood size is truncated at 100 for readability. As is well documented in the literature, the histogram follows a log-normal distribution, and the frequency decreases as an inverse power function of the flood size (Jackson 2013).

In panel (c), I plot the number of unique flood events for each community-year, conditional on having an event with flood size between 2 and 50. The histogram shows that about 70 percent of the community-years have exactly one event. This alleviates a concern over measuring flood exposure as the maximum flood size for a given year. More importantly, when I limit attention to floods with size over 10, which incurs disproportionately large damage, over 90 percent of the community-year pairs have only one such event.

Table A.2: Comparing the Estimated Flood Size Thresholds with the NWS Threshold

	2 Year Flood	10 Year Flood	50 Year Flood	100 Year Flood
Minor	0.778***	1.285***	1.74***	1.944***
	(0.052)	(0.071)	(0.102)	(0.124)
Moderate	0.594***	0.994***	1.36***	1.526***
	(0.042)	(0.06)	(0.085)	(0.103)
Major	0.45***	0.771***	1.081***	1.226***
	(0.034)	(0.043)	(0.051)	(0.06)

Note:
Note: The entries report the results from 12 separate regressions where each column represents four different dependent variables and each row represents three different regressors. Standard errors are clustered at the gauge level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

To better contextualize the recurrence interval based flood size, in Appendix Table A.2, I compare flood size with the gauge-specific NWS thresholds for minor, moderate, and major floods.[39] Specifically, I estimate equation (8) where Q_{ik} is the estimated flood threshold for site i for flood size k where $k \in \{2, 10, 50, 100\}$. NWS_{ij} is flood thresholds from the NWS for site i for flood severity j where $j \in$ {minor, moderate, major}.

$$Q_{ik} = \beta NWS_{ij} + \epsilon_{ijk} \tag{8}$$

β is the coefficient of interest which illustrates how comparable two thresholds are. Namely, the closer β is to 1, the more comparable two thresholds are. For this analysis, I use 2,093 sites that have both recurrence interval based flood size and the NWS flood thresholds. Appendix Table A.2 reports the estimated β for 12 separate regressions and provides useful insights. First, a minor flood from the NWS is comparable to a flood of size between 2 and 10. To see this, observe that when minor threshold increases by 1 unit, a 2-year flood threshold is increasing by only 0.78 units. Conversely, when minor threshold increases by 1 unit, a 10-year flood threshold is increasing by 1.29 units. Second, a 10-year flood threshold is tightly comparable to a moderate flood threshold ($\beta = 0.99$). Similarly, a 50-year flood closely matches with a flood with major impact ($\beta = 1.08$). Note, a 100-year flood threshold is 23 percent higher than a major flood threshold, which is plausible given that a 50-year flood threshold is comparable to the major category.

[39] NWS defines each flood category as the following (National Weather Service 2019). Minor: minimal or no property damage, but possibly some public threat (e.g., inundation of roads). Moderate: some inundation of structures and roads near stream, evacuations of people and/or transfer of property to higher elevations. Major:extensive inundation of structures and roads, significant evacuations of people and/or transfer of property to higher elevations.

A.2 Validation of Key Dependent Variables

Table 3.1 shows that key dependent variables in this paper have prevalence of zeros. These statistics are consistent with findings from external sources.

Block population. Bureau of the Census (1994) reports that a substantial number of blocks have zero population, with state-level proportions ranging from 14 percent (RI) to 65 percent (WY), and a median value of 31 percent (WA). In my sample, the numbers are slightly different at 17 percent for RI and 26 percent for WA (WY is a non-disclosure state). A minor discrepancy is not surprising given that blocks not included in the digitized flood map are excluded from the analysis.

Flood insurance counts. There is no prior work that has documented the fraction of communities with zero insurance policies. However, when I compare the total number of insurance policies by state in my sample with other studies, I find them highly congruent. For instance, in my sample, Louisiana had 504,641 policies as of 2007, a figure closely matching the documented 502,085 flood insurance policies as of December 2007 in Michel-Kerjan and Kousky (2010). Other disclosing states listed in Michel-Kerjan and Kousky (2010) Table 1 are also well matched: CA (258,808 vs. 266,171), NC (123,949 vs. 133,955), NY (141,525 vs. 144,253), SC (190,997 vs 197,334), and TX (508,348 vs. 666,920) where the first number is from my sample and the second number is from Michel-Kerjan and Kousky (2010). Note, for TX, there is a noticeable gap primarily because Harris County is not in my sample (the county is not included in the digitized flood map described in Section 3).

Flood damage. Similar to the flood insurance policy counts, no prior studies have cataloged the fraction of community-years with zero flood damage. However, a back-of-the-envelop calculation suggests that this statistic is in line with existing studies. For that, I take the average probability (1.45 percent) of filing a claim per policy over 1980–2012 from Kousky and Michel-Kerjan (2015) and multiply it with the number of flood insurance policies by the community in my sample. The result reveals that 17 percent of communities are predicted to have more than one claim in a given year (i.e., 83 percent of community-year observations are predicted to have zero claims). Note, while 83 percent is substantially lower than 95 percent as discussed in Section 3, this is a direct consequence of sample restriction: as I discuss in detail in Section 5.1, I remove floods with size 50 or above from my analysis for various economic and statistical reasons. When I undertake the same calculation without imposing these sample restrictions, I find that 86 percent of community-year observations have zero claims, a figure consistent with the 83 percent calculated based on Kousky and Michel-Kerjan (2015).

B Appendix B: Determinants of Home Seller Disclosure Requirement Adoption

Table B.1: State Characteristics in 1990

	Ever/Early		Never/Late		Difference	
Variables	Mean	SE	Mean	SE	Mean	P.Value
Panel A: Ever vs. Never States						
Population (millions)	6.57	1.31	3.43	0.651	3.143	0.048
Median Age	33.04	0.204	32.82	0.409	0.22	0.616
(%) White	0.827	0.019	0.879	0.018	-0.053	0.051
(%) BA	0.121	0.005	0.129	0.006	-0.007	0.324
Unemployment Rate	0.06	0.003	0.061	0.002	-0.001	0.773
GDP (billions)	152	34.38	74	14.95	78	0.057
N Housing Units (millions)	2.66	0.506	1.47	0.291	1.187	0.059
(%) Vacancy	0.095	0.005	0.132	0.008	-0.037	0
Democratic Party Vote Share	0.455	0.01	0.425	0.012	0.03	0.06
Average Flood Damage per Housing Unit	3.86	1.99	0.931	0.5	2.925	0.194
Flood Size	5.92	0.82	3.29	0.725	2.635	0.022
(%) in SFHA	0.16	0.012	0.132	0.013	0.028	0.117
Panel A: Early vs. Late States						
Population (millions)	5.53	1.29	7.8	2.42	-2.274	0.397
Median Age	33.07	0.286	33	0.302	0.071	0.865
(%) White	0.842	0.026	0.808	0.027	0.034	0.374
(%) BA	0.119	0.006	0.124	0.008	-0.005	0.592
Unemployment Rate	0.061	0.004	0.06	0.004	0.001	0.89
GDP (billions)	119	29.72	191	66	-72	0.306
N Housing Units (millions)	2.25	0.527	3.12	0.917	-0.87	0.402
(%) Vacancy	0.095	0.007	0.096	0.007	-0.001	0.908
Democratic Party Vote Share	0.47	0.013	0.438	0.014	0.031	0.118
Average Flood Damage per Housing Unit	3.82	1.99	3.9	3.75	-0.088	0.983
Flood Size	5.71	1.05	6.17	1.34	-0.465	0.784
(%) in SFHA	0.157	0.01	0.163	0.023	-0.006	0.788

Note:
This table compares key characteristics of ever-disclosed vs. never-disclosed (Panel A) and early-disclosed vs. late-disclosed (Panel B) states. All variables are as of 1990 except for the Democratic party vote share variable, which comes from 1988 presidential election. The last two columns show mean differences with p-values.

