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

Seunghoon Lee*

2023-10-09

Abstract

While flood damage is determined by both flood intensity and population exposure, the US has primarily focused on managing the former, with little success. This paper studies whether easing information frictions about flood risk could reduce the exposure and thus flood damage. By exploiting plausibly exogenous variations from the Home Seller Disclosure Requirement, I first show this policy lowers the population in high-risk areas. Further, using a hydrological measure of flood intensity, I find that the policy reduces the probability of flood damage by 33 percent from the baseline. These findings suggest that an information policy could facilitate voluntary adaptation.

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, 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 the high-risk areas), the US flood policy has focused primarily on managing the former by adding engineering structures (Changnon et al. 2000, Field et al. 2012, Tarlock 2012, Liao 2014). However, structural responses can attract more people and developments in floodplains (the so-called "levee-effect") by converting wetlands into habitable land (Pinter 2005, Kousky et al. 2006, Collenteur et al. 2015). The fundamental problem is that none of these structures are perfectly safe. When those engineering structures fail, either due to extreme weather conditions or improper maintenance, the risk of flood damage is heightened with the increased population exposure (Pinter et al. 2016).[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). Although some local governments impose development restrictions and related policies to limit exposure, in most of the US, the population in flood-prone areas is expected to grow rapidly (Wing et al. 2018).[2]

This paper studies whether easing information frictions about flood risk in the housing market could reduce the number of households in high-risk areas and thus flood damage. Although official flood maps have long been publicly available, earlier research and anecdotal evidence suggest a lack of flood risk awareness among home buyers. For instance, Chivers and Flores (2002) find only 14 percent of home buyers in high-risk areas learned about flood risk before closing. Such low awareness hinders home buyers from fully internalizing 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 Home Seller Disclosure Requirement (hereafter "the disclosure requirement") could alleviate the problem by efficiently delivering risk information.

[1] Over 1,000 levees failed during the 1993 Midwest Flood (LARSON 1996). A key factor behind frequent failure is the lack of maintenance. Pinter et al. (2016) find that only 1.9 percent of the levees in the US are rated "Acceptable".

[2] For many local governments, development restrictions are against the interest because it could hurt the tax base.

The policy mandates that home sellers must disclose known property defects 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-difference research design. In addition, the disclosure requirement treats properties located in and out of the SFHA differentially, 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 identifying variation deriving from the fact that the disclosure form considers flood risk in a discontinuous manner. Specifically, home buyers 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. The spatial discontinuity yields an opportunity to identify the information effect holding true flood risk constant. A concern is that being located in the SFHA could invite other treatments such as the mandatory purchase of flood insurance. To account for that possibility, I use the difference-in-discontinuity approach following Grembi et al. (2016).

To leverage these variations, I collect multiple datasets. I collect 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 a historical dataset of community-level past flood events based on a hydrological measure of flood intensity (Saharia et al. 2017, England Jr et al. 2019). These data overcomes the potential endogeneity of self-reported flood events, such as from the National Weather Service Storm Events data (Gall et al. 2009). Because main outcome variables used in 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. The first part investigates home buyer 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 on population distribution and flood insurance take up. The second empirical exercise estimates flood damage function to test if the disclosure policy reduces flood damage.

From the first part, I find that census blocks in the SFHA area (conditional on having a non-zero population) experiences a 7 percent decline in population after the disclosure policy. At the extensive margin, it lowers the probability of a block in the SFHA having any population by 0.01 percentage point (or 1.5 percent from the baseline of 0.67). Further, while most location adjustments occur within a community, home buyers seem to choose further away alternativess, which are likely to have meaningfully lower flood risk.

In contrast, I find a very small effect of the disclosure policy on insurance purchases: the probability of having a positive number of insurance policies at a community level decreases by 0.001 percentage point (or 0.1 percent from the baseline of 0.83; extensive margin) while insurance counts per housing unit decreases by 0.9 percent (intensive margin). Investigating these two potential responses 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, 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 2.5 percentage points (or 33 percent of the baseline probability). 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 the greatest flood risk.

This paper contributes to four different bodies of literature. First, it is related to prior work on factors mitigate damage from climate change. Whereas 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.

Second, it contributes 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 paper conceptually is Baylis and Boomhower (2021), which shows how building code policies can mitigate wildfire damage. A key difference is that the policies studied by Baylis and Boomhower (2021) directly mandate adaptation, whereas I show how a disclosure requirement can encourage voluntary adaptation.

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

Finally, I contribute to the literature 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, which allows me to document flood events objectively for a wide range of causes including rainfall, snow melt, or storm surge. This measure complements measures from existing flood damage functions that specialize in capturing the impact 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.

2 Background

2.1 Home Seller Disclosure Requirement

Publicly available Flood Insurance Rate Maps contain the information home buyers need to determine whether a property is located in an SFHA. Also, 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 home buyers 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 home buyers do not have a good understanding of the flood risk they face (Chivers and Flores 2002, Pope 2008, Bin and Landry 2013).[3] 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 Illinois Residential Real Property Disclosure Report in Appendix Figure C.1 illustrates, a typical form covers a wide range of property conditions such as structural issues (e.g., problems with walls, roofs, or plumbing) and surroundings (e.g., natural hazards such as flood risk).

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.[4] Because properties on the SFHA are more susceptible to flood, 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, the disclosure may raise home buyers' flood risk awareness for properties in SFHAs relative to those outside.

Disclosure background. Traditionally, home buyers were legally expected to exercise proper caution with regard to potential defects of a property (the so-called "*caveat emptor*" or "let the buyer beware" doctrine). However, with the rise of consumer protectionism (Lefcoe 2004), and with higher public attention to environmental contamination and health issues during the 1980s (Weinberger 1996), state courts increasingly held listing agents responsible for incomplete disclosures. In response, the National Association of Realtors issued a resolution in 1991 encouraging state associations to develop and support legislation regarding the statutory disclosure requirement (Tyszka 1995).

Consequently, between 1992 and 2003, 26 states in the contiguous US implemented the disclosure requirement with an explicit question on flood risk (see Figure 2.1). As its historic origin suggests,

[3] Anecdotal evidence also suggests that a large number of home buyers are not well aware of flood risk (Flavelle 2017, Satija et al. 2017). Also, Michel-Kerjan (2010) find only 20-30 percent of homeowners in the SFHA purchased flood insurance in 2000.

[4] 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.

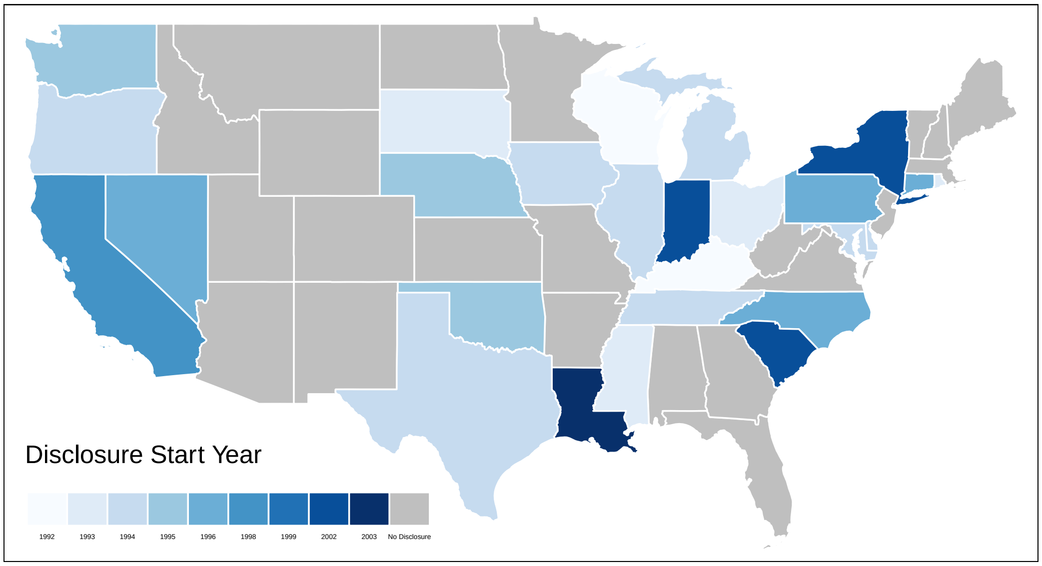


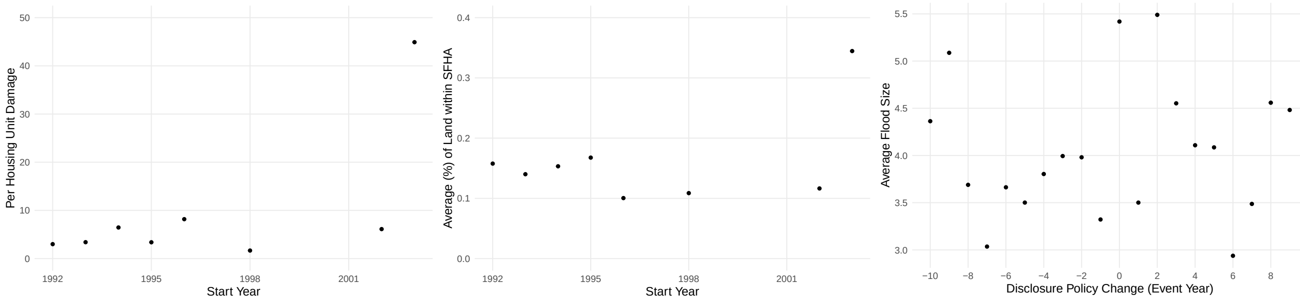
Figure 2.1: The Disclosure Requirement Implementation over Time

the disclosure policy was primarily an effort of realtors to deflect potential liability to sellers (Washburn 1995) and the timing of the policy implementation is related to the timing of the change in the state court's view on the *Caveat Emptor* doctrine (Roberts 2006).

Given the background and content of disclosure requirements, the implementation timing is unlikely to be correlated with each state's underlying flood profile. In Figure 2.2, 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.[5] In contrast to panels (a)–(b) where the x-axis is in calendar years, panel (c) shows the average flood size in event time. 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.

Does the disclosure requirement matter? Although many states levy fines or even allow buyers to rescind their purchase agreement without penalty to ensure compliance, the disclosure policy might

[5]Spikes in Figure 2.2 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.



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

Figure 2.2: 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 a flood risk disclosure policy.

still fail to raise home buyers' flood risk awareness. For example, home sellers might not comply with the regulation. If they furnish the form with inaccurate information or ignore the requirement despite the potential penalty, the disclosure policy's effectiveness could be seriously undermined. Home buyers might also fail to process the new information. While the disclosure forms consist of straightforward check box questions (see Appendix Figure C.1), the implication of living in an SFHA might not be fully understood. Thus, while there are good reasons to expect the disclosure policy to reduce information frictions around flood risk, whether it does is an empirical question.

Unfortunately, I cannot observe home buyers' perceptions of flood risk directly. However, in Appendix B I find disclosure reduces housing prices by 4.5 percent, which coincides with existing estimates of the effect of flood risk information on housing prices (e.g., see Hino and Burke (2021)). This finding provides evidence that at least some information is being conveyed to buyers through the disclosure policy. The extent to which this information translates to risk-mitigating behaviors is uncertain and the focus on my empirical work below.

States without disclosure requirements on flood. Before further proceeding, it is worth discussing the twenty-two states in the contiguous US (excluding Washington DC) do not mandate home sellers to disclose on flood risk. Despite the lack of mandatory requirements, more than 60 percent of these states have seen the formation of voluntary disclosure forms by their respective state realtor associations. In some states, realtor associations even require member realtors to use these forms.[6] Thus, a non-trivial number of home buyers in these "non-disclosure" states might have received information

[6] See Colorado, for instance, from Flood Disclosure Scorecard from NRDC https://www.nrdc.org/flood-disclosure-map (accessed on Sep 8, 2022).

on flood risk. The timing of these realtor-driven behavior is uncertain, as is compliance. For this reason, I exclude non-disclosure states and use later-treated states instead as the control group for earlier-treated states when I exploit the staggered adoption variation. Indeed, the inclusion of these states seriously attenuates the treatment effect and/or creates a pre-trend (see Appendix Figure B.1 panel (b) and Appendix Table C.5 and accompanying texts). Even in the late-adopting states, some realtor-driven disclosure behavior might have been in place prior to the state policy change, which can be viewed as attenuating my findings relative to a pure counterfactual of no disclosure.

