The Price and Distributional Impact of Flood Risk Disclosure: Evidence from US Housing Platforms *

Steve Billings † Sophie Calder-Wang ‡ Weiling Liu§

February 18, 2026

Preliminary

Abstract

We study the equilibrium price and distributional impact of flood risk disclosure on U.S. housing platforms. Starting 2020, major online platforms began displaying property-level flood risk assessments from First Street Technology. Using a regression discontinuity design that exploits discrete cutoffs in the disclosed risk categories, we find that homes labeled with "extreme" flood risk experience a 3.3% price discount and stay on the market for 6.8 days longer, relative to those labeled as "severe" risk. This effect is particularly pronounced in coastal areas outside of Special Flood Hazard Areas (SFHAs). The disclosure also creates significant distributional impacts: buyers of extreme-risk homes have 5.3% lower incomes and are 1.8 percentage points more likely to hold FHA loans. They are also 2 years older on average, though we find no significant effects on race or political affiliation. Using a discrete choice model, we find that the reallocation of flood risk to lower-income households is primarily driven by prices changes rather than heterogeneous preferences for flood risk. Our findings highlight that disclosure enables more informed decision-making, but it also shifts environmental risk toward financially constrained buyers.

Keywords: Climate Risk; Disclosure; Online Platforms; House Prices; Distributional Impacts.

*Support for this research was provided by the Zell-Lurie Real Estate Center at the Wharton School. Vivian (Mengmeng) Wang and Yu Hui provided excellent research assistance. All errors are our own.

†Leeds School of Business, University of Colorado at Boulder *stephen.billings@colorado.edu*

‡The Wharton School, University of Pennsylvania *sophiecw@wharton.upenn.edu*

§D'Amore-McKim School of Business, Northeastern University *w.liu@northeastern.edu*

1 Introduction

Flood-related natural disasters, such as Hurricane Katrina (2005), Sandy (2012), Harvey, Irma, and Maria (2017), Dorian (2019), Ida (2021), Ian (2022), and Milton (2024), have caused devastating losses of life, homes, and infrastructure. These events are of paramount concern for homeowners, investors, and policymakers due to their enormous economic impact. A report by the Joint Economic Committee estimates that flood damages cost the United States up to 496 billion in 2023 dollars every year (Heinrich, 2024). For the average household, increasing climate risks, such as floods, are re-pricing their most valuable asset: their house. However, understanding a property's future flood risk is not straightforward: it not only requires an understanding of the landscape and historical flood events, but it must also model future rainfall patterns and the likelihood of natural disasters such as storm surges, which may vary depending on climate change scenarios. Historically, potential homebuyers relied on Federal Emergency Management Agency (FEMA) flood maps to understand if their house is in a Special Flood Hazard Area (SFHA), defined as having a 1% or higher probability of flooding each year. However, these FEMA maps do not reflect forward-looking projections of changing weather patterns and are often outdated—75% of FEMA maps are five years or older (Eby, 2019). While insurance could fill this gap by helping homeowners understand their risks, homeowners' insurance does not typically cover flood damage. Furthermore, flood insurance outside of designated FEMA floodplains is generally expensive and difficult to obtain.

To fill in the gap in flood disclosure for homebuyers, in August 2020, Realtor.com began displaying First Street Flood Factor, a property-level flood risk estimate, for all homes listed on its site. In February 2021, Redfin.com followed suit. In September 2024, Zillow.com also started disclosing First Street's flood data. Given that these are the three largest online listing platforms, accounting for over 85% of online traffic,[1] it becomes relatively easy for a prospective home buyer to obtain information about a property's flood risk when using these online listing platforms. When a user views a home on one of the three search engines, they will be given a numerical flood risk score (between 1 for the lowest risk and 10 for the highest risk) as well as a qualitative flood risk label (minor, moderate, major, severe, or extreme) based on First Street's model. By disseminating information directly to the investor through online search platforms, First Street provides homebuyers with a novel, property-level measure of flood risk. Furthermore, unlike the information previously provided by FEMA maps, this measure incorporates forward-looking forecasts of climate change from big-data models, is updated frequently, and does not rely on insurance provision.

In a 2023 survey conducted by Zillow, over 80% of prospective homebuyers report that they consider climate risks when shopping for a home. Furthermore, among all climate risks, flood risk ranks as their

[1] https://www.rubyhome.com/blog/most-visited-real-estate-websites/

biggest concern (Zillow, 2023). The Flood Factor provided by First Street not only makes it easier for prospective homebuyers to understand a home's flood risk than ever before, it has also changed the location of perceived flood risk. For example, a report released in 2020 finds that over 40% of at-risk properties identified by First Street were not previously classified as risky by FEMA maps. If consumers notice and believe in First Street flood models, then nearly 6 million homes were suddenly reclassified as having high flood risk, despite being largely viewed as not risky before (Humphries, 2020). In equilibrium, the pricing and distributional effects of the new flood information provided by First Street remain unclear—our paper will address this gap.