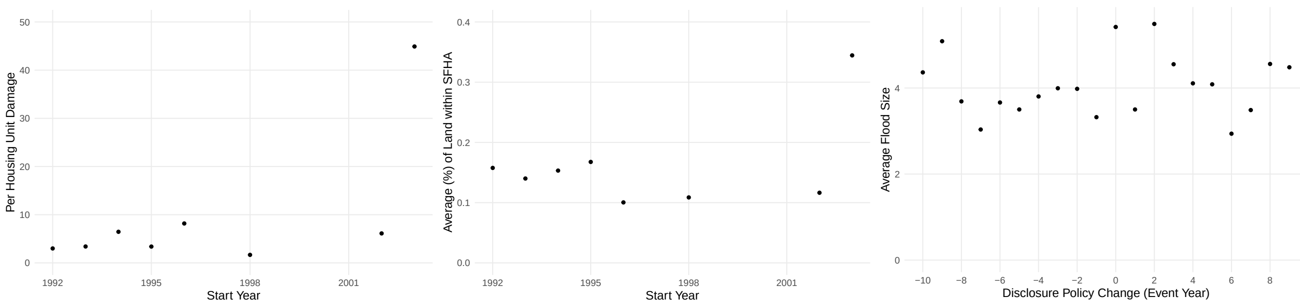
Given the background of disclosure requirement, the policy is much more likely to be adopted in states with higher needs or attention for consumer protection. In Table B.1, I compare never vs. ever (Panel A) and early vs. late (Panel B) adopted states on various characteristics.[40]

Results in Panel A shows that the 26 ever-adopted states tend to have statistically different demographic, economic, and political characteristics in comparison to the 22 states that never implemented such a policy. For instance, ever-disclosed states have a larger population, housing units, GDP, higher fraction of White population, and a lower vacancy rate. They're also more likely to

[40]The vote share for Democratic party comes from the 1988 Presidential election results compiled by Amlani and Algara (2021). State level GDP is acquired from BEA. Flood related variables are constructed as detailed in Section 3. The rest of the variables come from the 1990 decennial census.

support the Democratic party. Further, they're tend to experience larger floods.

Interestingly, no such difference is observed when I compare early—14 states that have implemented the policy by 1994—vs. late—12 states implemented after 1994—adopters within the 26 ever-disclosed states. In Panel B of Table B.1, I repeat the same exercise and find that not a single variable is statistically significantly different between these two groups, to a large part because the mean difference is much smaller in Panel B. For instance, mean difference for the vacancy rate and flood size in Panel B is less than 20% of the value in Panel A.[41] These observations suggest that the never-treated states may have meaningfully different trajectory from the ever-treated states. Thus, I use not-yet-treated states as a control group when employing panel regression models.



(a) Timing and Flood Damage (b) Timing and (%) of Land in SFHA (c) Timing (in Event) and Flood Size

Figure B.1: Correlation Between Disclosure Timing and Flood Profiles. These figures plot the disclosure policy timing against (a) past flood damage and (b) ex-ante flood risk profile. Panel (c) plots the average flood size in event time. Values on the y-axis is pooled across all states with the same treatment or event year.

While flood related variables in Panel B of Table B.1 indicates that the implementation timing is unlikely to be correlated with each state's underlying flood-risk profile, in Appendix Figure B.1, I provide additional evidence. In particular, I plot the relationship between disclosure year and (a) the average flood damage per housing unit and (b) the average proportion of land area inside of the SFHA. If the policy timing is correlated with underlying flood risk, we would expect to see a higher risk level for early adopters. However, both flood damage and the SFHA ratio are uncorrelated with the implementation year.[42] In contrast to panels (a)–(b) where the x-axis is in calendar years, panel (c) shows the average flood size in event time, namely time relative to passage of the state's disclosure legislation. If the decision to adopt a disclosure policy were a direct response to devastating flood events, the average flood size would be larger for event years right before 0. The plot also shows that flood size is essentially uncorrelated with policy adoption—if anything, flood size seems to be smaller in event years -1 and -2, which again suggests that the policy implementation is not driven by prior flooding. These figures reassure plaubible exogeneity of the disclosure requirement adoption timing.

[41] Democratic party vote share is larger for early-disclosed states. This is consistent with earlier findings that attention on consumer protection is much larger under Democratic regimes (Mullenix 2019).

[42] Spikes in Figure B.1 (a) and (b) are due to Louisiana, which has substantially higher flood damage per housing unit and a higher fraction of land area in SFHAs compared to other states.

C Appendix C: Disclosure Requirement and Housing Price

Housing price change to the disclosure policy is of interest in its own right, but it is also a first pass at testing the efficacy of the disclosure policy. That is, by comparing the estimated effect of flood risk information (through the disclosure requirement) to estimates from prior studies, I can indirectly test whether the disclosure requirement was effectively raising homebuyers' flood risk awareness.

For housing prices, I use the Zillow Transaction and Assessment Database (ZTRAX).[43] It documents transaction dates, sales prices, and housing characteristics such as type (e.g., single house, condominium, etc.), exact longitude and latitude, year built, and the number of bedrooms.[44]

A combination of the different policy implementation timing and the differential treatment of properties located in and out of an SFHA allows me to employ a triple difference design using the stacked DDD approach. I use not-yet-treated states as clean controls and exploit the policy implementation timing among the ever-treated states. Equation (9) estimates the impact of the disclosure policy on the housing price.

$$log(Price_{ijmstd}) = \beta T_{ijmstd} + \theta_{mjhld} + \mu_{jtd} + \lambda_{mtd} + \epsilon_{ijmstd} \tag{9}$$

$Price_{ijmstd}$ is the housing price for a property i with SFHA status j in community m in state s at time t in stack d and T_{ijmstd} is the treatment status dummy, which takes on a value of 1 when SFHA = Post = Disclosure = 1 where SFHA is a dummy for the SFHA status, Post is a dummy for the post-disclosure period, and Disclosure is a dummy for the treatment group assignment.[45]

I also include a complete set of two-way fixed effects μ_{jtd}: SFHA $\times$ Time $\times$ Stack, λ_{mtd}: Community $\times$ Time $\times$ Stack, and θ_{mjhld}: Community $\times$ SFHA $\times$ Building Age $\times$ Number of Beds $\times$ Stack to estimate β. These fixed effects allow me to estimate the policy effect using the sales price variation before and after the disclosure policy, inside and outside of the SFHA while controlling for the community by SFHA specific property characteristics. Further, these fixed effects are interacted with the stack d, to ensure that comparisons are made within each stack. For building age h, I group construction years into 10-year bins (e.g., 2000-2009, 1990-1999, etc.) and for the number of bedrooms l, I group them into 1-3, 4-6, 7-10, and 10-or-more bedrooms bins. The identification comes from plausibly exogenous disclosure policy change timings after conditioning on the set of fixed effects.

In Table C.1 column (1), I report the estimated coefficients of equation (9) to find that the disclosure requirement reduces the price of the properties in the SFHA by 4.5 percent in comparison to those outside of the SFHA. To put this number in context, I multiply the estimate from column (1) to the average price of properties located in the SFHA in the pre-disclosure period ($327,171) to produce the estimated reduction in the housing price of $14,598. Importantly, community by year level potential confounders such as flood exposure or flood insurance premiums are controlled by the community by year fixed effects in this specification. In column (2), I show that the estimate in column (1) is robust to removing properties in communities that have experienced a flood map update over the sample period. The estimate in column (2) is essentially identical, suggesting that map updates are uncorrelated with the disclosure policy implementation.

Figure C.1 presents an event study style graph, measuring the policy effect over event time. $\hat{\beta}_k$ in

[43] I thank Eyal Frank for his generous help with data access.

[44] I apply the following sample restrictions. First, I drop observations without longitude and latitude information. Second, I keep only single-family houses in the sample, reflecting the fact that the disclosure requirement in many states is applied only to one to four dwelling units. Third, I restricted the transaction price to be between $10,000 and $100,000,000 in nominal dollars.

[45] For details on creating data stacks, see Section 4.1.

Table C.1: Effect of Disclosure Requirement on the Housing Prices

	(1)	(2)
SFHA × Disclosure × Post	$-.045^{***}$	$-.046^{**}$
	(.015)	(.018)
Sample	Entire Communities	No-Revision Communities
Stack × Community × Year FE	X	X
Stack × Community × Year FE	X	X
Stack × Community × SFHA × Year Built × N Beds FE	X	X
Num. obs.	6249070	5931016

Note: Column (1) shows $\hat{\beta}$ from equation (9) from the 26 ever-disclosed states. In column (2), I remove observations from communities that have experienced Flood Insurance Rate Map, or an official flood map, update during the sample period. The dependent variable is log(sales price). All standard errors are clustered at the state level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