Importantly, five states (out of the 22 states) have adopted a variant of home seller disclosure policy, although it does not have a question on flood risk.[7] These "placebo" states are useful for checking the robustness of the main results.

2.2 Flood Map and Special Flood Hazard Area (SFHA)

An official flood map, or the Flood Insurance Rate Map, allows stakeholders to identify the boundary of an SFHA, determine a specific property's SFHA status, and determine the Base Flood Elevation (FEMA 2005). The SFHA, an area expected to be inundated with a 100-year flood, is a particularly important concept because flood risk communications frequently refer to it.[8]

The flood mapping process involves three key steps (FEMA 2005): (1) hydrologic analysis that determines the water amount in a stream channel for a given weather event; (2) hydraulic analysis that translates water amount to water surface elevation; and (3) floodplain mapping, which determines the boundary of inundation by comparing water surface elevation with the ground elevation. The procedure 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 each side of the border with almost identical true flood risk.

It is also worth noting that these maps are updated occasionally, which could potentially confound the disclosure policy. While the National Flood Insurance Reform Act of 1994 requires that FEMA

[7] For details regarding the extent of disclosure in these states, see the following. Idaho: 1994 Ida. HB 825 (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).

[8] Flood is defined as "a general and temporary condition of partial or complete inundation of two or more acres of normally dry land area or two or more properties from an overflow of inland or tidal waters, from unusual and rapid accumulation or runoff of surface waters from any source, or from mudflow" (FEMA 2005).

assess the need to revise and update all flood maps every 5 years, the vast majority of the maps fail to meet the required update cycle (DHS Office of Inspector General 2017). This is favorable for my research design because flood zone status for the majority of areas remains constant over the study period. Indeed, in Appendix B, I show that excluding properties from communities with map updates does not change the estimated effect of disclosure on the housing price.

The jurisdiction of each flood map is a National Flood Insurance Program "community", local political entities comparable to a US Census place. Appendix Figure C.2 shows a sample Flood Insurance Rate Map from a part of the Borough of Stonington, CT, and similar to this place, a typical entity has both SFHA and non-SFHA areas. Indeed, a histogram in Appendix Figure C.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

I compile multiple data sets on block and tract level demographics, community-level numbers of flood insurance policies, and flood damage. I also construct a community-level flood history dataset. In this section, I describe each data source and provide descriptive statistics.

Demography and flood insurance. Demographic characteristics come from two sources. I collect census block level population and occupancy data from the 1990, 2000, 2010, and 2020 decennial census. To account for changing block boundaries and resulting one-to-many matches across different decennial census years,[9] I calculate the weighted sum of count variables using interpolation weights from the NHGIS block-to-block crosswalk (Manson et al. 2022).[10]. This creates a geographically standardized time series. For other demographic characteristics such as income, age, race, and education, which are not available at the block level—the smallest census geographic unit—level, I utilize tract-level data from the Geolytics for 1990, 2000, and 2010 decennial census.[11] Data on the number of flood insurance policies by the NFIP community are available from FEMA from 1978–2008.[12]

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

[10] 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)

[11] Geolytics data provide tract level data after accounting for changing boundaries across different survey years (for more details, see www.geolytics.com).

[12] I thank Justin Gallagher for graciously sharing this data.

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 single-family houses and collapse it to the community-by-year-by-the-largest-flood-event level to match it with the annual maximum flood event.

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 community-level flood history data 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 Handbook of Management Group D of ASCE 1996): the expected number of years for a flood of the same magnitude to come back. Flood size is conveniently increasing in the recurrence interval. 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. This approach objectively measures the intensity of floods from various causes.

Practically speaking, the data is constructed 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. [13] Finally, I match each community to the three nearest gauges and calculate community-year-level flood size by taking the inverse-distance weighted

[13]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.

average of three closest gauges' recurrence intervals. More details on the flood data construction procedure and summary statistics are in Appendix A.1.

Other data sources. As the disclosure policy differentially treats the properties in and out of the SFHA, I spatially merge the NFIP community, block, and tract with the digitized flood map to determine the SFHA status of each geographic unit. Specifically, I use the Q3 map—the first generation of a digitized flood map—reflecting flood risk as of the mid-1990s. The map covers about half of all FEMA communities based on population density and the intensity of past flood incidents, and my main sample consists of these communities (FEMA 1996). Finally, the primary data source to track the disclosure requirement legislative history is the *Nexisuni* database. I cross-validate this database with prior works 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 (flood size) and dependent (population, flood insurance policy counts per housing unit, and flood damage per housing unit) variables of this paper. Population figures are for the Census blocks within the optimal bandwidth, while the last three values are from 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. Importantly, the prevalence of zeroes in the dependent variables are 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 these three dependent variables. 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.[14]

[14] 2.18 is slightly higher than 2 because I used annual peak flow data until 1990. More details are in Appendix A.1.

Table 3.1: Summary Statistics for Key Variables

Variables	Min.	Q25	Median	Mean	Q75	Max.
Census Block Population	0	0	10	34.4	40	7,597
NFIP Policies Per Housing Unit	0	0.001	0.006	0.03	0.019	6.53
Flood Damage Per Housing Unit	0	0	0	6.97	0	23,991
N of 10-Year Floods (For 20 Years)	0	1	2	2.18	3	15

4 Responses to the Disclosure Requirement

In this section, I investigate how home buyers respond to flood risk information. 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 on population distribution and flood insurance take up. Investigating both responses is important because they have starkly different implications for flood damage. That is, while choosing a safer location to live reduces the exposure to risk, flood insurance would simply redistribute income from the "dry state" to the "flooding state" without necessarily affecting the exposure (Ehrlich and Becker 1972).[15]

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 flood risk information effect from the true risk effect. However, a potential concern is that other policies such as flood insurance requirements also change at the border, which could confound the disclosure. To account for this issue, I leverage a difference-in-discontinuity approach. By taking the difference between the two spatial regression discontinuity estimates (for pre and post-disclosure periods), the design controls for time-invariant confounding factors.

Following Grembi et al. (2016), I estimate the policy effect on population and the vacancy rate using block-level decennial census data, which is the smallest Census geographical unit. 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. The distance to the border is defined by the distance

[15] 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.

between a block and the closest SFHA border.

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

To estimate δ_6, I first estimate the optimal bandwidth for each outcome variable. Then, I estimate equation (1) using observations within the optimal bandwidth (Calonico et al. 2014, Cattaneo et al. 2019).[16] In equation (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. X_{bs} is the distance from a border in meters (negative if in a non-SFHA area), $D_{bs} = 1$ (i.e., in the SFHA) 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 state s. δ_6 captures the impact of the disclosure policy for blocks located in close proximity to the SFHA border.

A potential concern of using a geographic area such as a block (namely, a polygon) rather than a property (namely, a point) is that a block might contain an SFHA border within it. In this case, the distance from a block to an SFHA border is not well defined. While this might be a serious problem for larger geographical units such as tracts, it will be less of a problem for blocks that are small. For instance, the median size of the census blocks in my sample is 0.009 square miles, and 83 percent of blocks are perfectly contained within or outside an SFHA as an example in Appendix Figure C.4 illustrates. To further alleviate the concern, when calculating the distance between a block and the nearest SFHA border, I take the difference of (1) the distance between block centroids and the closest SFHA border and (2) a block diameter.

Staggered Adoption. Flood insurance policy counts and demographic characteristics such as income are observed at either community or tract level. As these geographic units are typically much larger than the SFHA (Appendix Figure C.4 (b)), the distance to the nearest SFHA border is not well defined. Thus, I employ a triple difference design using equation (2). Here, Y_{mstd} denotes various outcome variables on NFIP and demographic characteristics for geographic unit m in state s in year t in stack d. H_{md} is an indicator variable equal to one if a geographic unit has an above-median

[16] 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.

fraction of the area covered by an SFHA, which proxies for the treatment intensity. I_{std} is a post disclosure indicator while D_{std} is a treatment group indicator. α_7 captures the disclosure effect.

$$\begin{aligned} log(Y_{mstd}) =& \alpha_1 H_{md} + \alpha_2 D_{std} + \alpha_3 I_{std} + \alpha_4 H_{md} D_{std} + \alpha_5 I_{std} D_{std} + \alpha_6 H_{std} I_{std} + \\ & \alpha_7 H_{md} D_{std} I_{std} + \omega_{td} + \psi_{md} + \epsilon_{mstd} \end{aligned} \tag{2}$$

As the stack subscript d alludes, 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 over 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 stacked dataset, I first keep each state's flood insurance policy counts for seven years before and seven years after the policy change. Inclusive of the change year, I use 15 years of data for each state.[17] Each stack consists of geographic units in the treated states, which have implemented the disclosure policy in year t^*, and geographic units in the control states, which have implemented the policy in year $\tilde{t} > t^*$.[18] I drop observations from the control states for $t >= \tilde{t}$ because it is no longer "not-yet-treated".

Equation (2) also include ω_{td}, the time $\times$ stack fixed effect to account for year-specific common shocks and a community or tract $\times$ stack fixed effect ψ_{md}, which captures an unobserved community or tract characteristics. Including fixed effects interacted with stack d ensures that the comparisons are made within each stack.

It is worth discussing one additional detail about the tract-level analysis. Because the decennial census is documented once every 10 years, states that implemented the policy after 2000 is used only as a control group, because there is no control group for them—every ever-treated state is treated in 2010. Thus α_3 is estimated from a single data stack consists of early treated states (treated before 2000) and late treated states (treated after 2000).[19] Because only two time periods are used the estimating equation (2) reduces to a standard triple difference specification with discretized treatment. Throughout the analysis, I use spatial-HAC standard errors that allow spatial correlation of

[17] The data from 1978–2008 are sufficient to cover 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.

[18] 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 $t^* < 1996$.

[19] Note the analysis on flood insurance take up does not have this problem as the observations are community-year.

up to 500 miles are estimated for inference (Newey and West 1987, Conley 1999). 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).

4.2 Findings

Self-protection. Table 4.1 reports the impact of the disclosure on population over an up to 20 year period. In column (1), I find that the disclosure reduces the probability of having any population in an SFHA block by 0.01 relative to a non-SFHA block (or 1.5 percent of the baseline value 0.68). In column (2), I limit the sample to the blocks with non-zero population and find that the disclosure reduces population in an SFHA block by 7 percent relative to a non-SFHA block. Taking these extensive and intensive margin effects together, the policy discourages not only in-migration to the existing properties in an SFHA but also new developments in previously uninhabited SFHA areas.

Importantly, these population effects are contained within a community. Using equation (2), I show in column (3) that the effect of disclosure on the community level population is only –0.6 percent—an order of magnitude smaller than column (2).[20] This is plausible because, as Appendix Figure C.3 shows, a typical community has a large fraction of land area outside of SFHA and thus flood risk can be easily avoided by within-community adjustments. Such a local location adjustment pattern in response to flood risk is documented in prior works as well (Noonan and Sadiq 2019).

Consistent with columns (1)–(2), in column (4), I report that the disclosure increases the vacancy rate for the blocks in the SFHA from 0.095 to 0.109.[21] 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.[22] Table 4.1 resonates with broader literature that people migrate away from negative environmental conditions, although the extent of migration in this paper seem local (Banzhaf and Walsh 2008, Boustan et al. 2012, Hornbeck 2012, Hornbeck and Naidu 2014).[23]

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

[20] I linearly interpolate community population using the decennial census.

[21] 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).

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

[23] 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.