In general, it is hard to disentangle the equilibrium effects of this information disclosure because many other concurrent factors could introduce bias. On one hand, increasing natural disasters and international efforts, such as the 2015 Paris Agreement, have brought more attention to climate change issues in recent years. This change in beliefs could cause concurrent price declines for flood-prone properties, as consumers become more sensitive to flood risks. However, these price declines should not be attributed solely to the information release of the Flood Factor by First Street. On the other hand, homes that have high flood risk may also have valuable amenities like ocean views. If these home features are changing in value over time, then they could also bias our estimates of prices.

Importantly for our paper, First Street's Flood Factor has a unique feature. While a home's true underlying flood risk is a continuous number, it is not displayed to homebuyers on the online search platform. Instead, as Figure 1 shows, this continuous risk (shown on x-axis) is mapped into the discrete Flood Factors (shown on y-axis) based on a set of numerical cutoffs (shown as dotted vertical lines) before being published on a search platform. By using the true underlying continuous measure of risk, we can compare homes with nearly identical flood risks but different disclosed Flood Factors on either side of each cutoff, where a regression-discontinuity (RD) design can untangle the effects of disclosure from other potential omitted factors (e.g., ocean views, housing quality, etc).[2] While a standard RD design in the post-disclosure period provides consistent results (shown in the Appendix), one may still be concerned that certain geographic features, such as having a waterfront location, are disproportionately more likely to be found on one side of the cutoff. Since these geographic features are fixed and do not change over time due to disclosure, we can address them by estimating our main results using a difference-in-RD design. By calculating the difference in prices on either side of the discontinuity, before versus after disclosure, we can provide plausibly exogenous estimates of the pricing and distributional impacts of disclosure.

In this paper, we combine First Street, CoreLogic, HMDA, and Voter Registration data to answer the following question: *What is the equilibrium price effect when property-level flood risks are disclosed*

[2]See Bell et al. (2024) for evidence as well as an econometric solution to the problem of omitted property attributes.

directly to consumers through online search platforms? There is an ongoing debate about whether homebuyers care about flood information and, thus, discount homes with flood risk. Prior studies have focused on specific geographies, utilized flood maps, leveraged flood events, or relied on flood insurance (Gibson et al., 2017; Murfin and Spiegel, 2020; Bin and Landry, 2013; Ortega and Taşpınar, 2018; Beltrán et al., 2019; Baldauf et al., 2020). Closest to our paper, Fairweather et al. (2024) studies the effect of First Street's flood factor by utilizing a randomized experiment on the Redfin platform. In order to understand how flood disclosure influences individuals' search behavior, the authors randomly assign treatment at the user level. This analysis provides causal evidence that flood disclosure changes homebuyers' search behavior and leads to changes in home prices. Our work naturally extends this analysis to examine the overall effects of information disclosures across multiple platforms on price and the attributes of homebuyers living in properties with different flood risk by utilizing a natural experiment that pseudo-randomizes the disclosed flood risk at the property level.

To summarize the average price changes before and after the rollout of the Flood Factor, we begin with a naive difference-in-differences analysis using all homes sold in the U.S. between 2016 and 2024. Using various geographic controls, ranging from county-period fixed effects (most relaxed) to block-group and FEMA-map fixed effects (most stringent), we find a significant decline in sale price of roughly 1% for homes with "extreme" flood risk following the release of First Street's Flood factor. While this provides correlative patterns over our full sample, it is nearly impossible to interpret these preliminary estimates as causal. One may worry that the Flood Factor is correlated with omitted factors, such as the proximity of the property to the sea, which is a desired amenity and may actually increase equilibrium prices. To address this, our main empirical analysis employs a difference in regression discontinuity (Diff-in-RD) model. By backing out the continuous but proprietary risk variable that is not disclosed to the public, we can compare nearly identical houses that were just below versus just above the cut-offs used to assign the discrete flood factors. An example of how this treatment is pseudo-randomly assigned to a small neighborhood of houses in Hillsborough County, Florida is shown in Figure zoomin$_m ap$.

In these RD models, we now restrict our sample to homes that are very similar in their underlying flood risk. Using the RD specification, we find a baseline price discount of 3.3% for homes with "extreme" flood risk; strikingly, we find a sizable discount of 11% for the subsample of properties exposed to a coastal source flooding (and up 16% for non-SFHA properties that are exposed to coastal flooding). This suggests that the market is particularly sensitive to flood risk when the expected damage is high and when the risk is tied to expected sea level rise. Consistent with a general drop in demand, we find that these properties also remain on the market longer–an average of 6.8 days longer and up to 13.1 days longer for the subsample of properties exposed to coastal flooding. This suggests that the market for extremely risky properties become more shallow following disclosure,

making matching between buyer and seller more difficult as a result.