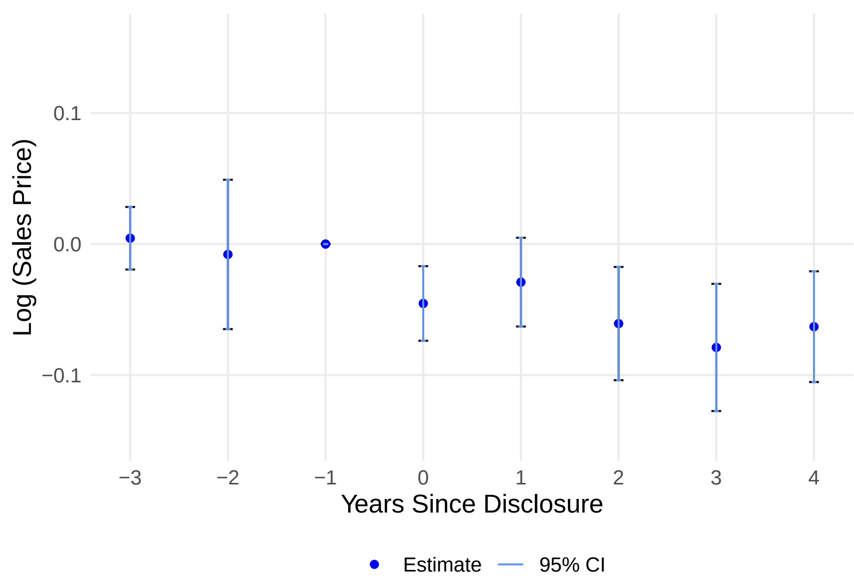


Figure C.1: The Effect of the Disclosure Requirement on Housing Price. These figures plot the coefficients of interaction terms between the SFHA status and disclosure policy dummies in event time. The dependent variable is the log of housing price. Standard errors are clustered on state.

the pre-disclosure periods are almost zero, satisfying the parallel trend assumption. Since the first year of the policy change, the price of affected properties has fallen by about 4 percent. The effect is persistent up until five years after the policy implementation.

D Appendix D: Additional Tables and Figures

Table D.1: Balance Table (Tracts by the SFHA Status)

	No SFHA		With SFHA		Difference	
Variables	Mean	SE	Mean	SE	Mean	t-stat
N Housing Unit	1377	6.31	1408	4.2	32	1.2
N Home Age Below 6	92	1.58	160	1.14	68	4.64
N Home Age Above 42	558	5.09	343	2.44	-216	-3.7
(%) Home Age Below 6	0.078	0.0012	0.1296	9e-04	0.052	4.1
(%) Home Age Above 42	0.3961	0.0028	0.2337	0.0014	-0.162	-3.8

Note:
This table compares the proportion of older and newer housing stocks in census tracts with and without SFHAs. The last two columns show mean differences with t-statistics. Standard errors are clustered at the state level.

Back to 2.1.

Table D.2: Effect of Discosure Requirement on Household Responses (Placebo States)

	Prob. of Any Population	Log Population	Vacancy Rate
	(1)	(2)	(3)
SFHA × Post	.004	.036	.004
	(.004)	(.061)	(.012)
Avg D.V. (Within BW)	0.595		0.096
Bandwidth	395	438	400
Num. obs.	311076	198685	186471

Note: This table is produced from equation (1). Columns (1)–(3) are estimated using the decennial census block-level data in 1990, 2000, 2010, and 2020. Standard errors are clustered at the state level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 4.2.

Table D.3: Effect of Discosure Requirement on Population and Vacancy Rate (Doughnut Specification)

	Prob. of Any Population	Log Population	Vacancy Rate	Prob. of Any Population	Log Population	Vacancy Rate
	(1)	(2)	(3)	(4)	(5)	(6)
SFHA × Post	$-.011^{**}$	$-.079^{**}$	$.013^{***}$	$-.007$	$-.080^{**}$	$.014^{**}$
	(.004)	(.031)	(.004)	(.005)	(.033)	(.006)
Avg D.V. (Within BW)	0.692		0.093	0.704		0.092
Doughnut Size	20	20	20	40	40	40
Num. obs.	1228215	1765788	1550895	984969	1609485	1395789

Note: This table is produced from equation (1) after excluding observations closest to the SFHA border. In columns (1)–(3), doughnut sizes are 20 meters and in columns (4)–(6) doughnut sizes are 40 meters. Standard errors are clustered at the state level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 4.2.

Table D.4: Effect of Discosure Requirement on Net Population Flow (Excluding Blocks with Map Revision)

	Prob. of Any Population	Log Population		Vacancy Rate
	(1)	(2)	(3)	(4)
SFHA × Post	$-.011^{***}$	$-.071^{**}$		$.014^{***}$
	(.003)	(.031)		(.004)
High SFHA × Disclosure × Post			$-.009$	
			(.006)	
Avg D.V.	0.67			0.098
Year × Stack FE			X	
Community × Stack FE			X	
Bandwidth	138	301		262
Num. obs.	1313619	1682626	481467	1495039

Note: Estimates are based on equation (1) and (2) after removing geographic units that have experienced flood map update. Columns (1)–(2) and (4) are estimated using the decennial census block-level data in 1990, 2000, 2010, and 2020. Columns (3) is estimated using community-level population data. Standard errors are clustered at the state level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 4.2.

Table D.5: Effect of Discosure Requirement on Flood Insurance Take-Up (Excluding Communities with Map Revision)

	Prob. of Any Insurance	Log Insurance Per Housing Unit
	(1)	(2)
High SFHA × Disclosure × Post	.003	−.023
	(.008)	(.030)
Avg D.V.	0.819	
State × Year × Stack FE	X	X
Community × Stack FE	X	X
Num. obs.	390382	319639

Note: This table is produced from equation (2) using community-level National Flood Insurance Program data after removing communities that have experienced map update during the sample period. Standard errors are clustered at the state level. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 4.2.

Table D.6: Effect of Disclosure Requirement on Flood Damage

	Prob. of Any Damage Per Housing Unit			Log Damage Per Housing Unit
	(1)	(2)	(3)	(4)
Flood Size 2-10	.020***	.022***	.019**	.223***
	(.007)	(.008)	(.008)	(.054)
Flood Size 10-20	.054***	.071***	.039***	1.140***
	(.012)	(.016)	(.010)	(.104)
Flood Size 20-30	.078***	.108***	.048***	2.095***
	(.023)	(.032)	(.013)	(.419)
Flood Size 30-40	.068***	.082***	.052*	1.584***
	(.024)	(.031)	(.027)	(.338)
Flood Size 40-50	.096**	.144***	.026	1.850***
	(.042)	(.046)	(.040)	(.254)
Disclosure × Size 2-10	.028***	.039***	.014***	.087
	(.009)	(.015)	(.005)	(.164)
Disclosure × Size 10-20	.094***	.117***	.065***	−.003
	(.017)	(.026)	(.009)	(.086)
Disclosure × Size 20-30	.114***	.144***	.075***	−.164
	(.018)	(.028)	(.017)	(.138)
Disclosure × Size 30-40	.106***	.112***	.100***	.077
	(.030)	(.033)	(.030)	(.156)
Disclosure × Size 40-50	.107**	.115**	.100***	−.048
	(.044)	(.054)	(.031)	(.406)
Post × Size 2-10	.028**	.035***	.018*	.452**
	(.012)	(.012)	(.010)	(.184)
Post × Size 10-20	.094***	.096***	.088***	.086
	(.030)	(.034)	(.030)	(.076)
Post × Size 20-30	.105***	.117***	.093***	−.349
	(.028)	(.032)	(.032)	(.326)
Post × Size 30-40	.191***	.246***	.140*	.668*
	(.054)	(.044)	(.078)	(.361)
Post × Size 40-50	.239***	.356***	.135***	.629**
	(.040)	(.042)	(.046)	(.312)
Post × Disclosure × Size 2-10	−.039*	−.056*	−.021	−.063
	(.023)	(.030)	(.015)	(.325)
Post × Disclosure × Size 10-20	−.072*	−.086*	−.051*	.189
	(.039)	(.050)	(.029)	(.189)
Post × Disclosure × Size 20-30	−.080***	−.131***	−.018	.170
	(.029)	(.038)	(.031)	(.562)
Post × Disclosure × Size 30-40	−.141*	−.172**	−.111	−.360
	(.073)	(.072)	(.082)	(.442)
Post × Disclosure × Size 40-50	−.197***	−.339***	−.054	−.425
	(.055)	(.061)	(.068)	(.540)
Sample	All	High SFHA	Low SFHA	Damage > 0
Year × Stack FE	X	X	X	X
Community × Stack FE	X	X	X	X
Num. obs.	505383	242458	262925	22100

Note: This table shows the full sets of coefficients for Table 5.1. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 5.2.

Table D.7: Effect of Disclosure Requirement on Flood Damage (Placebo States)

	Prob. of Any Damage		
	(1)	(2)	(3)
Post × Disclosure (Size 2-30)	.007	.003	.010
	(.006)	(.006)	(.008)
Post × Disclosure (Size 30-50)	.045	−.046	.175
	(.138)	(.152)	(.132)
Sample	All	High SFHA	Low SFHA
Year × Stack FE	X	X	X
Community × Stack FE	X	X	X
Num. obs.	29626	14864	14762

Note: This table repeats Table 5.1 using the placebo states. The dependent variables in columns (1) to (3) are the probability of having any flood damage per housing unit. Column (1) is based on the entire set of communities while in columns (2) and (3), I repeat (1) using the subsample of communities with an above median SFHA ratio and below median SFHA ratio. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference for columns (1)–(3). $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 5.2.