Table 4.1: Effect of Discosure Requirement on Net Population Flow

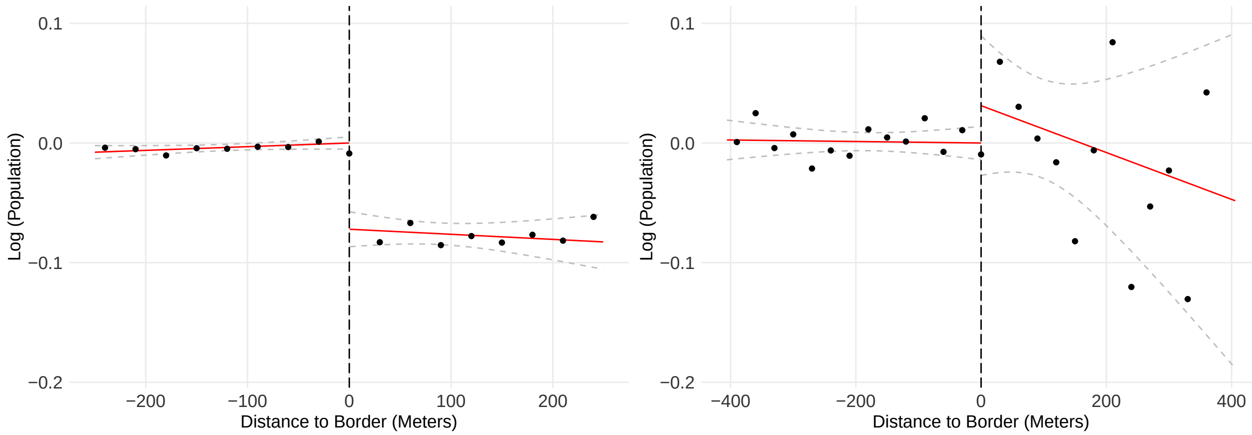
	(1)	(2)	(3)	(4)
SFHA × Post	$-.011^{***}$	$-.072^{**}$		$.014^{***}$
	(.003)	(.034)		(.005)
High SFHA × Disclosure × Post			$-.006^{**}$	
			(.003)	
D.V	Prob. of Any Population	Log Population	Log Population	Vacancy Rate
Avg D.V.	0.674			0.097
Year × Stack FE			X	
Community × Stack FE			X	
Bandwidth	136	250		190
Num. obs.	1465392	1651843	499075	1306372

Note: This table is produced from equation (1) and (2). 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. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

running variable of the spatial RD design, which is 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 for the non-SFHA blocks is normalized such that $\Delta Y^{-} = 0$. I also overlay a scatterplot, which shows the means of the difference in pre and post disclosure periods log population with a bin size of 30 meters.

The figure indicates that 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, which is consistent with column (2). Additionally, the figure shows that the regression line tightly fits the scatter plot, which suggests that the choice of functional form for the running variable is unlikely to have a substantial impact on the RD estimates.

Table 4.1 implies that most location adjustments are local, which raises a concern about disclosure's deterrence effect. That is, if diverted buyers choose a property on outside the SFHA but very close to the border, the risk exposure would be essentially identical despite the population reduction effect discussed earlier. To gain insights on this issue, I repeat the analysis in columns (1)–(2) and (4) of Table 4.1 after removing blocks that are within 20 and 40 meters from the border. In Appendix Table C.2, I show the directions and magnitudes of my doughnut specification estimates are similar to Table 4.1. Moreover, Appendix Figure C.5 shows that there is no diminishing policy effect



(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) treated and (b) placebo states. The discontinuity at the threshold (dotted 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.

as I expand the bandwidth. These findings suggest when home buyers are deterred from an SFHA, they choose further away places, which are likely to have meaningfully lower flood risk.[24]

In Appendix Table C.3, I explore the demographic composition change as a result of the disclosure. The estimated coefficients suggest that higher-risk tracts become less affluent and less old in comparison to low-risk tracts. This finding coincides with Bakkensen and Ma (2020) in that those with more resources tend to choose a safer place to live.

While difference-in-discontinuity design controls for time-invariant confounders, a concurrent policy change can still bias the estimates in Table 4.1. One possibility is flood insurance premiums, which could have changed differentially over time for the SFHA and non-SFHA areas. Another possibility is a change in enforcement of the flood insurance purchase requirement over time. To rule out this possibility, I use the five placebo states that have implemented a disclosure policy without a question about flood risk. If my findings are driven by concurrent policy changes rather than the disclosure, I would expect to find a similar effect in the placebo states.

In Appendix Table C.1, I find no evidence of a reduction in population or an increase in the vacancy rate in the placebo states. If anything, my estimates are either precisely estimated null effects (for the extensive margin population and the vacancy rate) or positive (for the intensive margin pop-

[24]These findings also rules out a SUTVA assumption violation, which could overestimate the effect size in Table 4.1.

ulation). Figure 4.1 (b) visually confirms this point: in contrast to panel (a), there is no reduction in the population after the disclosure policy and if anything, the point estimate is positive.[25]

While choosing a safe location represents an extensive margin self-protection strategy, elevating structures to prevent inundation is one of the most common intensive margin responses (Montgomery and Kunreuther 2018, 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.

Market Insurance. Because the level of observation in this exercise is a community, the distance to the SFHA border is not well defined. Therefore, I leverage the staggered adoption of the disclosure policy to study its impact on insurance take up in a triple difference framework. The differences are over time, treatment group assignments, and between communities with high versus low exposure to SFHA areas (measured by the fraction of total area covered by an SFHA).
In Table 4.2 column (1), I show the disclosure policy reduces the probability of having a positive flood insurance policy per housing unit for high-risk communities by 0.001 percentage point (or 0.1 percent from the baseline of 0.83).[26] Similarly, column (2) suggests the intensive margin effect of the disclosure policy on the number of insurance policies per housing unit is –0.9 percent. Given tight standard errors and small point estimates, home buyers do not seem to respond to the disclosure policy by purchasing flood insurance.

Why do home buyers engage in self-protection despite the option to buy flood insurance? One possibility is that the benefit of flood insurance could be small. For instance, the NFIP offers an incomplete insurance with a coverage capped at \$250,000. Further, a flood could negatively affect an individual's health or employment status, and even seriously disrupts daily lives, all of which are

[25] In Appendix Figure C.7, I conduct a more formal test by taking the difference of the diff-in-disc terms for the treated and placebo states by estimating an augmented version of equation (1): $Y_{bmst} = \delta_0 + \delta_1 X_{bms} + \delta_2 D_{bms} + \delta_3 X_{bms} * D_{bms} + T_{st}[\delta_4 + \delta_5 X_{bms} + \delta_6 D_{bms} + \delta_7 X_{bms} * D_{bms}] + H_s[\delta_8 + \delta_9 X_{bms} + \delta_{10} D_{bms} + \delta_{11} X_{bms} * D_{bms} + T_{st}[\delta_{12} + \delta_{13} X_{bms} + \delta_{14} D_{bms} + \delta_{15} X_{bms} * D_{bms}]] + \epsilon_{bmst}$. Here δ_{14} is the coefficient of interest. The estimates suggest that the effect size (in magnitude) is even larger when we consider the trends in the placebo states.

[26] While columns (1) and (2) report coefficients of the High Risk $\times$ Disclosure $\times$ Post term only, I include a full set of interaction terms for estimation.

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

	(1)	(2)
High SFHA × Disclosure × Post	$-.001$	$-.009$
	(.006)	(.008)
D.V	Prob. of Any Insurance	Log Insurance Per Housing Unit
Avg D.V.	0.826	
Year × Stack FE	X	X
Community × Stack FE	X	X
Num. obs.	439822	363476

Note: This table is produced from equation (2) using community-level National Flood Insurance Program data. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

not covered by flood insurance (Kahn 2005, Deryugina 2017, Lee et al. 2023). Another important reason is that the cost of location adjustment is substantially lower for home buyers—especially in comparison to non-moving households. Indeed, Zumpano et al. (2003) documents that home buyers actively seek for alternatives and an average buyer physically visits 17 properties before closing.

Findings in this section collectively indicate that households primarily respond to the risk information by selecting a safer location. If instead, the primary response is purchasing more flood insurance (and thus muting self-protection measures), the disclosure policy would have a limited impact on damage reduction.

5 The Effect of the Disclosure Requirement on Flood Damage

5.1 Estimation Framework

For a given flood 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 to 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.[27] Surprisingly, there has been limited attention directed towards damage functions specific to floods, despite the substantial disruptions they cause. That is partly because objective measurement of flood size is challenging, and I overcome this by constructing a hydrology-based flood history

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

dataset detailed in Appendix A.1.

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

Before delving into the specifics of the estimation process, it's beneficial to conceptualize the damage function. Consider equation (3), where the dependent variable is per housing unit flood damage. D is a dummy for the treated group assignment and F^k is a dummy that takes 1 when the annual maximum flood size is in flood size bin k where $k \in \{$2-10, 10-20, 20-30, 30-40, 40-50$\}$. I take a non-parametric approach following Barreca et al. (2016) to let the data, rather than the functional form assumption, determine the shape of the function.

There are a couple of points to discuss regarding the F^k variable. First, I use the annual maximum flood size as a proxy for flood exposure for a given community-year. While this implies that smaller floods occurred in the same year are ignored, it is unlikely to be critical given that the majority of the community-years in the dataset had just one flood as Appendix Figure A.3 (c) illustrates. Further, 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 and thus measurement error becomes a serious issue (Kron et al. 2012). Further, as Appendix Figure A.3 (b) shows, the frequency of flood events reduces exponentially as size increases. This implies that non-parametrically identifying statistical relations for highly infrequent floods is challenging. Also, flood sizes 2–50 cover a wide enough band to capture floods of different severities. Indeed, 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.

Lastly, the key assumption behind binning is 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).

Flood size between 1 and 2 is the baseline size and is the omitted category. Thus, α_1^k in equation (3) indicates the additional flood damage per housing unit when a community in the control group experiences a flood of size k as opposed to the baseline flood. α_2 allows a different slope for the treated group accounting for potential differences in flood exposure or policies.

Now posit that a disclosure policy is implemented. Equation (4), which mirrors a canonical difference-in-difference model, shows how equation (3) changes when post period dummy I is introduced. The coefficient for the interaction term (β_4^k) captures the treatment effect.

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

For estimation, I use equation (5). Y_{mtd} is either $P(\text{Per Housing Unit Damage} > 0)$ or $log(\text{Per Housing Unit Damage})$ conditional on having positive damage for community m at year t for data stack d. Because the damage variable has a point mass at zero and long right tails, I estimate the extensive and intensive margin effect separately, although an emphasis is given to the extensive margin due to greater generalizability—only a small fraction of communities experience repeated damage—and higher statistical power.

Similar to Section 4, I use a stacked approach for the estimation and thus every term in equation (5) has a subscript representing the stack d. Moreover, as I exploit the timing of the disclosure requirement for identification, not-yet-treated states form the control group. One important difference is that I run the stacked DD (i.e., ignoring differences in treatment intensity) for tractability.[28]

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

Equation (5) also includes year $\times$ stack (ω_{td}) and community $\times$ stack (θ_{md}) fixed effects, to control for overall time trend and unobserved community characteristics. I use 20 years of observation for each state around the disclosure policy change year. Throughout the analysis, I use spatial-HAC standard errors that allow spatial correlation of up to 500 miles (Newey and West 1987, Conley 1999).[29] Similar to Section 4, when the variance-covariance matrix is not positive-semidefinite, I use eigendecomposition and convert any negative eigenvalue(s) to zero following Cameron et al. (2011).

Before further proceeding, it is worth 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

[28] I exploit the differences in treatment intensity by estimating heterogeneous treatment effects.

[29] Weights in this matrix are uniform up to that cutoff distance.

limitations for estimating a more aggregate level 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. Thus, estimating flood damage at an aggregate level using a depth-damage function requires a detailed hydraulic study, which translates weather events into an inundation level for each property (Scawthorn et al. 2006). However, these studies are costly and thus many communities either lack access to them or are reliant on outdated versions (FEMA 2005, Bakkensen and Ma 2020, Weill 2021).