Next, we plot dynamic coefficients from the difference in RD regression and show that there is no pretrend of price change declines prior to flood disclosure. In contrast, the price gap between "extreme" and "severe" properties falls after disclosure, becoming more pronounced over the three years of our post-disclosure sample. For any concurrent event that may cause price declines during the post-period to be an identification concern, it must *both* explain the drop in price right at the discontinuity point in the proprietary First Street Flood score *and* match the empirical pattern that price effects are most pronounced in *areas outside of SFHA zones exposed to coastal flooding*.

As robustness tests, we check that there is no bunching in a histogram of properties around the cutoff point, and we ensure that results are unchanged if we alternatively use a tract-level instead of block-group-level fixed effects. Finally, in a covariate balance test, we reassuringly find that most hedonic housing characteristics, such as the number of bedrooms and bathrooms, are similar on either side of the cutoff. There could still be an imbalance around the cutoff in other unobserved attributes, such as geographic location (eg, ocean front properties are disproportionately more likely to be on the right of the cutoff). However, these imbalances are fixed for each property and should not change between the pre- and post-disclosure period.

Lastly, in order to benchmark the economic magnitude of the observed price declines, we compare the dollar value of these declines to the actuarial cost estimates of future flooding provided by First Street. Since the average home sale price in our data is $382,850, the observed 3.3% drop in prices translates into a net value decline of $12,634.05. Based on First Street's cost model, the actuarial decline in value due to future expected flooding, based on a 5% discount rate and staying in the home for 10 years, should actually be about $34,752. Thus, consistent with a story of retail inattention, the observed price drop does not seem to fully account for future flood damages; instead, it accounts for just 36% of the actuarial value.

The second portion of our paper focuses on distributional impacts, in order to answer the question: *How does flood risk disclosure change who lives in flood-prone properties?* So far, we find that homes with an extremely high Flood Factor (implying extreme flood risk) dropped in value and stayed on the market longer after First Street released its data. Since housing supply is generally inelastic, this is consistent with lower demand for flood-prone properties. However, the change in demand may also be heterogeneous and vary depending on consumer characteristics. For example, flood-prone properties may attract more climate-risk-tolerant buyers, and this risk tolerance may be correlated with observable characteristics such as age, race, and political beliefs. On the other hand, it may simply attract more first-time or lower-income buyers, because they are more financially constrained and cannot afford to buy the less flood-prone properties, which are becoming relatively more expensive. To

better understand the distributive impact of disclosure, we regress various homebuyers' demographics on flood-risk disclosure.

Using the same difference in RD regression design, we find a strong and significant decrease in average income (roughly 5.3%) for homebuyers of extremely risky properties following disclosure. Similarly, we find that homebuyers of extremely risky properties are 1.8% more likely to hold a FHA loan. Since FHA loans are intended to help first-time buyers with lower credit quality, this similarly suggests that, in equilibrium, buyers are becoming more financially constrained. On the other hand, even when we control for income effects, we also find resorting along non-pecuniary lines. Following disclosure, buyers of the riskiest properties are, on average, 1.4 years older and are 0.4% less likely to be registered as Democrat.[3]

Ultimately, both heterogeneous characteristics that are correlated with flood-risk tolerance (eg age) as well as the ability to pay appear to factor into the redistribution that we observe after disclosure. Since it is unclear how these preferences interact with each other, we estimate a simple choice model that accounts for both dimensions of resorting observed in our data. Calibrating the model reveals that resorting on income is almost entirely driven by prices, while resorting on age is mostly due to preference heterogeneity.

Our work naturally complements the growing literature on natural disasters and frictions in insurance pricing. Traditionally, homeowners embed the cost of natural disasters, including floods, through the price of their homeowners' insurance (Ge et al. 2022, Mulder 2024). However, as the number of natural disasters continues to increase, the property and casualty insurance market faces increasing frictions, including sharply increasing and uneven premiums due to reinsurance (Keys and Mulder 2024); riskier mortgage originations that amplify systemic and taxpayer risks (Sastry et al. 2023); insufficient coverage in the event of total loss (Cookson et al. 2025); as well as misclassified risk leading to both suboptimal insurance take-up and adaptations (Mulder 2024). In addition, widespread under-insurance (Ge et al. 2025) and increasing rates of delinquency in insurance (Sastry et al. 2024) have distributional impacts by shifting risk to the most financially constrained households.

We also contribute to the literature on the impact of digitization in the housing market. Despite the real estate sector's slow adoption of technology in general, online housing platforms continue to play a crucial role for households' search behavior, altering the information environment for market participants. For example, Buchak, Matvos, Piskorski and Seru (2022) finds that algorithmic pricing deployed by iBuyers[4] provides automated intermediation by producing more accurate predictions in

[3] Although estimates are not significant, we also find that buyers are more likely to identify as non-white (as shown in the Appendix).