Table D.8: Effect of Disclosure Requirement on Flood Damage (Excluding Communities with Map Revision)

	Prob. of Any Damage Per Housing Unit			Log Damage Per Housing Unit
	(1)	(2)	(3)	(4)
Post × Disclosure (Size 2-10)	$-.038^{*}$	$-.054^{*}$	$-.020$	.089
	(.023)	(.030)	(.015)	(.256)
Post × Disclosure (Size 10-20)	$-.060$	$-.070$	$-.043^{*}$	.249
	(.037)	(.051)	(.024)	(.214)
Post × Disclosure (Size 20-30)	$-.081^{**}$	$-.132^{***}$	$-.015$	.034
	(.037)	(.042)	(.037)	(.590)
Post × Disclosure (Size 30-40)	$-.139^{*}$	$-.151^{**}$	$-.124$	$-.596$
	(.078)	(.062)	(.102)	(.430)
Post × Disclosure (Size 40-50)	$-.219^{***}$	$-.352^{***}$	$-.072$	$-.377$
	(.068)	(.065)	(.083)	(.562)
Annual Effect	-0.023**	-0.032***	-0.012	0.009
	(0.01)	(0.011)	(0.009)	(0.053)
Sample	All	High SFHA	Low SFHA	Damage > 0
Year × Stack FE	X	X	X	X
Community × Stack FE	X	X	X	X
Num. obs.	487704	233225	254479	20619

Note: This table repeats Table 5.1 after removing communities that have experienced map updates during the sample period. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 5.2.

Table D.9: Effect of Disclosure Requirement on Flood Damage (State Level Clustering)

	Prob. of Any Damage Per Housing Unit			Log Damage Per Housing Unit
	(1)	(2)	(3)	(4)
Post × Disclosure (Size 2-10)	$-.039^{*}$	$-.056^{*}$	$-.021$	$-.063$
	(.020)	(.029)	(.013)	(.354)
Post × Disclosure (Size 10-20)	$-.072$	$-.086$	$-.051$	.189
	(.045)	(.070)	(.031)	(.277)
Post × Disclosure (Size 20-30)	$-.080^{*}$	$-.131^{*}$	$-.018$	.170
	(.044)	(.066)	(.037)	(.608)
Post × Disclosure (Size 30-40)	$-.141$	$-.172^{*}$	$-.111$	$-.360$
	(.090)	(.095)	(.096)	(.624)
Post × Disclosure (Size 40-50)	$-.197^{***}$	$-.339^{***}$	$-.054$	$-.425$
	(.066)	(.084)	(.053)	(.582)
Annual Effect	-0.023**	-0.034**	-0.012	-0.012
	(0.01)	(0.014)	(0.008)	(0.072)
Sample	All	High SFHA	Low SFHA	Damage > 0
Year × Stack FE	X	X	X	X
Community × Stack FE	X	X	X	X
Num. obs.	505383	242458	262925	22100

Note: This table repeats Table 5.1 with state level clustering. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 5.2.

Illinois REALTORS®
RESIDENTIAL REAL PROPERTY DISCLOSURE REPORT
(765 ILCS 77/35)

NOTICE: THE PURPOSE OF THIS REPORT IS TO PROVIDE PROSPECTIVE BUYERS WITH INFORMATION ABOUT MATERIAL DEFECTS IN THE RESIDENTIAL REAL PROPERTY. THIS REPORT DOES NOT LIMIT THE PARTIES' RIGHT TO CONTRACT FOR THE SALE OF RESIDENTIAL REAL PROPERTY IN "AS IS" CONDITION. UNDER COMMON LAW, SELLERS WHO DISCLOSE MATERIAL DEFECTS MAY BE UNDER A CONTINUING OBLIGATION TO ADVISE THE PROSPECTIVE BUYERS ABOUT THE CONDITION OF THE RESIDENTIAL REAL PROPERTY EVEN AFTER THE REPORT IS DELIVERED TO THE PROSPECTIVE BUYER. COMPLETION OF THIS REPORT BY THE SELLER CREATES LEGAL OBLIGATIONS ON THE SELLER; THEREFORE SELLER MAY WISH TO CONSULT AN ATTORNEY PRIOR TO COMPLETION OF THIS REPORT.

Property Address: ____

City, State & Zip Code: ____

Seller's Name: ____

This Report is a disclosure of certain conditions of the residential real property listed above in compliance with the Residential Real Property Disclosure Act. This information is provided as of ____________________________, 20___, and does not reflect any changes made or occurring after that date or information that becomes known to the seller after that date. The disclosures herein shall not be deemed warranties of any kind by the seller or any person representing any party in this transaction.

In this form, "am aware" means to have actual notice or actual knowledge without any specific investigation or inquiry. In this form, a "material defect" means a condition that would have a substantial adverse effect on the value of the residential real property or that would significantly impair the health or safety of future occupants of the residential real property unless the seller reasonably believes that the condition has been corrected.

The seller discloses the following information with the knowledge that even though the statements herein are not deemed to be warranties, prospective buyers may choose to rely on this information in deciding whether or not and on what terms to purchase the residential real property.

The seller represents that to the best of his or her actual knowledge, the following statements have been accurately noted as "yes" (correct), "no" (incorrect), or "not applicable" to the property being sold. If the seller indicates that the response to any statement, except number 1, is yes or not applicable, the seller shall provide an explanation, in the additional information area of this form.

	YES	NO	N/A	
1.	___	___	___	Seller has occupied the property within the last 12 months. (No explanation is needed.)
2.	___	___	___	I am aware of flooding or recurring leakage problems in the crawl space or basement.
3.	___	___	___	I am aware that the property is located in a flood plain or that I currently have flood hazard insurance on the property.
4.	___	___	___	I am aware of material defects in the basement or foundation (including cracks and bulges).
5.	___	___	___	I am aware of leaks or material defects in the roof, ceilings, or chimney.
6.	___	___	___	I am aware of material defects in the walls, windows, doors, or floors.
7.	___	___	___	I am aware of material defects in the electrical system.
8.	___	___	___	I am aware of material defects in the plumbing system (includes such things as water heater, sump pump, water treatment system, sprinkler system, and swimming pool).
9.	___	___	___	I am aware of material defects in the well or well equipment.
10.	___	___	___	I am aware of unsafe conditions in the drinking water.
11.	___	___	___	I am aware of material defects in the heating, air conditioning, or ventilating systems.
12.	___	___	___	I am aware of material defects in the fireplace or wood burning stove.
13.	___	___	___	I am aware of material defects in the septic, sanitary sewer, or other disposal system.
14.	___	___	___	I am aware of unsafe concentrations of radon on the premises.
15.	___	___	___	I am aware of unsafe concentrations of or unsafe conditions relating to asbestos on the premises.
16.	___	___	___	I am aware of unsafe concentrations of or unsafe conditions relating to lead paint, lead water pipes, lead plumbing pipes or lead in the soil on the premises.
17.	___	___	___	I am aware of mine subsidence, underground pits, settlement, sliding, upheaval, or other earth stability defects on the premises.
18.	___	___	___	I am aware of current infestations of termites or other wood boring insects.
19.	___	___	___	I am aware of a structural defect caused by previous infestations of termites or other wood boring insects.
20.	___	___	___	I am aware of underground fuel storage tanks on the property.
21.	___	___	___	I am aware of boundary or lot line disputes.
22.	___	___	___	I have received notice of violation of local, state or federal laws or regulations relating to this property, which violation has not been corrected.
23.	___	___	___	I am aware that this property has been used for the manufacture of methamphetamine as defined in Section 10 of the Methamphetamine Control and Community Protection Act.

Note: These disclosures are not intended to cover the common elements of a condominium, but only the actual residential real property including limited common elements allocated to the exclusive use thereof that form an integral part of the condominium unit.

Note: These disclosures are intended to reflect the current condition of the premises and do not include previous problems, if any, that the seller reasonably believes have been corrected.

 Page 1 of 4

Figure D.1: Example of the Home Seller Disclosure Form (IL)

Back to 2.1.