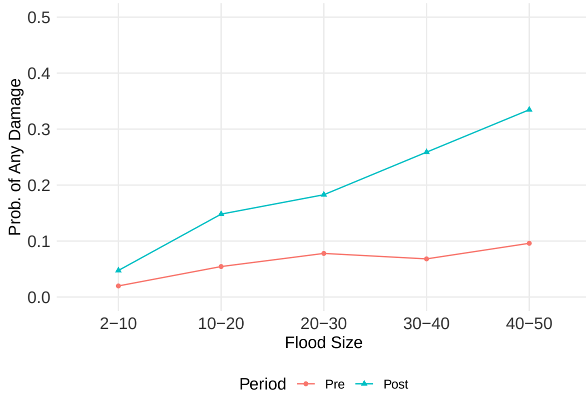
Second, and presumably most importantly, these studies typically lack the capability to account for adaptations at the property level, which is likely to cause biases in the estimated damage function. In theory, this issue can be addressed by (1) modeling how various defensive measures such as property 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.^[For instance, commonly considered resistance parameters in the engineering literature are building type, building material, and usage of building (Merz et al. 2010). It would be extremely hard to tell differences in the adaptation level across different single family houses using these crude measures.} This is a major drawback of depth-damage functions given that the main purpose of constructing a damage function is reliable flood damage projection, which is a key ingredient for the cost-benefit analysis of any flood management policy.

This paper takes a "reduced-form" approach and overcomes these issues. By directly relating flood size, which is measured at the community level using the water gauge records, to the community-level flood damage, this approach can be applied even in areas without up-to-date hydraulic studies. Moreover, the community level damage metric readily factors in the number of properties damaged and the impact of any pre-existing adaptation measures.

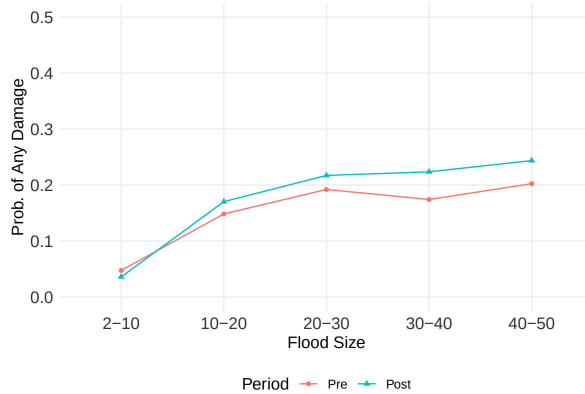
5.2 Findings

In Figure 5.1, I plot the damage functions for (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 periods (red line) in panel (a) and (b), respectively. The dependent variable in Figure 5.1 is the probability of having any damage, and thus the vertical axis indicates the additional

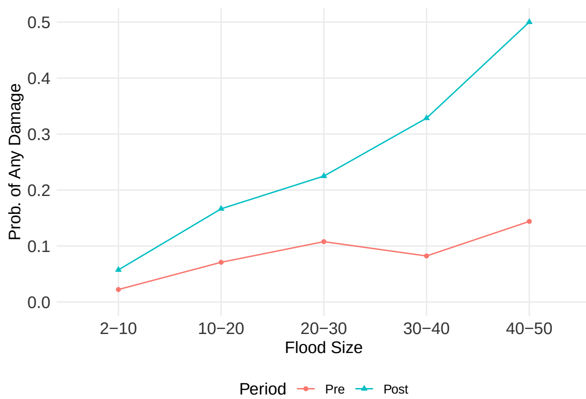
[30] Appendix Figure C.8 reproduces Figure 5.1 with a 95 percent confidence interval.



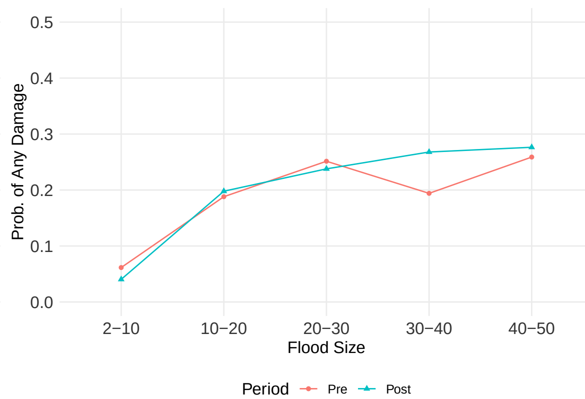
(a) Control Group (All Communities)



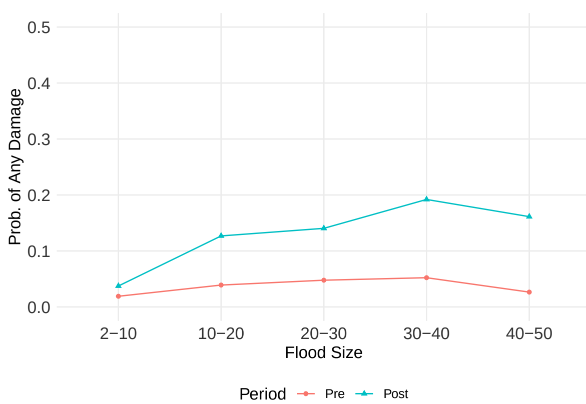
(b) Treatment Group (All Communities)



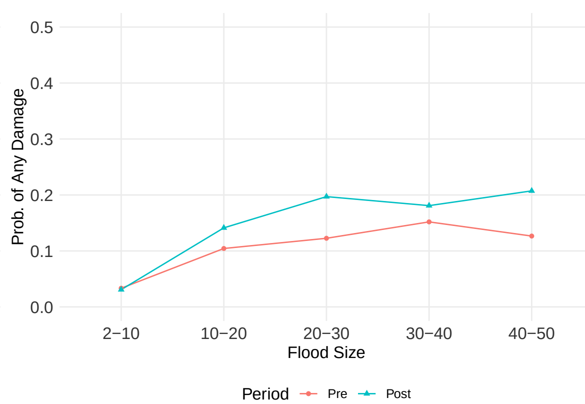
(c) Control Group (High Risk Communities)



(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 a set of 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 C.8 reproduces Figure 5.1 with corresponding 95% confidence intervals.

probability of any damage when the baseline flood (k=1–2) is replaced by a flood of size k.

Before discussing the disclosure policy effect, I test if the estimated damage function is valid. For that, I limit the attention to the pre-treatment period (red lines in Figure 5.1 (a) and (b)). The red lines indicate that as flood size increases, the probability of having any flood damage increases monotonically. The largest bin (k=40–50) suggests that on average, communities with a flood of size 40–50 have 10 percentage points (panel (a)) and 20 percentage points (panel (b)) higher chances of having any damage in comparison to the baseline floods.

Another source of a validity test is the heterogeneity in the damage function. Even when faced with floods of the same size (as defined by community-specific recurrence intervals), communities with higher risk should exhibit higher levels of damage. To illustrate this, consider two communities, A and B, characterized by starkly different risk profiles: A is entirely situated within the SFHA, an area expected to be inundated during a 100-year flood event, 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, whereas B would remain unaffected. Indeed, in Figure 5.1 (c)–(f), high risk communities (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.

Table 5.1 highlights the impact of the disclosure requirement on flood damage. For the interest of space, I only report $\hat{\beta}_4^k$ from equation (5) but a full set of coefficients are in Appendix Table C.4. In column (1), 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 each k for the communities in the disclosed states relative to the ones in the not-yet-disclosed states.[31] The damage reduction effect can be verified visually as well: the distance between red and blue lines in Figure 5.1 is much larger in panel (a) than (b).

To put the coefficients in Table 5.1 in context, I take a probability-weighted average of them as equation (6). Here, $Pr(K = k)$ is the likelihood of flood occurrence for bin size k each year.[32] Also, because equation (5) is a linear probability model, I can conveniently interpret equation (6) as the

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

[32] 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).

Table 5.1: Effect of Disclosure Requirement on Flood Damage

	(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	-0.025**	-0.036***	-0.013	-0.014
	(0.01)	(0.012)	(0.009)	(0.074)
Dep.Var	$P(Y>0)$	$P(Y>0)$	$P(Y>0)$	$log(Y)$
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 damage). 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. $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$.

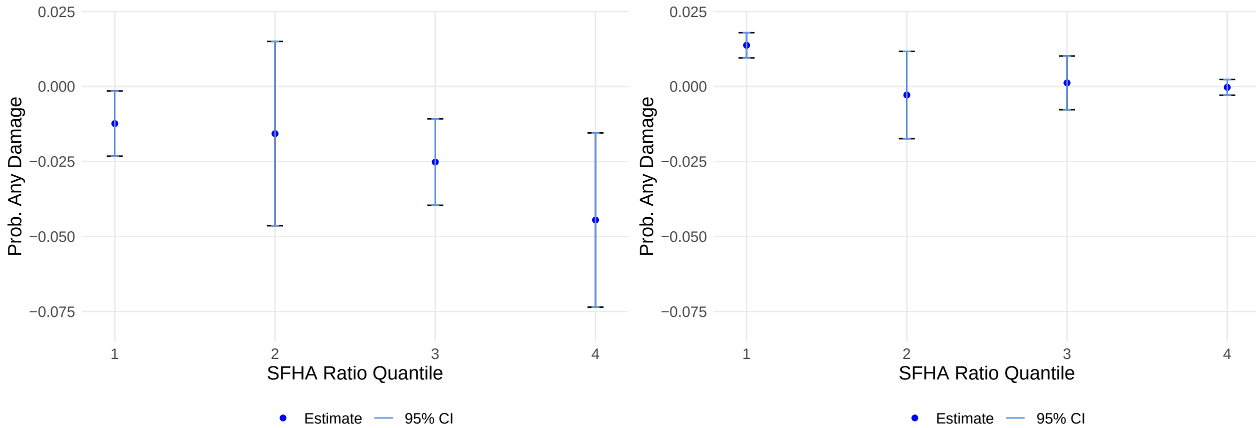
estimated change in the probability of having any flood damage due to the disclosure requirement.

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

In Table 5.1, I report that the annualized reduction in the probability of having any damage is 2.5 percentage points. When I compare this with the average probability of having any damage conditional on exposure to a flood of size 2 or larger (7.4 percentage points), the effect size is a 33 percent reduction from the benchmark.

In columns (2) and (3), I split the sample into communities above and below the median fraction of the area covered by an SFHA to explore the heterogeneous treatment effect. Because the disclosure targets properties in an SFHA, the policy effect should be driven by the high-SFHA communities. Indeed, the annualized effect from equation (6) is three times larger for high-SFHA communities than low-SFHA communities (–3.6 vs. –1.3 percentage points).[33] Figure 5.1 (c)–(f) mirrors this: the distance between the red and blue lines between the control and treated groups is much larger for

[33]These are 37.5 percent and 23 percent reduction from the respective benchmark probabilities.



(a) Disclosed States

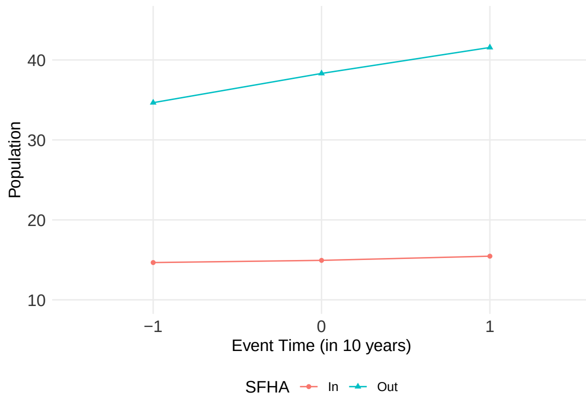
(b) Placebo States

Figure 5.2: Annual Expected Damage Reduction Effect by the Ratio of SFHA. These figures plot the annual expected damage reduction effect from disclosure against SFHA ratios. I estimated equation (5) for subsample of communities in different quantile of SFHA ratio, and aggregated coefficients using equation (6).

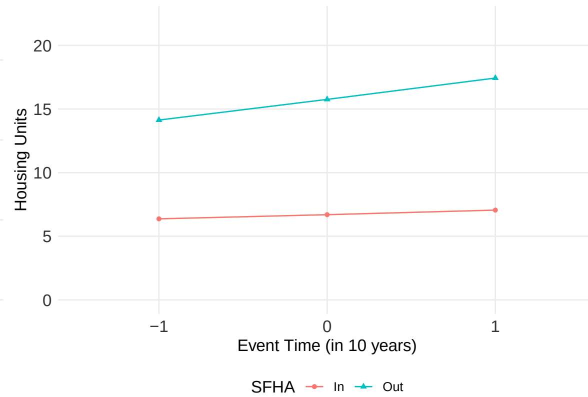
the high-SFHA communities (panels (c)–(d)). Moreover, Figure 5.2 (a) further splits columns (2) and (3) and presents the annualized effect for four groups of communities with differential SFHA ratios, which clearly shows monotonically increasing magnitude in the SFHA ratio.