[4] iBuyers stands for "instant buyer," referring to real estate companies "that use algorithms and technology to buy and resell homes quickly." https://www.nerdwallet.com/article/mortgages/understanding-ibuyers, accessed August 12, 2024.

liquid housing markets, and Harrison et al. (2024) finds both elevated equilibrium prices as well as localized pricing spillovers after the entrance of iBuyers. In addition, Fu, Jin and Liu (2022) shows that Zillow's automated valuations ("Zestimates") influence both listing and sale prices, while Raymond (2024) finds that algorithmic valuation can have market-level effects, closing historical racial gaps in housing values. On the household side, Bailey et al. (2018) finds that networks on social media platforms significantly impact residential home-buying behavior. Finally, Fuster et al. (2022) finds that machine learning algorithms have distributional impacts on the mortgage market by being the least helpful for Black and Hispanic borrowers.

Finally, since a house's flood risk is both intangible and often difficult to understand without disclosure, our studies also tie into a literature on the importance of information salience in finance. Studies show that effective disclosure needs to be simple and salient in order to influence household financial choices (Agarwal et al. 2015), especially regarding complicated issues such as global warming (Choi et al. 2020). When it comes to attention to climate-related issues, there is a divergence as institutions appear to be highly aware of their risks (Krueger et al. 2020). In contrast, regulators, academics, and practitioners generally agree that market prices underreact to climate risks—suggesting that most investors are not properly attentive (Stroebel and Wurgler 2021). Consistent with the theory of market underreaction, we find that the average price drop that we observe represents only about 41% of expected flood costs, suggesting that not all consumers are equally attentive to First Street's Flood Factor.

Results for both prices and homebuyer attributes provide evidence that markets are capitalizing on information about flood risk, and older, more Republican, lower-income households will bear a higher risk of flooding. Our context is important, as information on flood risk (and other natural hazards) has become quite accessible to all homebuyers and, thus, should lessen the costs of unknown risks to housing investments. Ultimately, inequality in who bears the near-term cost of providing flood disclosures is important, as those living in high risk homes may require additional financial assistance given the escalating costs of insurance and the role of the public sector as the lender of last resort.

2 Data and Summary Correlations

2.1 Summary of Main Data Sources

First Street

The measure of flood risk that we focus on is the Flood Factor, a score published by the First Street Foundation and released on popular real-estate platforms such as Zillow and Redfin (see Figure 3 for an example of how the scores are displayed for a platform user on Redfin). The initial goal of the First Street Foundation was to provide property-level information on future risks of flooding for all homes

in the US. Subsequently, they also expanded their database to include information on other climate risks, including wildfires, heat, wind, and air. Their flood factor is produced using the First Street Flood Model (FSF-FM) — a nationwide model that calculates the potential risk of flooding based on a combination of rainfall, streamflow, rise in sea level, tidal expectations, and storm surge.

The FSF-FM model is forward-looking in nature and incorporates both historic flood risk as well as climate change-driven changes in flood risk. One notable component of their model that deviates from standard flood models, including the FEMA floodplain maps, is the use of the First Street Precipitation Model (FS-PM). The FS-PM takes into account early 21st century rainfall records and calculates forward-looking projections, capturing the notion that the United States will experience more occurrences of extreme rainfall in the future. As a result, inland properties that were traditionally classified as safe based on floodplain maps are more likely to be reclassified as risky under FSF-FM.[5]

We obtained a dataset from the First Street Foundation that provides a composite measure of future flood risk for each property, resulting in a score between 1 (lowest risk) and 10 (highest risk). As shown in Table 1, our final merged dataset consists of over 25.6 million sales transactions. 5% of all properties are at risk of coastal flooding, and 13% are at risk of pluvial (i.e., rainfall) flooding. In general, the causes of flooding do not have to be mutually exclusive. In the rest of this paper, we define Coastal properties as all properties that are exposed to coastal flooding as a source of risk. To help visualize their general locations, Figures A6a and b plot the coastal properties as red, semi-transparent dots, along with local bodies of water for Hillsborough County, Florida (panel a) and a zoomed-in subset of Hillsborough County (panel b). We see that coastal properties are those that lie in close proximity to a local body of water, whether that is a lake, river, or ocean.

Table 2 breaks down the transaction data by the property's Flood Factor (column 1) and qualitative risk category (column 2). We see that while over 84% of all sold properties have a minimal risk score of 1, there is also a substantial number of properties with either severe or extreme flood risk. For example, we observe 591,844 (424,407+167,437) homes sold that had "extreme" flood risk.