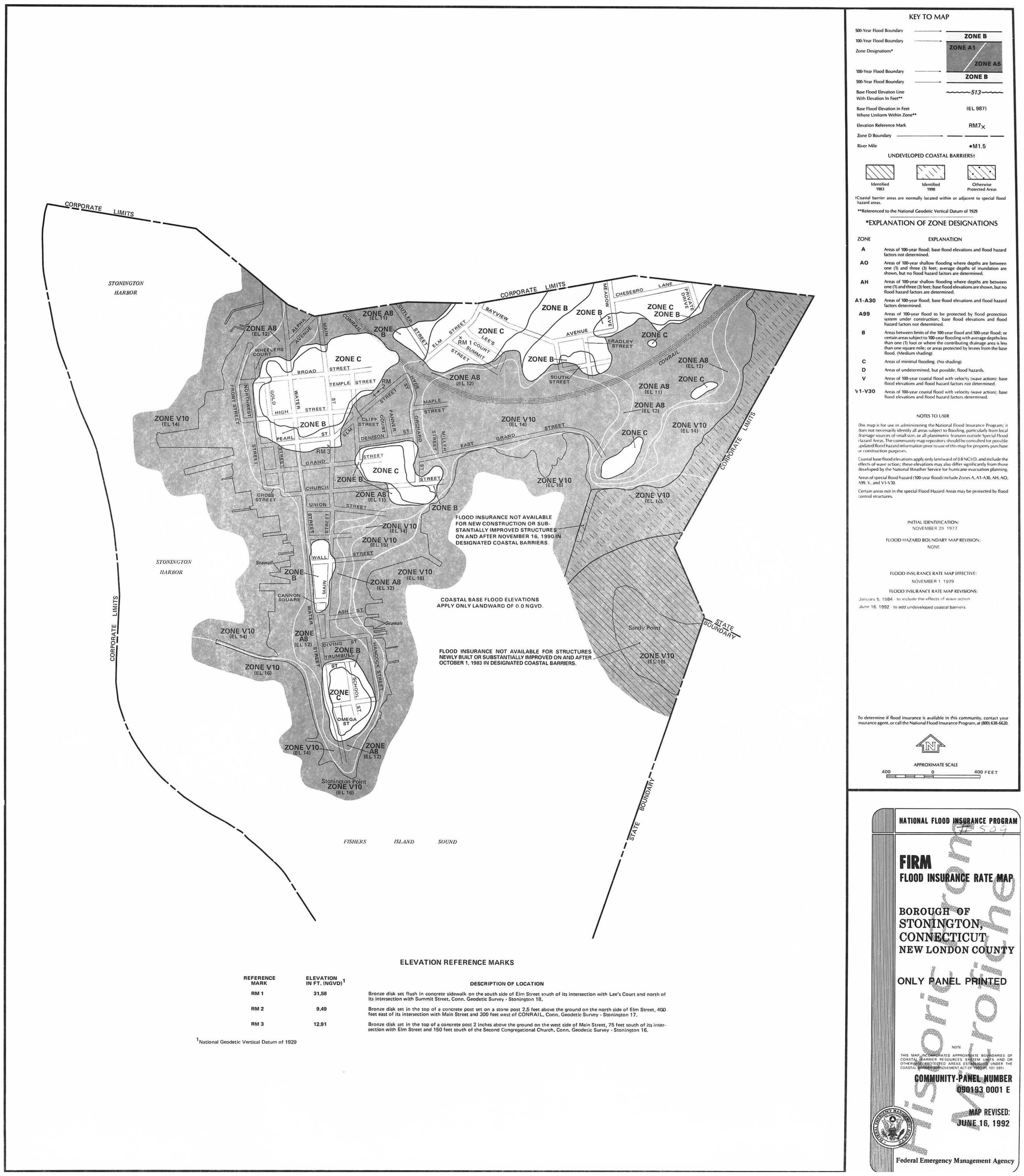


Figure D.2: Sample Flood Insurance Rate Map (Borough of Stonington, CT)

Back to 2.2.

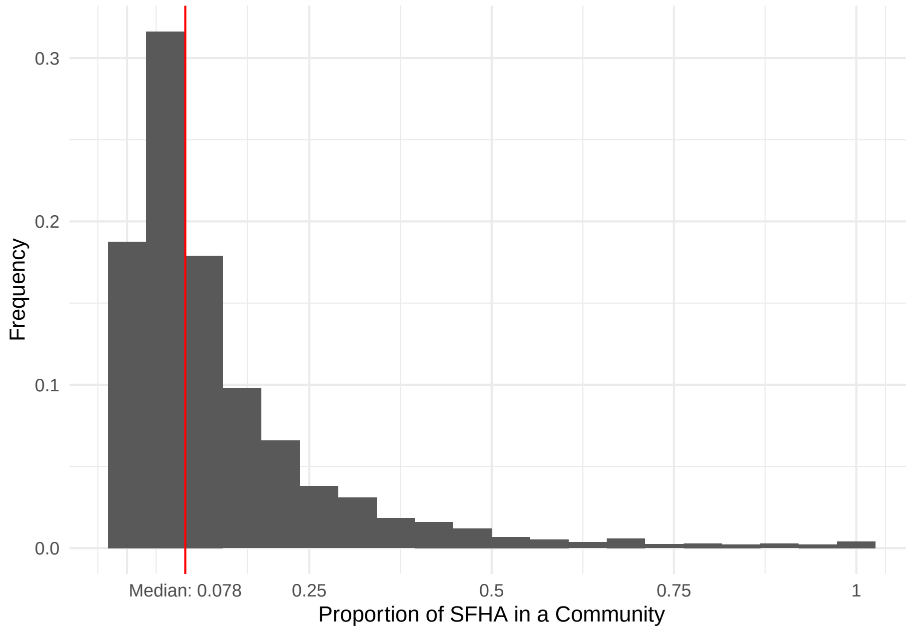
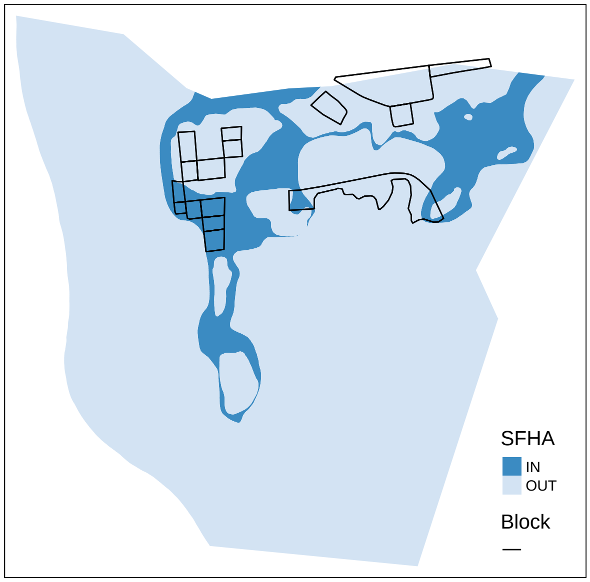
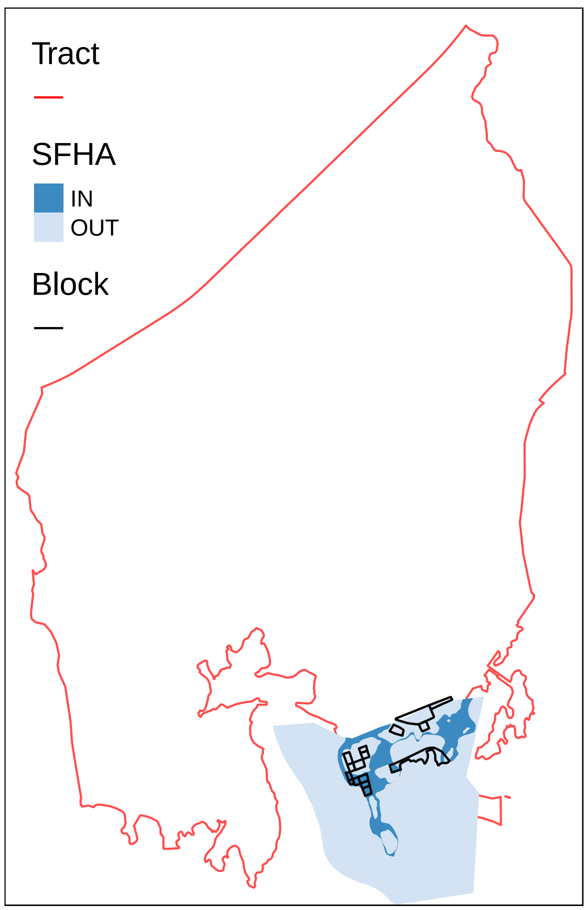


Figure D.3: Histogram of the Proportion of the SFHA at the Community Level. The plot shows the distribution of the SFHA ratio for the 8,194 communities that are on the Q3 map (first generation of digitized flood map) and in the 26 ever-disclosed states.

Back to 2.2.



(a) Block



(b) Tract

Figure D.4: Census Geographies and the SFHA Status (Borough of Stonington, CT)

Back to 4.1.