Column (4) reports the intensive margin effect, where the dependent variable is the log of per housing unit damage. Because the sample for this exercise is restricted to a community-year with positive damage, the model does not have the power to draw a strong conclusion. Still, I find suggestive evidence that the disclosure reduces damage for communities with repetitive flood events.

Taken together, Table 5.1 indicates that the disclosure curbs flood damage for the treated communities relative to the control communities, a finding consistent with the reduction in flood risk exposure in Section 4.2. To investigate the mechanism further, in Figure 5.3 (a), I plot the average population for census blocks inside and outside of an SFHA—the same blocks used to generate Table 4.1 column (2)— in event time, which is defined as -1 for pre-treatment periods, 0 for post-treatment periods up to 9 years, and 1 for post-treatment period beyond 10 years. The figure illustrates that the relative population of SFHA blocks is decreasing because of an increasing population of non-SFHA blocks rather than a shrinking population of SFHA blocks. Similarly, in panel (b), I find a rapid expansion of housing units in non-SFHA blocks and a stagnation for SFHA blocks. These empirical patterns suggest that the damage reduction effect is driven by diverted in-migration (and resulting suppressed development) rather than active out-migration from SFHA areas.



(a) Population



(b) Housing Units

Figure 5.3: Population and Housing Unit Trends in Event Time. These figures plot the (a) average population and (b) 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 or more years.

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

Rosuteness check. I checked the robustness of the policy effect by conducting a placebo test using five states (ID, ME, MN, NH, and VA) that had implemented a disclosure policy but without a question on the flood risk. In Appendix Table C.6, I estimate a version of equation (5) with coarser flood bins based on these five states.[34] In columns (1) to (3), the coefficients suggest that a disclosure without flood risk information does not reduce the probability of flood damage at all. The estimates are statistically insignificant and economically small, which are consistent with Figure 5.2 (b): the effect is zero for all four groups of communities with varying SFHA exposures.

Another robustness check is in Figure C.9, which plots the marginal effect of disclosure for larger floods in event time. Similar to the placebo analysis, I use coarser flood bins. Also, I impose an endpoint restriction at -5 and 4. It shows no pre-trend and a clear reduction in the probability of flood damage after the disclosure. This effect corresponds to a flatter damage function in Figure 5.1.

Before concluding, in Appendix Table C.5, I report how the results in Appendix Table C.4 change

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

when I add never-treated states to the control group.[35] A comparison of the two tables produces two observations. First, the shape of the damage functions coincide. For instance, the coefficients for the control group in the pre-treatment period (Flood Size terms) monotonically increase in flood size in both tables. Similarly, the Post × Flood Size terms are positive and increasing in flood size for both tables. Second, the magnitudes of the interaction terms, which capture the treatment effect, in Appendix Table C.5 is 20–40 percent of those in Appendix Table C.4. Such an attenuated effect is consistent with the discussion in Section 2.1 that a non-trivial number of home buyers in these non-disclosure states are likely to have received some form of information even in the absence of the required disclosure policy. As such, I use not-yet-treated states as a primary control group.

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. A prevalent policy prescription in the US has been structural flood water control, namely, adding more physical structures. However, this approach discourages adaptation—it rather attracts more people to areas with flood risk, by distorting location choices.

In this paper, I study whether alleviating information friction regarding flood risk in the housing market can be a more effective way to foster adaptation by exploiting plausibly exogenous variations created by the disclosure requirement. I explore if and how home buyers respond to the disclosure policy and investigate its implications for flood damage. The results show that the disclosure requirement reduces the population and increases the vacancy rate in high-risk areas. As a smaller number of households are exposed to flood risk, the probability of having any flood damage also decreases by 2.5 percentage points (or 33 percent from the average probability).

The findings of this paper have important policy implications as disclosure is getting more attention as a flood risk management tool. For instance, FEMA has recently proposed an NFIP reform that conditions a community's flood insurance eligibility on the implementation of a mandatory flood risk disclosure (U.S. Department of Homeland Security 2022). My analysis shows that such a policy can facilitate voluntary adaptation by helping home buyers make a more informed choice.

[35] Out of 22 never-treated states, 5 are placebos. For this exercise, I consider placebo states as never-treated states.

References

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

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.

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].

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.

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

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.

Flavelle, C. 2017. The Nightmare Scenario for Florida's Coastal Homeowners. Bloomberg.

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., 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. THE 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.

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.

Merz, B., H. Kreibich, R. Schwarze, and A. Thieken. 2010. Assessment of economic flood damage. Natural Hazards and Earth System Sciences 10:1697–1724.

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.

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.

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.

Satija, N., K. Collier, and A. Shaw. 2017. Houston officials let developers build homes inside reservoirs. But no one warned buyers. The Texas Tribune.

Scawthorn, C., P. Flores, N. Blais, H. Seligson, E. Tate, S. Chang, E. Mifflin, W. Thomas, J. Murphy, C. Jones, and M. Lawrence. 2006. HAZUS-MH Flood Loss Estimation Methodology. II. Damage and Loss Assessment. Natural Hazards Review 7:72–81.

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. THE 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.

Wing, O. E. J., P. D. Bates, A. M. Smith, C. C. Sampson, K. A. Johnson, J. Fargione, and P. Morefield. 2018. Estimates of present and future flood risk in the conterminous United States. Environ. Res. Lett. 13.

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 data 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 meteological 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 measure 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, snowmelt, 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).[36]

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-

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

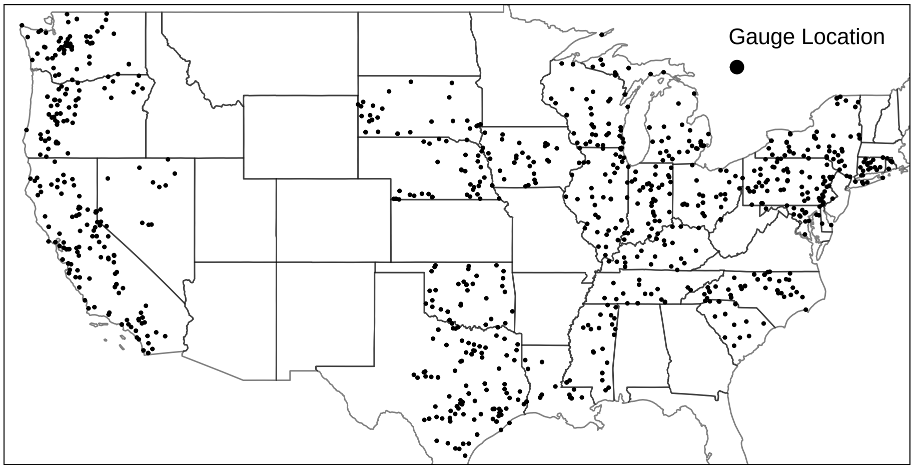


Figure A.1: The Distribution of a Sample of USGS/NOAA Gauges

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).[37] I use three different levels of geographic units, namely state, HUC4, and HUC2 and separately estimate Fuller coefficients. 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]

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

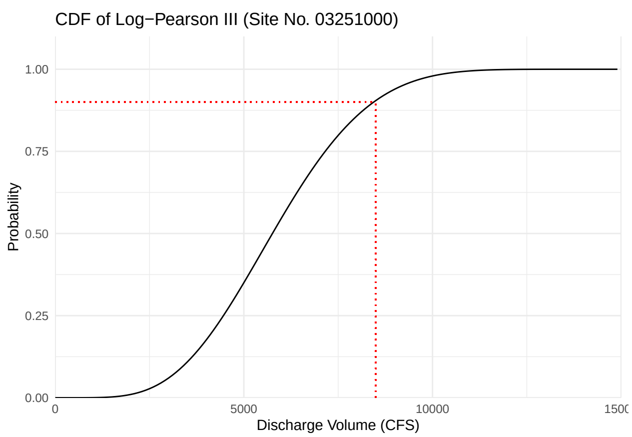
Now, by converting the estimated instantaneous peak flow to the quantile of the estimated Log-

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

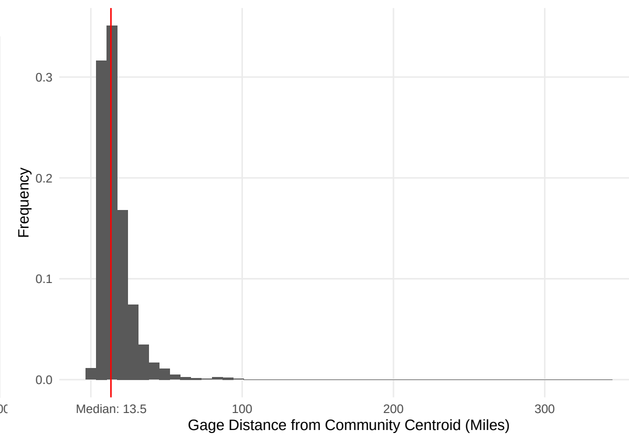
[38] Practically, I apply the following hierarchy among state, HUC4, and HUC2 models: (1) When a site has the best match (which means that a site has both daily mean flow and instantaneous flow records), I use it. (2) If a site does not have site-specific match (meaning that this site did not have instantaneous 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).

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 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 gauges and community centroid. Over 90% of them are within 20 miles with the median distance 13.5 miles.

Appendix Figure A.2 (a) illustrates step 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.

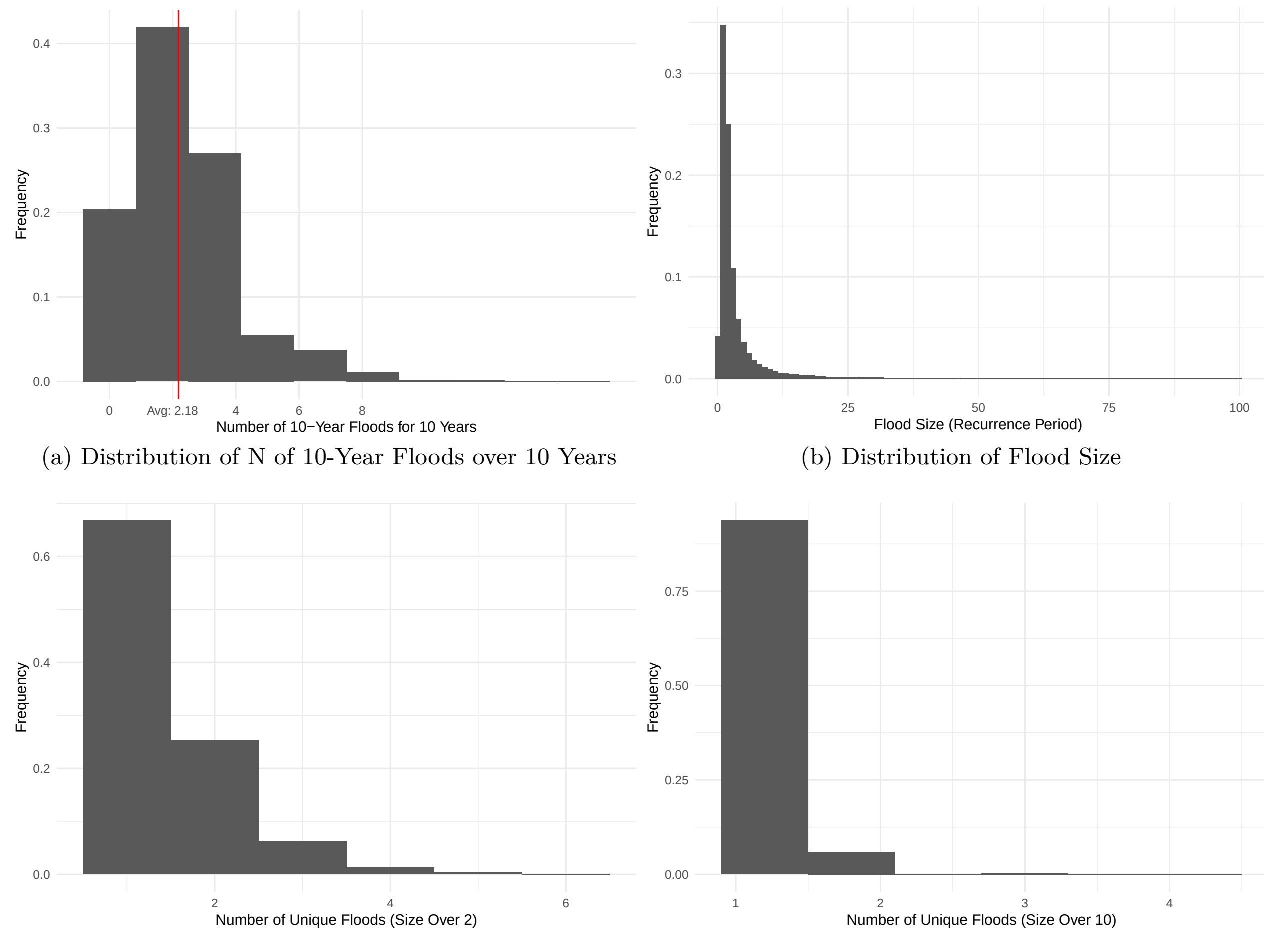
Note, because the USGS gauge stations rarely cover coastal areas, I add 45 additional NOAA sites to the gauge station data. Zervas (2013) documents the flood threshold for the entire NOAA sites by fitting GEV distribution, so I adopt them directly. NOAA water level data are retrieved using the R package "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, solely relying 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 is going to happen 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 the 10-year flood 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) Distribution of N of 10-Year Floods over 10 Years (b) Distribution of 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 on most communities had 1 or 2 10-year floods over the 20 years whereas 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 well documented in the literature, the histogram follows a log-normal distribution, where as the flood size increases the frequency decreases as an inverse power function (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-year 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 Yrs Flood	10 Yrs Flood	50 Yrs Flood	100 Yrs 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 $j \in \{\text{minor, moderate, major}\}$.