In addition to the qualitative category and quantitative risk score, we also back out the internally calculated running variable— the expected flood depth under different time horizons and the expected costs of repairs at different time horizons. Using a weighted average of the expected flood depths under various horizons (aka "return periods" under the FSF-FM), we are able to construct a continuous score called the expected flood depth. Figure 1 shows how we can cleanly infer the cutoff points by plotting the observed flood factor (y-axis) against the logarithm of the expected depth (x-axis). As Figure

[5]The First Street model improves upon existing flood map methodology by being more recently updated and forward-looking. However, a team of White House scientific advisers has warned against climate models such as First Street's, saying they are of "questionable quality," and Bloomberg has also noted that catastrophe models have large standard errors, often coming to very different conclusions (Meyer, 2024). Our paper does not take a stand on the validity of the First Street measure but instead focuses on the impact of the information that is given to consumers.

4 depicts, this allows neighboring homes that are in similar neighborhoods with nearly identical risk of flooding to be assigned very different scores. We will use these cutoff points for our regression discontinuity analysis in Section 3.1.

First Street vs FEMA

We obtain floodplain data from FEMA, and we merge this with our First Street data to see how the two measures of flood risk compare. To illustrate the differences, Figure A4 shows the fraction of home transactions per census tract that are labeled as risky by FEMA in panel a (i.e., those in a SFHA, or equivalently, a 100-year floodplain; [6]) versus First Street in panel b (i.e., those "with "extreme" flood risk) for the state of Florida. In general, we see that the two data sources seem closely related, but they are not identical. For instance, more properties sold in the upper left corner are labeled as risky by First Street compared to FEMA.

In addition to different categorizations of risk, FEMA's flood data is often less granular than First Street's flood data. To illustrate, we zoom into a select number of census tracts in Hillsborough County, Florida, in Figure A5, and we plot every transaction as an individual dot colored by either its FEMA risk (i.e., whether it is in a SFHA) or its First Street risk (a numerical factor ranging from 2 to 10, from least to most risky). First, by construction, we see much more variation in degrees of risk under the First Street data. Second, for example, if you look in the lower half of tract 006200 (middle right side of the graph), it is totally blue in panel a, or designated as not SFHA by FEMA. However, in panel b, we see a large variation in First Street's flood risks, ranging anywhere from 3 (moderate) to 8 (severe). If you look closely, you can see that houses right next to each other can have very different scores under both FEMA and First Street, but it is especially prominent in First Street's model due to the range of scores possible.

We formalize the comparison for all transactions in the United States in column 5 of Table 2. More specifically, we calculate the proportion of homes sold under each Flood Factor that is also classified within a FEMA flood zone. As Table 2 shows, the Flood Factor and FEMA zones are positively correlated, such that larger Flood Factors have a greater fraction of homes in FEMA flood zones; however, there is still considerable disagreement. For example, even for homes in the riskiest category of 10, roughly 38% of them would **not** have been classified as risky by FEMA.

CoreLogic

Our transaction data comes from the Corelogic Multiple Listing Enterprise Solutions (MLS) Database. The MLS database comprises data that real estate agents use to match buyers with sellers. In addition to the address and geolocation, we also have hedonic characteristics such as the number of bedrooms,

[6] FEMA's 100-year floodplain is an area that has a 1% or higher probability of flooding each year.

bathrooms, and lot size. In addition, we have information on the sale price and date, as well as limited agent information. Table 1 summarizes several key variables from our transactions data.

After merging the Corelogic database with First Street data, we filter the data by removing properties with sales prices below $50,000 or with more than one building. In our main estimation sample, we further restrict the sample to two time periods of interest: between January 2016 and August 2020 (aka the "pre-period") and between August 2021 and December 2024 (aka the "post-period"). As Figure 2 shows, these dates align with the period before First Street data was publicly disseminated and the period after the final version of First Street data was disseminated, respectively.

After applying these filters, we are left with roughly 22.2 million transactions. The average sale price (see Table 1) was $372,170; although there is substantial variation. The 25th and 75th percentiles of sale prices are $172,000 and $439,000, respectively. The median house sold in our sample has 3 bedrooms and 2 bathrooms, with a living area of 1,680 square feet.

HMDA

For homes that are purchased with a mortgage, we obtain additional borrower characteristics through the HMDA (Home Mortgage Disclosure Act) data. We use transaction level information on census tracts, year, loan amount, and sales price to match individual mortgages for home purchases reported as part of HMDA to individual property transactions.[7] For each mortgage that we successfully match, we gather the borrower's age, ethnicity, income, and a dummy variable equal to one if the loan is an FHA loan. Since the FHA insures the loan, it allows first-time and low-credit borrowers to attain better deals on their mortgages. Thus, in general, we use the FHA loan dummy as a proxy for first-time or low-credit homebuyers.

Table 1 summarizes the HMDA data in the second to last panel. We were able to match roughly 8.4 million mortgage applications. For this matched sample, the average income was about $128,936 and 17% were FHA loans. In addition, about 17% of buyers were aged 55 and above, and 39% identified themselves as non-white.