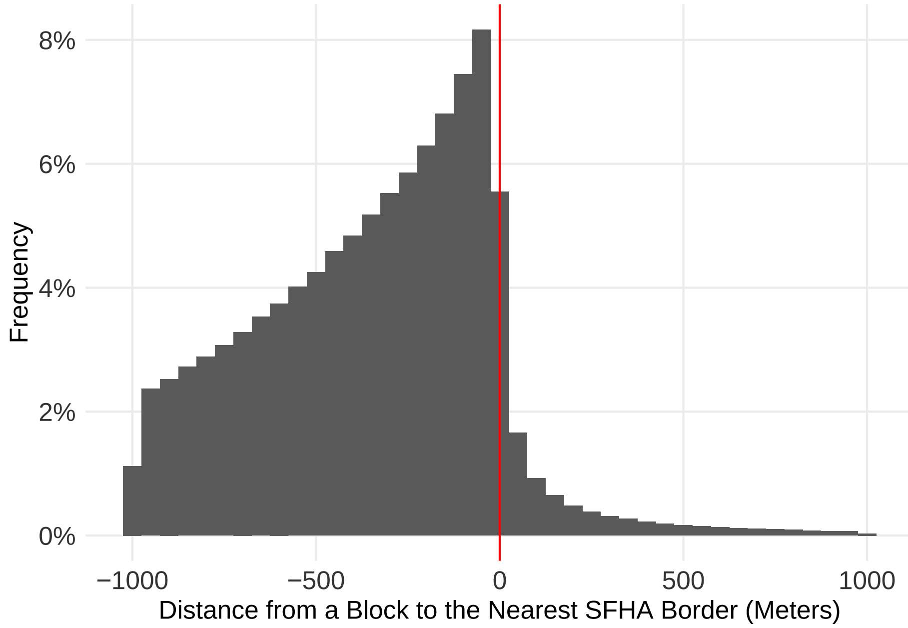
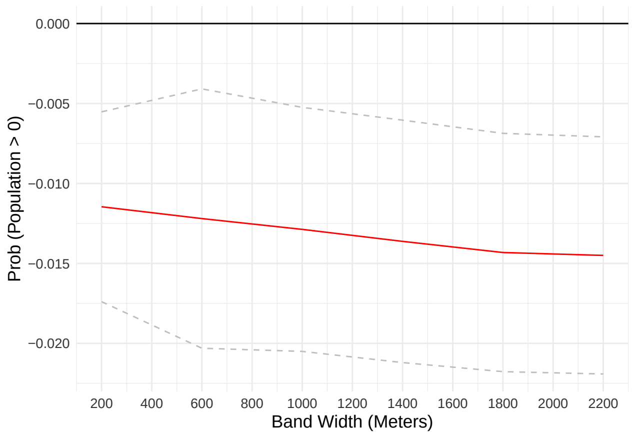
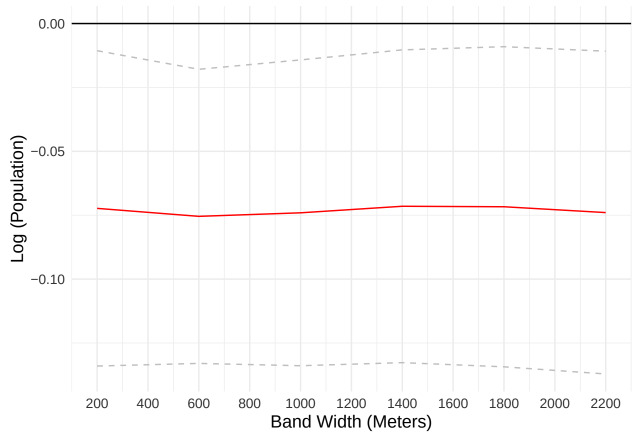


Figure D.5: Histogram of Running Variable (Distance to the SFHA Border)

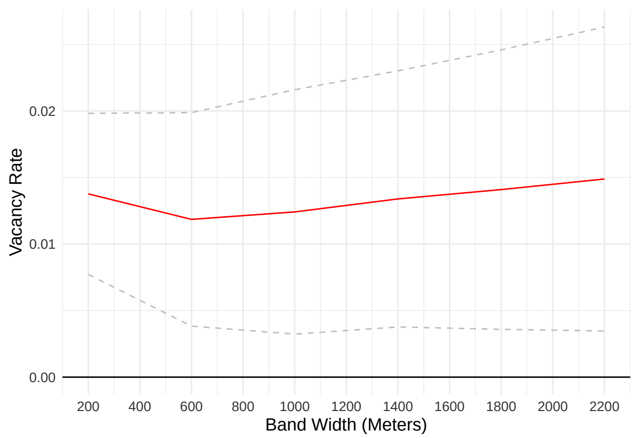
Back to 4.1.



(a) Prob. of Any Population



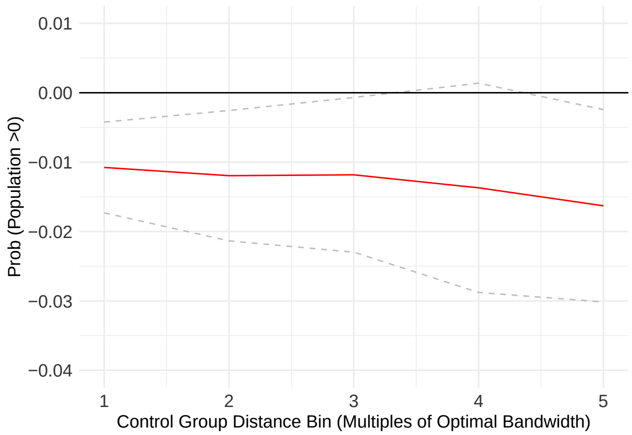
(b) Log Population



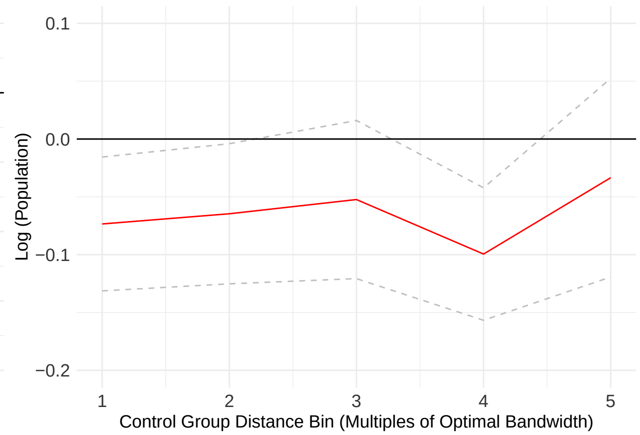
(c) Vacancy Rate

Figure D.6: The Effect of the Disclosure Requirement on Population and Vacancy Rate for Different Bandwidths. These figures plot $\hat{\delta_6}$ from equation (1) for a range of bandwidths. The level of observation is census block, which is the smallest census geographical unit. Standard errors are clustered at the state level.

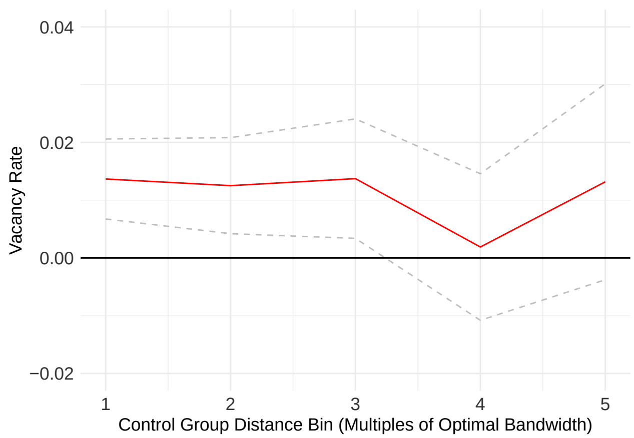
Back to 4.2.