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

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.

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.

[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.

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 works. 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 in Table 3.1, 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 due to 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: 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 works, I can indirectly test whether the disclosure requirement was effectively raising home buyers' flood risk awareness.

For housing prices, I use the Zillow Transaction and Assessment Database (ZTRAX).[40] 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.[41]

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 control 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 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. Importantly, Post and Disclosure dummies are defined for each stack.

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 B.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 occasional flood map updates, which can coincide with the disclosure policy change. Specifically, I repeat column (1) after removing 6 percent of 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.

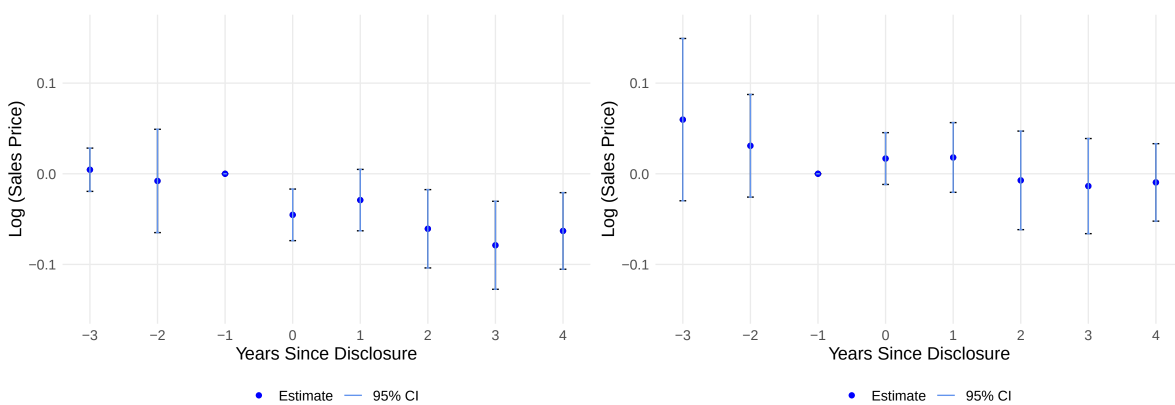
[40] I thank Eyal Frank for his generous help with data access.

[41] 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 (before CPI adjustment) to be between $10,000 and $100,000,000.

Table B.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. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.



(a) Early vs. Late

(b) Ever vs. Never

Figure B.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. Panel (a) uses late-treated states as a control group whereas panel (b) uses never-treated states as a control group. Standard errors are clustered on state. See the text for additional details.

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

Importantly, this is a notable contrast to the figure in Panel (b), which has included never-treated (namely, states that have not implemented a home seller disclosure requirement on flood risk) states as a control group. In contrast to Panel (a), the plot exhibits a clear pre-trend. Further, the impact of the disclosure policy on the treated properties is substantially attenuated, which is consistent with an earlier discussion that home buyers in the non-disclosure states are also exposed to information treatment from realtor association-driven disclosure efforts.

Finally, it is worth pointing out that the violation of the SUTVA assumption would not be a major concern in this setting because such an effect is likely to be small. The number of properties inside of the SFHA is equal to or less than 10 percent (average: 4.8 percent) for every state except Louisiana. Thus even if home buyers sort into the non-SFHA area after the disclosure policy, the effect would not be large enough to change the counterfactual price of the non-SFHA properties. Further, even if the SUTVA assumption is violated, the estimated housing price change can still show that the disclosure policy is effective and would be a policy relevant parameter because it reflects an actual housing market response to the flood risk information.

C Appendix C: Additional Tables and Figures

Table C.1: Effect of Discosure Requirement on Household Responses (Placebo States)

	(1)	(2)	(3)
SFHA × Post	.004	.031	.004
	(.004)	(.062)	(.010)
D.V	Prob (Pop > 0)	Log Population	Vacancy Rate
Avg D.V. (Within BW)	0.602		0.095
Bandwidth	494	406	477
Num. obs.	357459	188304	209918

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. 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 4.2.

Table C.2: Effect of Discosure Requirement on Population and Vacancy Rate (Donut Specification)

	(1)	(2)	(3)	(4)	(5)	(6)
SFHA × Post	$-.011^{***}$	$-.077^{**}$	$.013^{***}$	$-.007^{*}$	$-.079^{**}$	$.015^{**}$
	(.003)	(.035)	(.005)	(.003)	(.038)	(.006)
D.V	P(Pop > 0)	Log Population	Vacancy Rate	P(Pop > 0)	Log Population	Vacancy Rate
Avg D.V. (Within BW)	0.691		0.095	0.703		0.094
Doughnut Size	20	20	20	40	40	40
Num. obs.	1209186	1499805	1155519	966183	1343715	1000621

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. 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 4.2.

Table C.3: Effect of Discosure Requirement on Demographic Compositions

	(1)	(2)	(3)	(4)
High SFHA × Disclosure × Post	$-.034^{***}$	$-.770^{***}$	.512	$-.004$
	(.012)	(.268)	(.376)	(.657)
D.V	log(Median income)	(%) 65+	(%) BA	(%) Black
Avg D.V.		12.3	24	14.8
Year FE	X	X	X	X
Tract FE	X	X	X	X
Num. obs.	73702	73702	73702	73702

Note: This table is produced from equation (2) using the decennial census data in 1990 and 2000. Outcome variables and their average values can be found in the table text. 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 4.2.

Table C.4: Effect of Disclosure Requirement on Flood Damage

	(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)
Dep.Var	$P(Y > 0)$	$P(Y > 0)$	$P(Y > 0)$	$log(Y)$
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 variables in columns (1) to (3) are the probability of having positive flood damage (per housing unit damage). 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. Dependent variables in columns (4) is log transformed per housing unit damage conditional on non-zero damage. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 5.2.

Table C.5: Effect of Disclosure Requirement on Flood Damage (Including Never Disclosed States)

	(1)	(2)	(3)	(4)
Flood Size 2-10	.026***	.035***	.019***	.146
	(.005)	(.006)	(.004)	(.130)
Flood Size 10-20	.091***	.116***	.064***	.531***
	(.009)	(.015)	(.008)	(.192)
Flood Size 20-30	.132***	.169***	.094***	.841***
	(.012)	(.015)	(.017)	(.296)
Flood Size 30-40	.154***	.179***	.123***	1.394***
	(.022)	(.029)	(.017)	(.133)
Flood Size 40-50	.191***	.251***	.119***	1.776***
	(.022)	(.027)	(.019)	(.199)
Disclosure × Size 2-10	.019**	.022**	.016***	.046
	(.009)	(.011)	(.006)	(.200)
Disclosure × Size 10-20	.058***	.066***	.049***	.534***
	(.017)	(.023)	(.012)	(.114)
Disclosure × Size 20-30	.060***	.084***	.034**	.995***
	(.020)	(.031)	(.015)	(.241)
Disclosure × Size 30-40	.028	.022	.037	.271***
	(.021)	(.023)	(.030)	(.104)
Disclosure × Size 40-50	.025	.031	.015	.052
	(.022)	(.033)	(.023)	(.328)
Post × Size 2-10	.020**	.025***	.013*	.440***
	(.008)	(.009)	(.007)	(.131)
Post × Size 10-20	.062**	.062**	.059**	.563***
	(.026)	(.027)	(.026)	(.105)
Post × Size 20-30	.080**	.078*	.082***	.593**
	(.037)	(.046)	(.030)	(.234)
Post × Size 30-40	.103***	.100***	.107**	.373***
	(.038)	(.034)	(.048)	(.104)
Post × Size 40-50	.119**	.119**	.117***	.374**
	(.047)	(.058)	(.033)	(.165)
Post × Disclosure × Size 2-10	−.016	−.022	−.009	.113
	(.011)	(.015)	(.010)	(.158)
Post × Disclosure × Size 10-20	−.017	−.022	−.005	−.174*
	(.020)	(.032)	(.020)	(.101)
Post × Disclosure × Size 20-30	−.028	−.057	.010	−.599**
	(.030)	(.038)	(.036)	(.234)
Post × Disclosure × Size 30-40	−.028	.008	−.055	.095
	(.046)	(.042)	(.052)	(.240)
Post × Disclosure × Size 40-50	−.055	−.068	−.020	−.183
	(.040)	(.050)	(.050)	(.260)
Dep.Var	$P(Y > 0)$	$P(Y > 0)$	$P(Y > 0)$	$log(Y)$
Sample	All	High SFHA	Low SFHA	Damage > 0
Year × Stack FE	X	X	X	X
Community × Stack FE	X	X	X	X
Num. obs.	1084023	542442	541581	57343

Note: This table repeats Appendix Table C.4 after including 22 non-disclosure states as control group. The dependent variables in columns (1) to (3) are the probability of having any damage per housing unit. Dependent variables in columns (4) is logged per housing unit damage conditional on non-zero damage. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference. $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Back to 5.2.

Table C.6: Effect of Disclosure Requirement on Flood Damage (Placebo States)

	(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)
Dep.Var	$P(Y > 0)$	$P(Y > 0)$	$P(Y > 0)$
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 positive flood damage (per housing unit damage). 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.

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

Back to 2.1.