L2 Voter Registration Data

We obtain buyer's age, college attainment, as well as political party affiliation using L2 Voter Registration data. For each sales transaction in Corelogic, we match its address to the address of a household recorded in the L2 data in the next general election following the sale. Additionally, we check to make sure that at least one of the names registered to that household matches the buyer's name in Corelogic.

[7] This matching process has been implemented by other researchers (Bayer et al., 2016, 2025; Bernstein et al., 2022) and is fuzzy in nature. In some cases, our matching process leads to non-unique matches (¡5%), which we exclude. In general, we are able to match over 60% of mortgages to unique transactions.

Sometimes, registrations take a while to update, so we also use the following year of L2 data, if we cannot find an adequate match.

Table 1 summarizes the voter registration data in the last panel. We were able to match roughly 11 million sales to a household in the L2 data. Among these households, the average age is 44 and about 43% of them have at least one member with a college degree. In addition, the data looks very politically balanced; roughly 39% of households have at least one member registered as a Democrat, and the same fraction (39%) of households have at least one member registered as a Republican. For all analyses in the paper, we use the dummy indicators rather than a percentage of household calculations because we do not wish to be biased by dependents who may be registered but ultimately do not make home purchase decisions. However, the results are very similar if we use a percentage instead. While we omit those from the paper for space, we can share those results upon request.

2.2 Summary of Price Changes in Risky Homes Following Disclosure

As a baseline, we provide a naive estimate of log price changes across homes of different flood-risk levels, both before and after the release of First Street's Flood Factor on online platforms. While this only shows non-causal, correlative patterns, one advantage of this analysis is that it provides a broad overview of the market across all homes sold in the U.S.–rather than focusing on a narrow band of homes around a threshold, which is necessary when we move to a regression discontinuity design to provide causal estimates.

More specifically, we present the averages estimated from the following DID specification:

$$Ln(Price)_{ijt} = \alpha_t + \mu_{jt} + \vec{\beta_1}\vec{F_i} + \vec{\beta_2}\vec{F_i} \cdot Post_t + \vec{\beta_3}\vec{X_{ijt}} + \varepsilon_{ijt} \tag{2.1}$$

where each observation is one sales transaction and i, j, and t represent property, location, and time indices, respectively. α_t and μ_{jt} represent time (year-month) and time-location fixed effects, respectively; across specifications, the unit of location varies from county (most broad) to block-group (most granular). $\vec{X_{ijt}}$ represents a vector of hedonic controls, including the building age, number of bedrooms, number of bathrooms, lot size, and living area.

Our key explanatory variables are represented by the vector $\vec{F_i}$, which is a 5x1 vector of dummy variables, each equal to one if the property's flood factor fell into one of the five respective risk categories or above (ie minimal, minor, moderate, major, or severe). The coefficients summarize the average *incremental* price change between homes of that category and homes in the category directly

above.

$$\vec{F_i} = \begin{pmatrix} I(> Minimal)_i \\ I(> Minor)_i \\ I(> Moderate)_i \\ I(> Major)_i \\ I(> Severe)_i \end{pmatrix}$$

To summarize the price changes resulting from the release of First Street's Flood Factor, we multiply the treatment $\vec{F_i}$ by $Post_t$, a scalar dummy equal to one if the transaction took place in the post period. As Figure 2 shows, we define the post period as the period following the official release of the final version of First Street's Flood Factor. Thus, $\vec{\beta_2}$ captures the change in prices from before to after flood disclosure.[8] If consumers cared about flood risk, this should lower prices for high-risk homes. Subsequently, we would expect the price changes in $\vec{\beta_2}$ associated with high-risk homes, such as homes with extreme flood risk, to be negative and significant.

Table 3 summarizes price changes across the entire United States. Going left to right, from column (1) to column (6), we add progressively more stringent controls. For example, in column (1), we only include time fixed effects. Going from columns (2) to (4), we add increasingly more granular geographic fixed effects (county, tract, and blockgroup, respectively). Starting in column (5), we add building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft) as control variables. Lastly, in column (3), we add a dummy equal to one if a property is inside a SFHA. In general, we find that properties listed as "extreme" (which is the same as being above "severe") have an incremental price decline of about 1%. There are no other significant and consistent price declines at other lower risk cutoffs, suggesting that home buyers only pay attention to flood risk when it is especially high. This "extreme" effect is statistically significant at the 1% level for all specifications with location controls and remains unchanged with the SFHA indicator. While we cannot interpret these correlations as causal, they illustrate that the significant price impact of flood disclosure across the United States.[9]

3 Price and Turnover Effects

To understand the impact of nationwide flood disclosure on the residential housing market, we begin by estimating its effect on equilibrium prices as well as turnover rates. To do this in a plausibly causal fashion, in Section 3.1 we estimate a difference-in-regression-discontinuity (Diff-in-RD) regression for different risk factor cutoffs. Overall, consistent with the correlative evidence shown in Table 3, we

[8] We can also calculate the average (incremental) price in the post-disclosure period by adding the difference to the pre-disclosure period's price.