(a) Prob. of Any Population



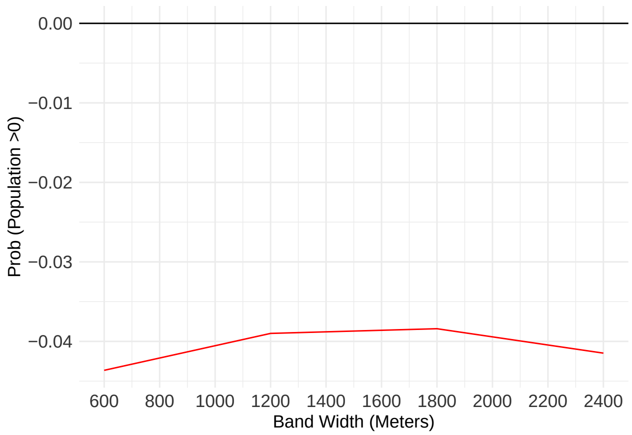
(b) Log Population



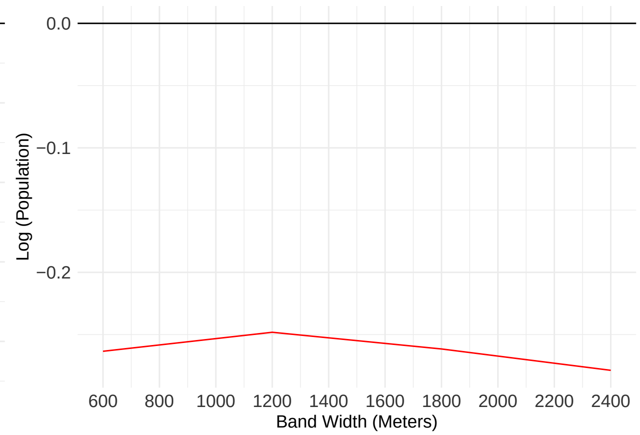
(c) Vacancy Rate

Figure D.7: The Effect of the Disclosure Requirement on Population and Vacancy Rate by Control Group Distance Bin. These figures plot $\hat{\delta_6}$ from equation (1) for control groups of varying distance. X-axis indicates the distance bin of control group in multiples of variable specific optimal bandwidth that has been used for the analysis (e.g., distance bin r on x-axis indicates that blocks that are within $(r - 1)$ and r times optimal bandwidth have been used to form a control group). The level of observation is census block, which is the smallest census geographical unit. Standard errors are clustered at the state level.

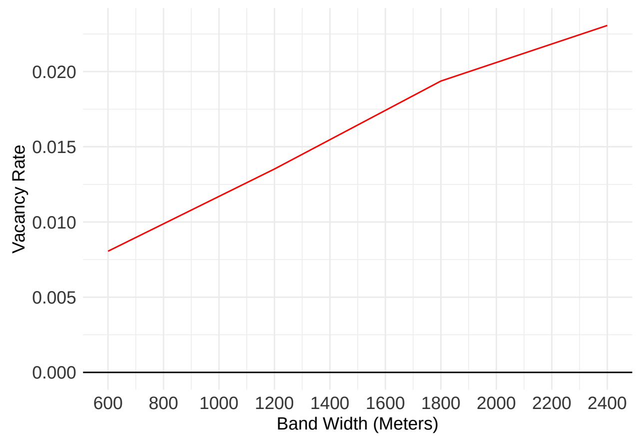
Back to 4.2.



(a) Prob. of Any Population



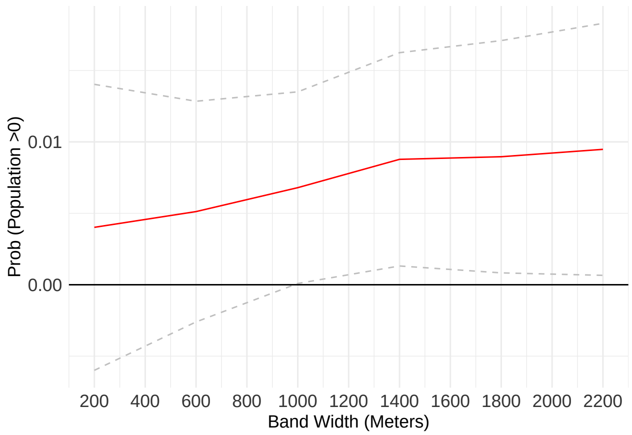
(b) Log Population



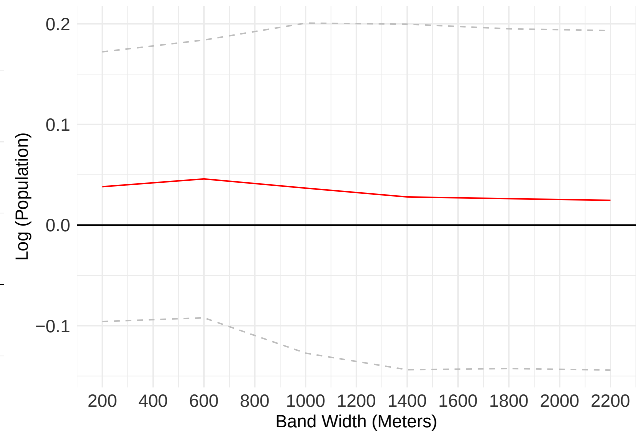
(c) Vacancy Rate

Figure D.8: The Effect of the Disclosure Requirement on Population and Vacancy Rate with Time Varying Discontinuity. These figures plot $\hat{\delta_6}$ from a variant of equation (1) that allows time varying discontinuity at the border. The level of observation is census block, which is the smallest census geographical unit. Standard errors are clustered at the state level but omitted due to readability.

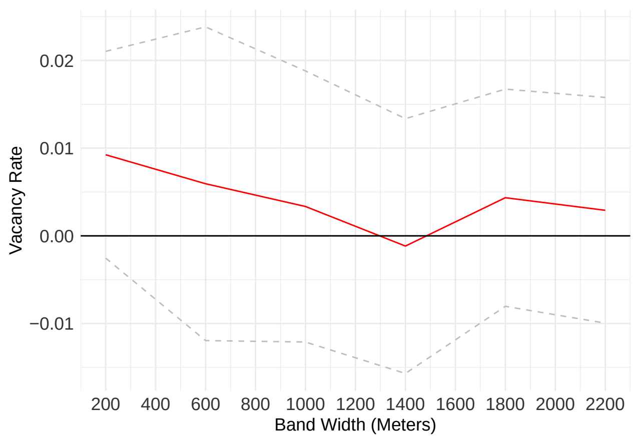
Back to 4.2.



(a) Prob. of Any Population



(b) Log Population



(c) Vacancy Rate

Figure D.9: The Effect of the Disclosure Requirement on Population and Vacancy Rate for Different Bandwidths (Placebo States). These figures plot $\hat{\delta}_6$ from equation (1) for a range of bandwidths for the placebo states. The level of observation is census block, which is the smallest census geographical unit. Standard errors are clustered at the state level.

Back to 4.2.

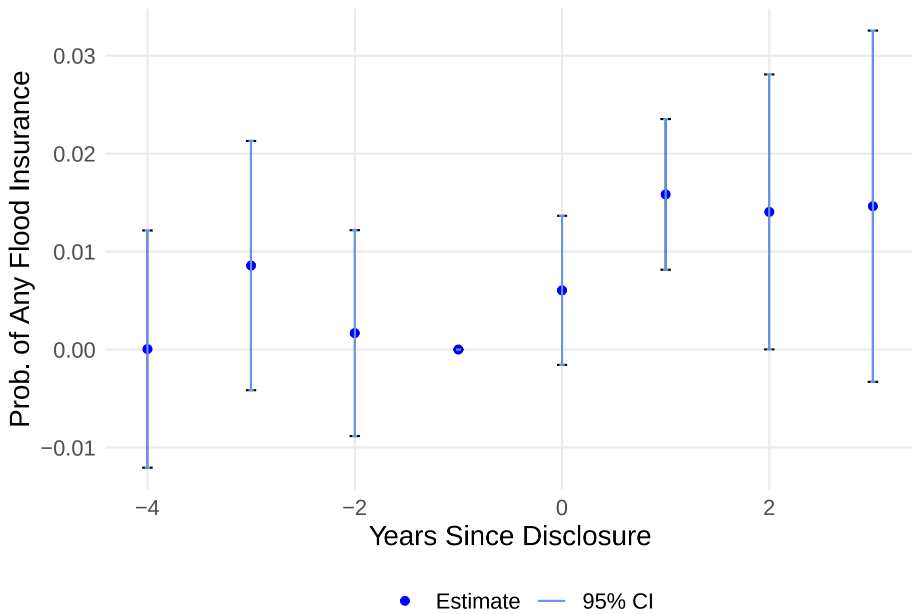
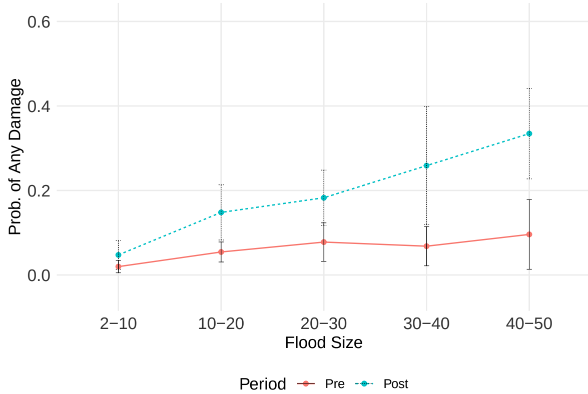
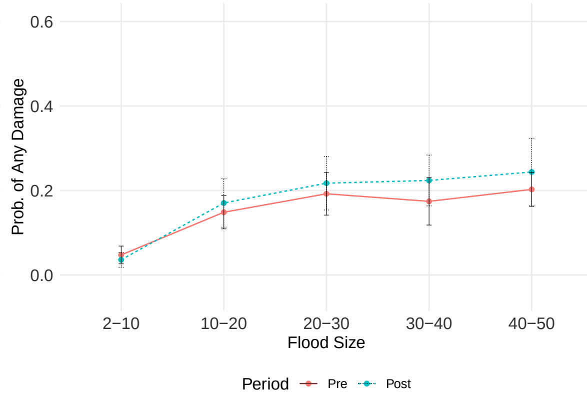


Figure D.10: The Effect of Disclosure on the Probability of Having Any Flood Insurance. This figure depicts the impact of disclosure on the probability of having any flood insurance policy at the community level using an event study version of equation (2). The error bar represents the 95% confidence interval.