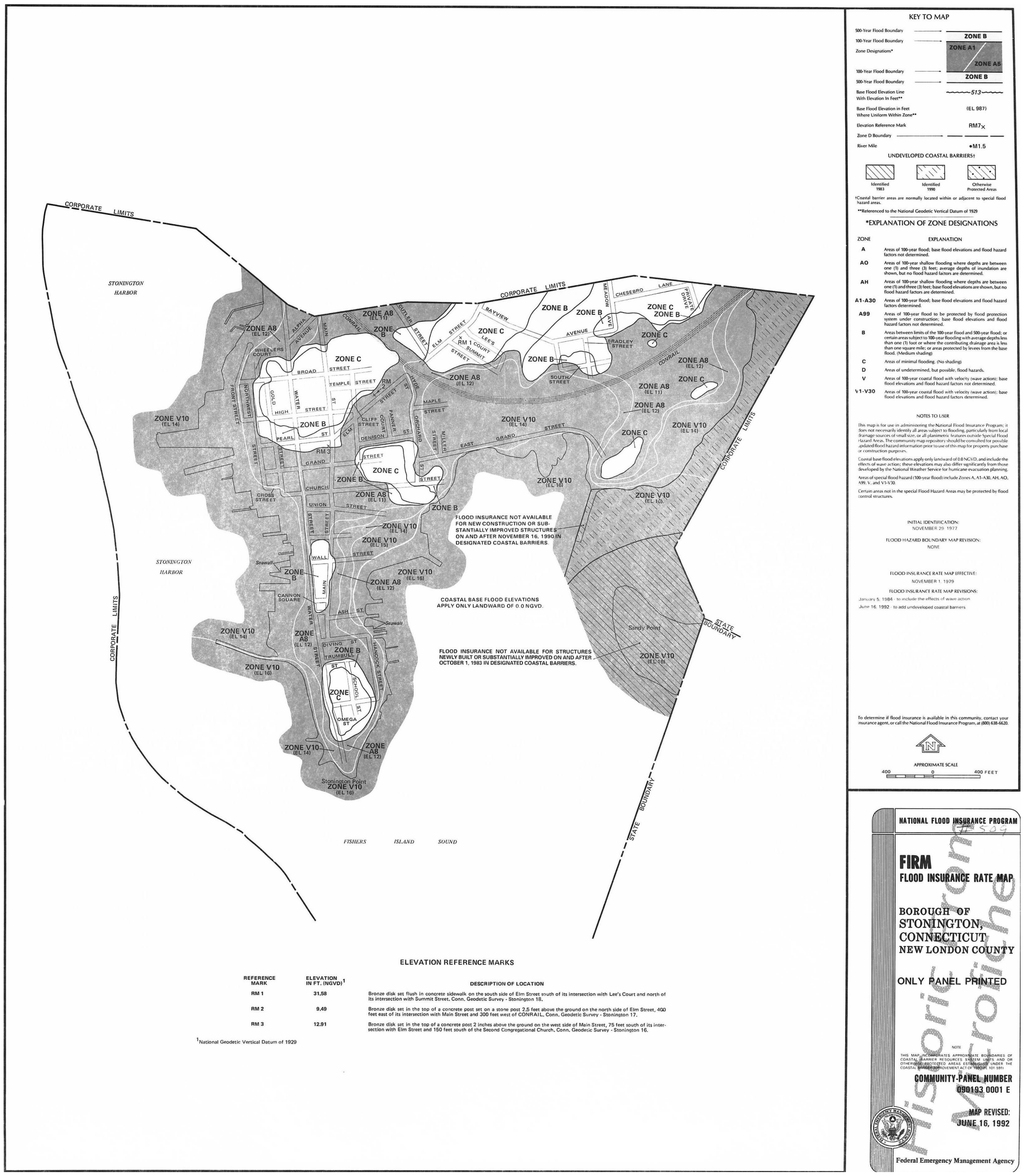


Figure C.2: Sample Flood Insurance Rate Map (Borough of Stonington, CT)

Back to 2.2.

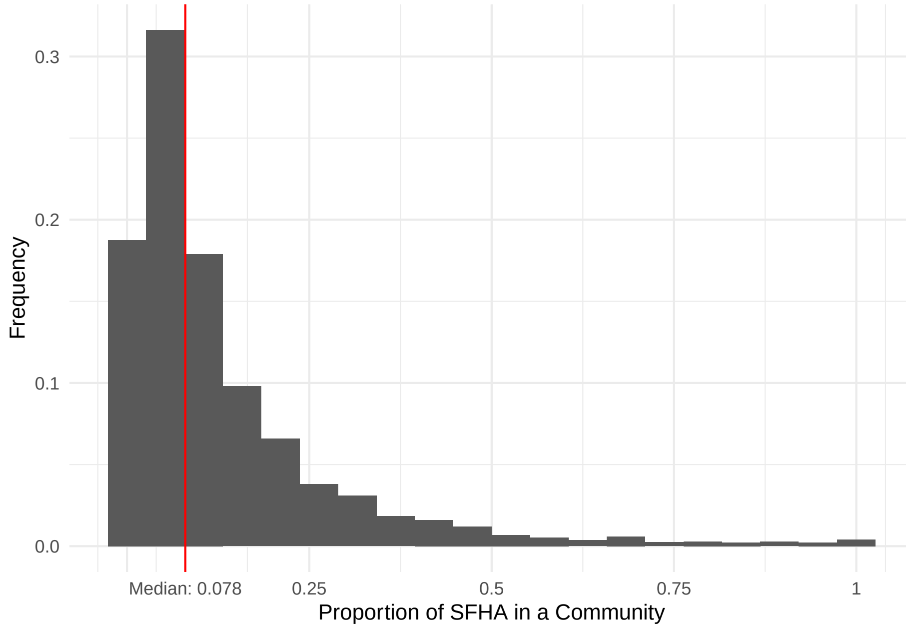
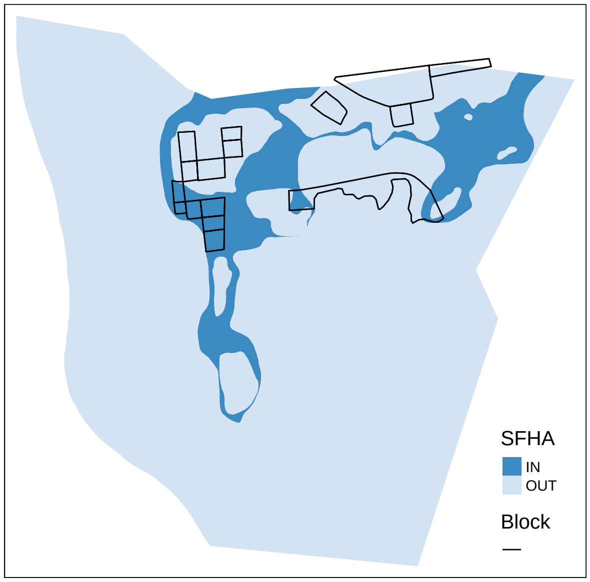
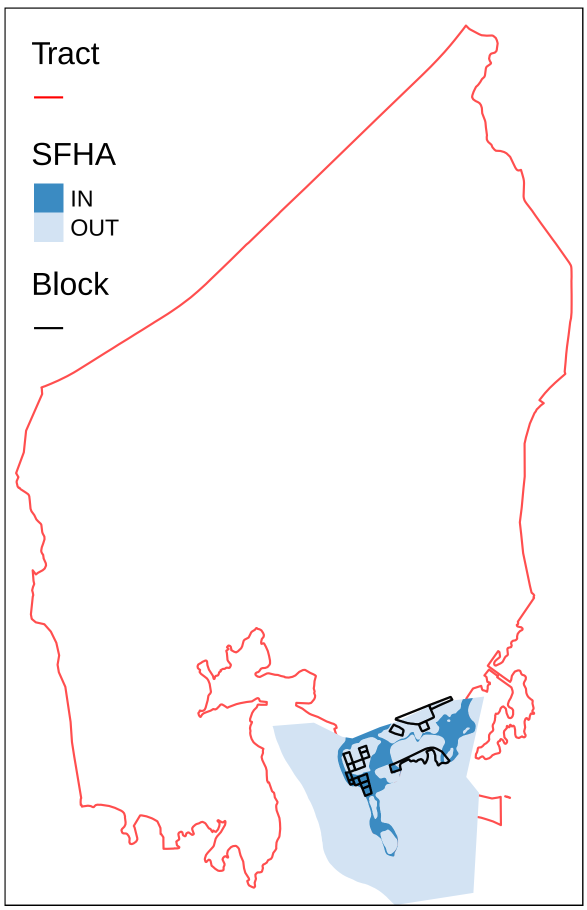


Figure C.3: Histogram of the Proportion of the SFHA at the Community Level. The plot shows the distribution of the SFHA ratio for the 8,194 communities that are on the Q3 map and in the 26 ever-disclosed states.

Back to 2.2.



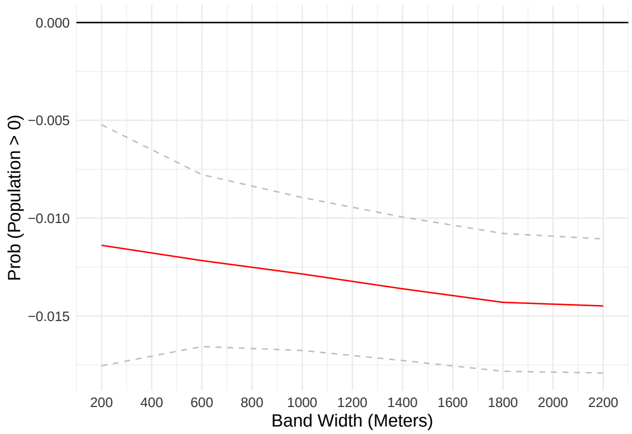
(a) Census Block vs. SFHA Status



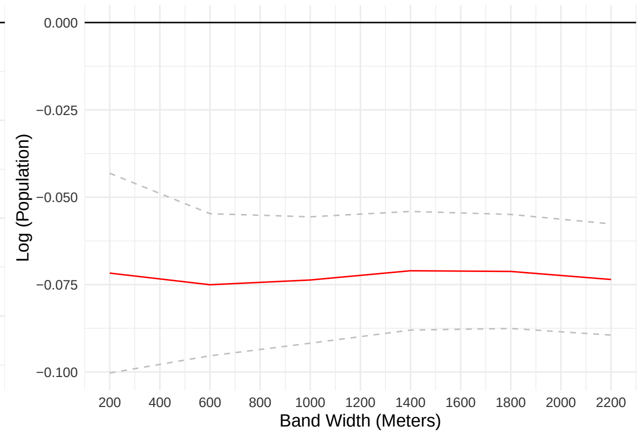
(b) Census Tract vs. SFHA Status

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

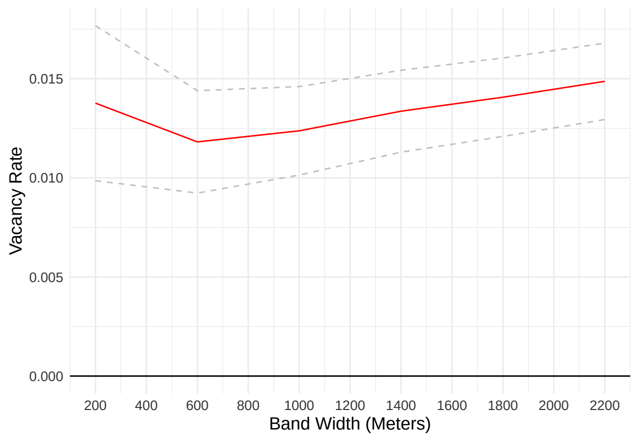
Back to 4.1.



(a) Prob. of Any Population



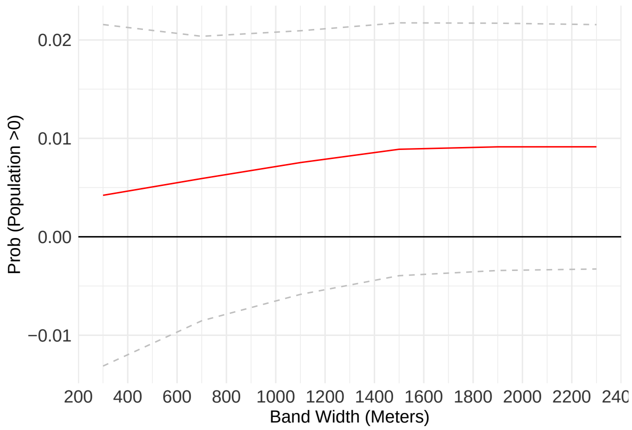
(b) Log Population



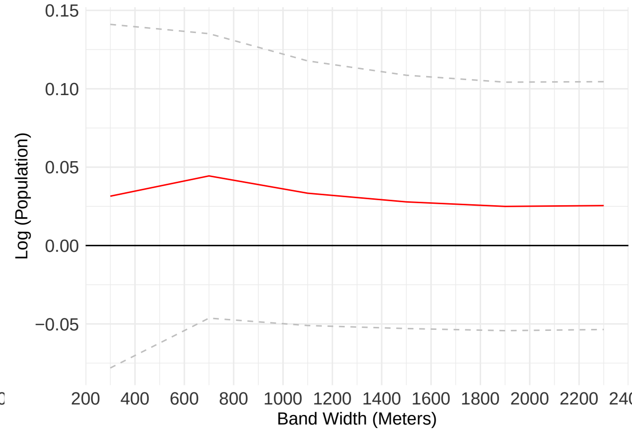
(c) Vacancy Rate

Figure C.5: The Effect of the Disclosure Requirement on Population and Vacancy Rate for Different Bandwidths. The figure plots $\hat{\delta_6}$ from equation (1) for a range of bandwidths. The level of observation is census block, which is the smallest census geographical unit. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference

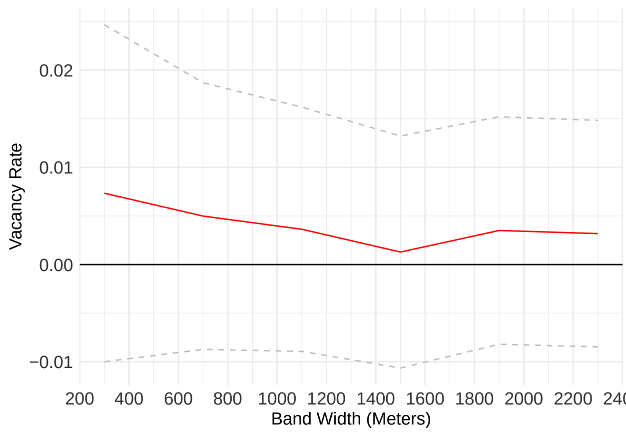
Back to 4.2.