[9] As robustness, we check and verify there is no pre-trend that can explain these findings. This is shown in Figure A1, which plots the dynamic coefficients from the DID regresson.

find that homes with extreme flood risk face a significant decline in prices. This effect is primarily concentrated in Coastal areas (outside of SFHA zones). In Section 3.4, we test whether the drop in price appears to be actuarially fair.

3.1 RD Estimator

While the summaries provided in Table 3 based on DID regressions utilize our full sample and possess the most statistical power, they suffer from concerns regarding changing beliefs about flood risk, as well as unobserved time-varying changes in the value of amenities (e.g., coastal views, water access) that may be correlated with flood risk. For instance, between our pre- and post-period, several hurricanes drew national attention to flooding and increased the salience of flood risk in homes more broadly. As a result, homes with flood risk, which are naturally correlated with First Street's Flood Factor, may drop in price simply due to the increasing salience of climate risk. While the pretrend analysis and controls for existing FEMA information ameliorate some of these concerns, they may not perfectly capture how homebuyers' perceptions of climate risk are changing over time. If so, then our DID estimates may still have little to do with First Street's flood disclosure.

To address the identification concerns above, for our main empirical analyses, we use a difference in regression-discontinuity (Diff-in-RD) design.[10] First, we recreate a continuous measure of a property's flood risk called "expected depth" (see Figure 1) that is monotonically mapped into the discrete First Street Flood Factor using a set of empirically observable cut-offs. Then, using a Diff-in-RD analysis, we compare how prices changed for homes just above and just below the empirical cutoff values. In doing so, we assume that homes right around the cutoff are otherwise very similar and only differ depending on their assigned Flood Factor. While a standard RD-specification also works well in this setting (results are shown in Appendix Section B.1), we utilize a difference in RD approach because some unobserved house characteristics may still be discontinuous at the cutoff. To interpret our estimates as plausibly exogenous, we assume that the potential buyers of these homes suddenly have different beliefs about climate risk at the cutoff. Any observed price changes at the cut-off can then be reasonably assumed to be driven by the Flood Score disclosure. If we expect consumers to discount homes because they learn a property has a higher Flood Score, then we would expect to see a price drop right at the cutoff.

Formally, for each change in flood factor scores, we estimate the following difference-in-RD regression:

[10] Results using a standard RD regression are shown in A1 and depicted in A2. These results are even stronger and more robust, but they may be prone to omitted variable vias. In particular, some unobservable traits of properties (eg having an oceanfront view) may bunch discontinuously at the cutoff. To control for these traits, which should remain static for a given property before and after disclosure, we utilize the more stringent difference-in-RD design throughout the paper.

$$
\begin{aligned}
\text{Ln(Price)}_{ijt} &= \beta_1 \text{I(high-risk)} \cdot \text{I(Post)} \\
&+ \beta_2 \text{I(Post)} + \beta_3 \text{I(high-risk)} + \sum_{m=1}^{3} \rho_m \text{exp depth}_i^m \\
&+ \sum_{j=1}^{3} \gamma_j \text{exp depth}_i^j \cdot \text{I(high-risk)} + \sum_{k=1}^{3} \nu_k \text{exp depth}_i^k \cdot \text{I(Post)} \\
&+ \sum_{n=1}^{3} \delta_n \text{exp depth}_i^n \cdot \text{I(high-risk)}_i \cdot \text{I(Post)} + \vec{\beta_4} \vec{X_{ijt}} + \mu_j + \varepsilon_{ijt}
\end{aligned} \tag{3.1}
$$

where each observation is a sales transaction and i, j, and t represent property, location, and time indices, respectively. μ_j represents block-group fixed effects and $\vec{X_{ijt}}$ represents a vector of hedonic controls, including the building age, number of bedrooms, number of bathrooms, lot size, and living area. exp depth$_i$ is our continuous measure of flooding risk, and I(high-risk) captures the discontinuous jump. It is set equal to one if a property fell into the riskier category (ie has a flood score of 9) and is zero otherwise (ie has a flood score of 8). If we expect consumers to discount properties with higher Flood Factors, then we would expect β_1 to be significantly less than 0.