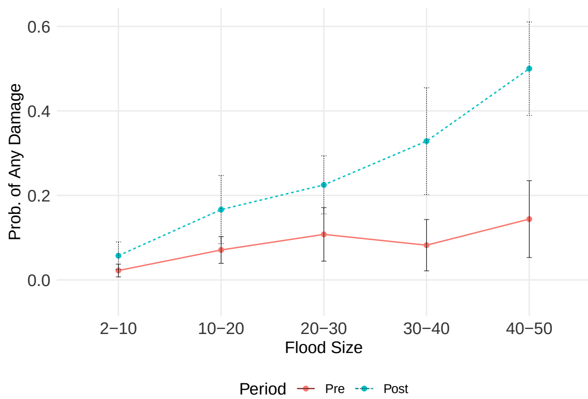
Back to 4.2.



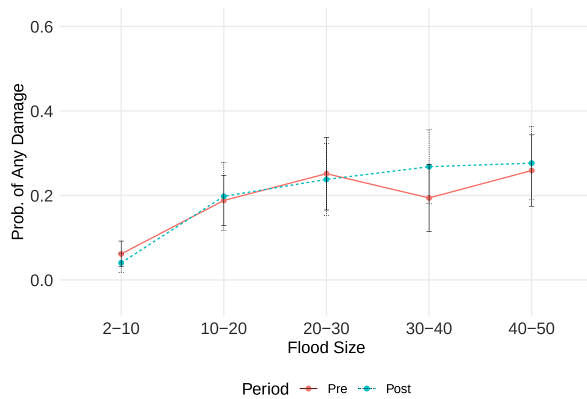
(a) Control Group (All Communities)



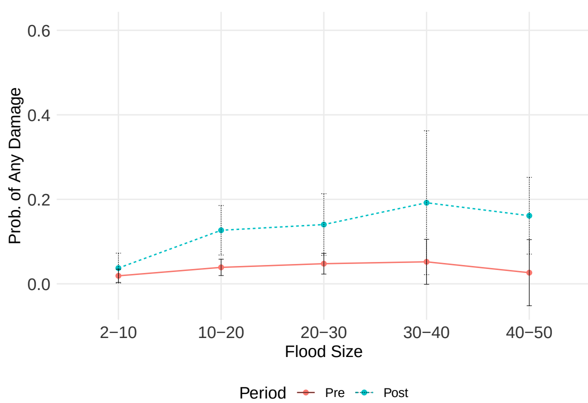
(b) Treatment Group (All Communities)



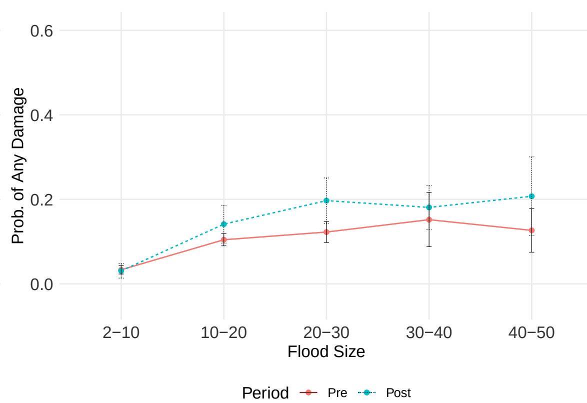
(c) Control Group (High Risk Communities)



(d) Treatment Group (High Risk Communities)



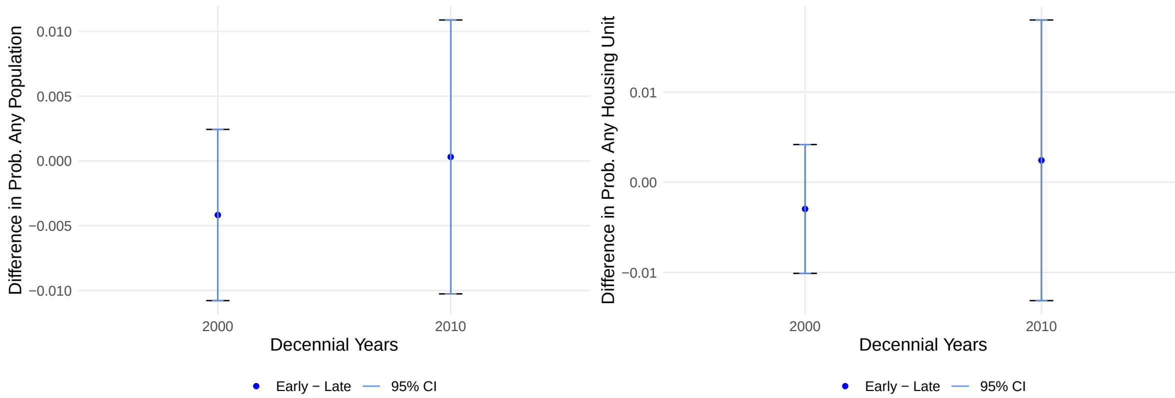
(e) Control Group (Low Risk Communities)



(f) Treatment Group (Low Risk Communities)

Figure D.11: The Effect of Disclosure on the Damage Function with 95% Confidence Intervals. These plots reproduce Figure 5.1 with corresponding confidence intervals.

Back to 5.2.



(a) Prob. of Any Population (b) Prob. of Any Housing Units

Figure D.12: Population and Housing Unit Changes within SFHAs by Treatment Status. These figures plot the (a) difference in probability of having any population and (b) difference in probability of having any housing unit between blocks that were treated early (before 2000) and late (after 2000) over time. The difference in 1990 between two groups are normalized to 0.

Back to 5.2.

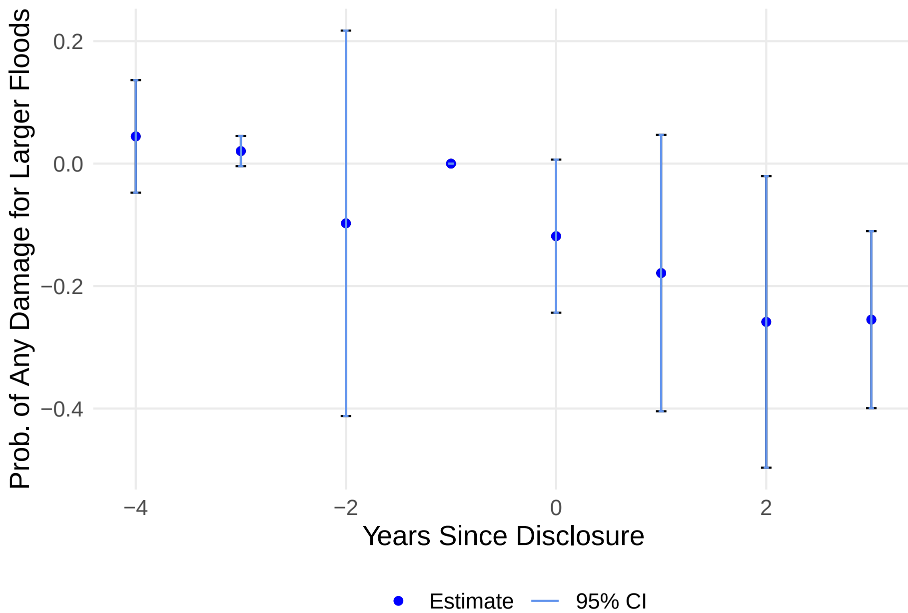


Figure D.13: The Effect of Disclosure on the Damage in Event Time. This figure depicts $\hat{\beta_{4,t}^{30-50}}$ for flood size of 30-50 in event time t where the dependent variable is probability of having any damage. The error bar represents the 95% confidence interval.

Back to 5.2.