(a) Prob. of Any Population



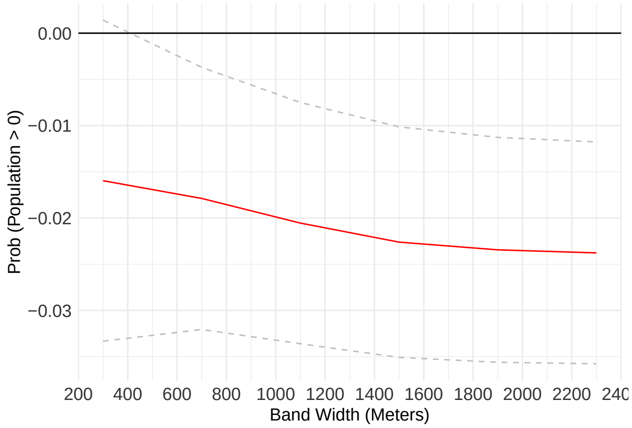
(b) Log Population



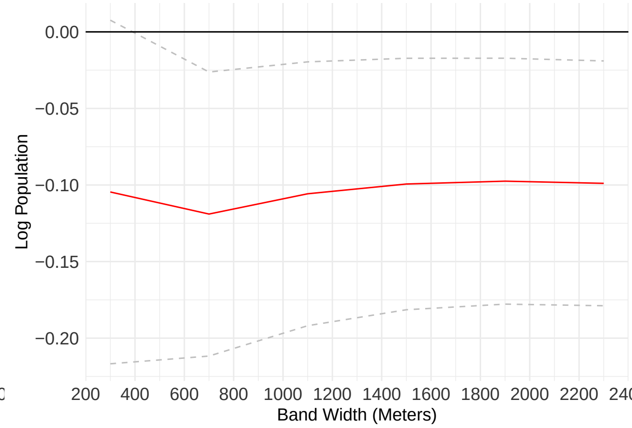
(c) Vacancy Rate

Figure C.6: The Effect of the Disclosure Requirement on Population and Vacancy Rate for Different Bandwidths (Placebo States). The figure plots $\hat{\delta_6}$ 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. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference.

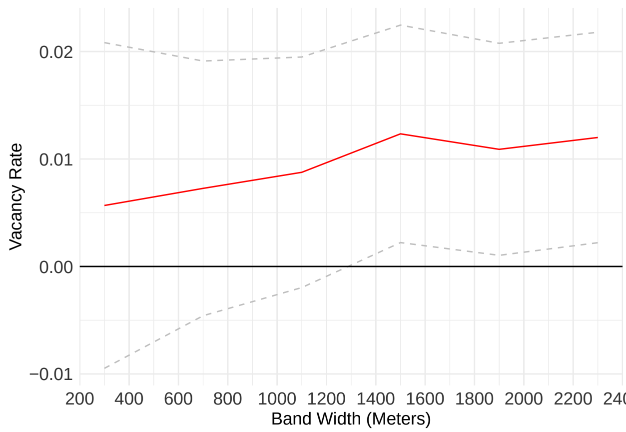
Back to 4.2.



(a) Prob. of Any Population



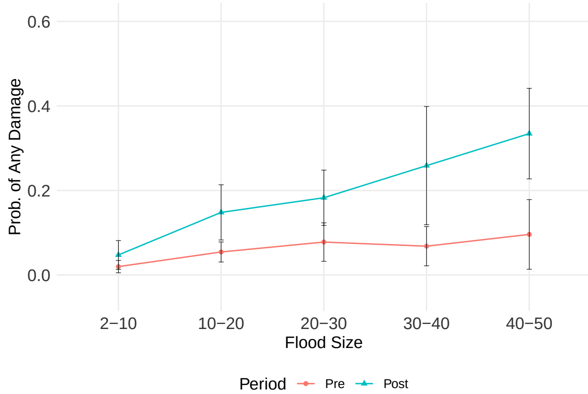
(b) Log Population



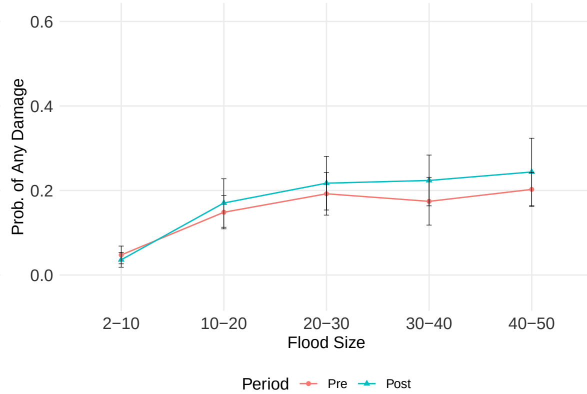
(c) Vacancy Rate

Figure C.7: The Difference in Diff-in-Disc Estimates Between Treated and Placebo States. These figures plot the difference in diff-in-disc estimates between the treated and placebo states for a range of bandwidths. The level of observation is census block, which is the smallest census geographical unit. Spatial-HAC standard errors that allow spatial correlation of up to 500 miles are estimated for inference. See the text for additional details.

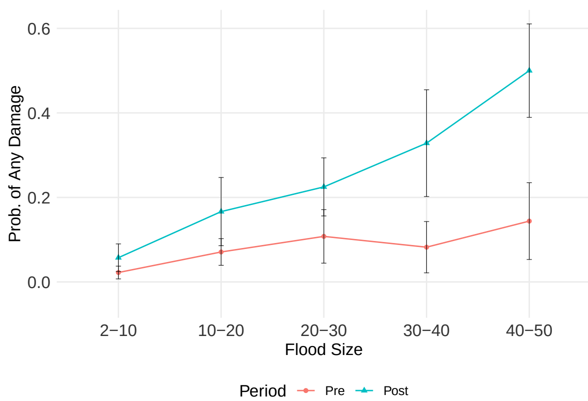
Back to 4.2.



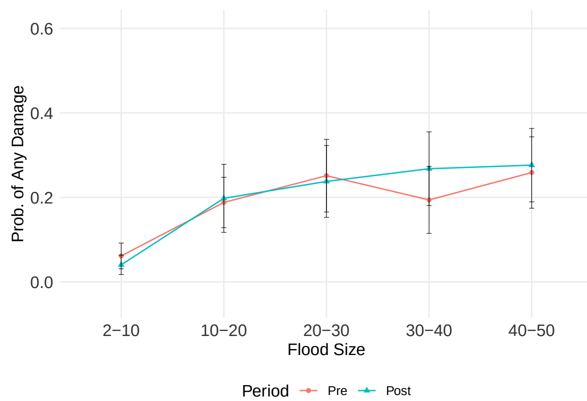
(a) Control Group (All Communities)



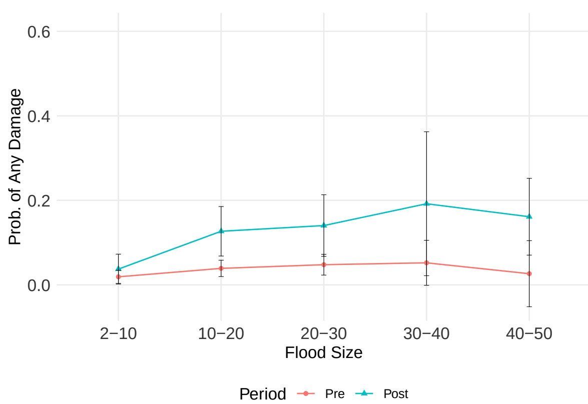
(b) Treatment Group (All Communities)



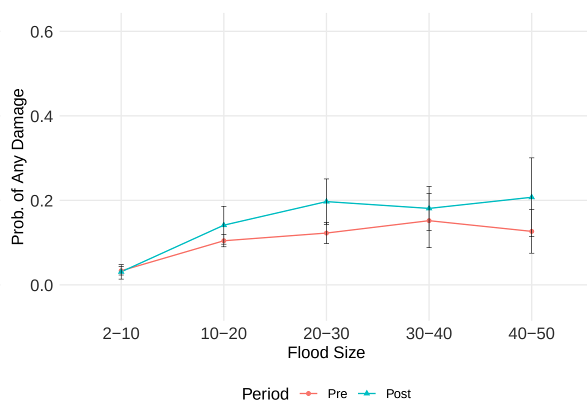
(c) Control Group (High Risk Communities)



(d) Treatment Group (High Risk Communities)



(e) Control Group (Low Risk Communities)



(f) Treatment Group (Low Risk Communities)

Figure C.8: The Effect of Disclosure on the Damage Function with Corresponding 95% Confidence Intervals. These plots reproduce Figure 5.1 with confidence intervals.

Back to 5.2.

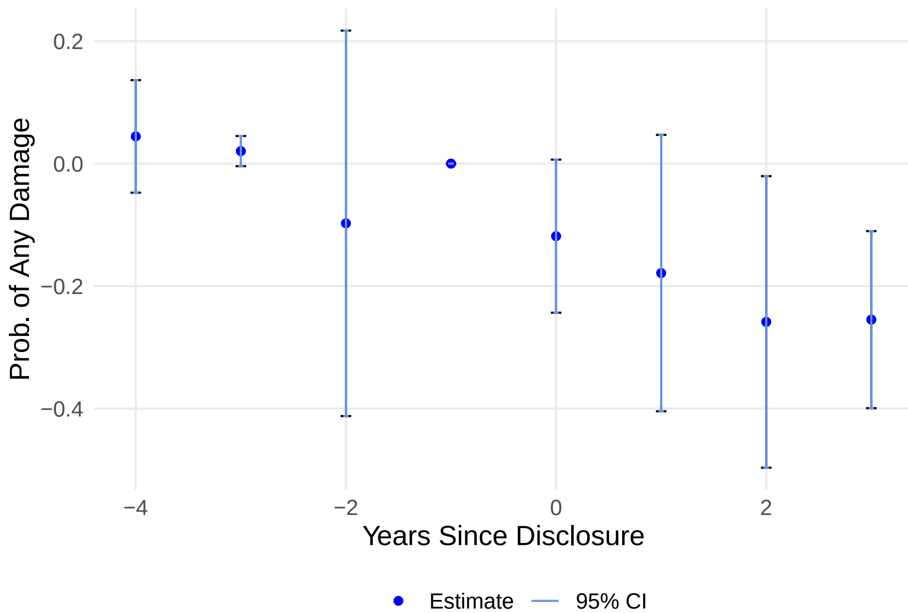


Figure C.9: The Effect of Disclosure on the Damage in Event Time. This figure depicts $\hat{\beta_{4,t}^{30-50}}$ for flood size of 30-50 in event time t where the dependent variable is probability of having any damage. The error bar represents the 95% confidence interval.

Back to 5.2.