Table 4 shows the results from estimating Eq 3.1 with each panel displaying the disclosure effect estimated over the cutoff points from different pairs of flood risk factors: Panel A shows 3 vs 4, Panel B shows 4 vs 5, Panel C shows 5 vs 6, etc. In particular, Panels B, D, F may be particularly interesting because they focus on thresholds that also correspond to a change in qualitative risk (eg Panel B compares factors 4 vs 5, or moderate vs major risk homes), which is all that the user can see when they scroll through a home's listing on the online platform. To see the differences between the other risk factors, the user would have to click on the box containing the risk factor to get more information. Based on our fndings using the DID regression, we hypothesize that users pay more attention to the risk factors when the homes are especially risky. [11]

The columns of Table 4 correspond to estimations using a linear auto-selected band, a polynomial of order 2, and a polynomial of order 3, respectively.[12] Consistent with Section 2.2, we only find a significant price drop for the riskiest homes. Panel F shows that we find a significant price drop of 3.3% (linear specification shown in column 1), 2.4% (second order polynomial shown in column 2), or 3.0% (third order polynomial shown in column 3) post disclosure for homes in the severe versus above severe categories. Similarly, Panels E and G shows that there is a significant and similar-magnitude price drop at other large flood risk values: 2.7% and 3.4% for the 7v8 and 9v10 cutoffs respectively

[11] We omit the results for the minimal and minor categories, because they pose very little flood risk for the homeowner. In addition, the minor category holds very few homes so the estimation around this cutoff is especially noisy.

[12] The sample size varies between each estimate based on the auto-selected bandwidths for the points to the left and the right of the risk cutoff (ie the control and treatment groups, respectively). These two bands are optimized by Stata within the *rdrobust* package to minimize the Mean Squared Error (MSE) of the RD treatment effect estimators.

(using a linear bandwidth, column 1). No other risk pair is consistently significant across all three specifications; risk factors 3 vs 4 also find significant drops, but only with the second and third order polynomial specifications.

Figure 5 Panel A illustrates the Diff-in-RD for homes with scores of 8 (labeled as "severe") and 9 (labeled as "extreme"), for which we find strong, treatment effects.[13] Specifically, we plot a binscatter graph with expected depth on the x-axis and log price on the y-axis. Properties with a flood score of 8 are shown in blue, while properties with a flood score of 9 are shown in red. We control for location, time, and basic hedonics by residualizing log sale prices to block-group fixed effects and hedonic controls. Furthermore, each dot represents the price difference between the post-disclosure period and the pre-disclosure period. [14]

For conciseness, we will focus on 8v9 for the remainder of the main paper. Not only does this sample have the highest number of total housing transactions, but more importantly, it is one of the most interesting cut-off points to examine due to its salience. This cutoff point captures the most salient change in label from "severe" to "extreme" rating–the only First Street information that is shown to consumers on the home search splash page.

3.2 Heterogeneous Effects

In Table 5, we test whether the baseline effect for 8 vs 9 homes varies by geographic location. In particular, we will look at two different subsample cuts: homes exposed to a coastal source of flooding versus not, and SFHA versus non-SFHA properties. First, we test whether certain types of flood risk are priced higher than others. In the First Street data, the main flood types in our time sample are flooding from rain (i.e., pluvial) and flooding from exposure to the coast (i.e., fluvial). We hypothesize that since coastal flooding is both potentially more costly and more salient, it would have stronger price effects. Additionally, looking at nonSFHA homes not only allows the First Street data to provide novel flood information, but it also ensures that observed price effects are not confounded by the expected payoffs from subsidized flood insurance.

For convenience, we repeat the main specification using the full sample in column 1. Next, columns 2 and 3 repeat the analysis for just the properties exposed to coastal flooding (column 2) versus not (column 3). Consistent with our hypothesis, we find that properties that subject to coastal flooding risk experience a large price drop of roughly 11%. This suggests that the release of First Street data has an especially strong price effect in areas where flood risk was already salient, rather than adding new pricing information to inland areas that were traditionally considered less risky. This finding

[13] Appendix Figure A3 show the RD across alternative and significant cutoffs of 7v8 and 9v10.

[14] Ideally, we would not require any controls and would rely solely on the RD design. While we also found effects without any controls, the estimates were quite noisy, and significant jumps right at the cutoff were harder to identify. Intuitively, even properties with similar flood risks can vary significantly in price depending on their location and characteristics, such as size. By residualizing on the full sample first, we can reduce much of the noise and make the RD less fuzzy.

largely aligns with several behavioral papers, which have found that flooding is only priced in areas where individuals believe in the flood risk (Baldauf et al., 2020) or that individuals largely respond to disaster risk only when it is salient (Dessaint and Matray, 2017; Gallagher, 2014). Next, we look at SFHA homes (column 4) and nonSFHA homes (column 5) separately, finding equally strong but more significant results *outside* of SFHAs. This confirms that our results are not driven by areas with pre-existing information, where flood insurance is prevalent. Lastly, when we break down these subsamples further in Appendix Table A4, we see that the strongest price effects are in nonSFHA homes exposed to coastal flooding, with a striking price discount of 16%. Showcasing that the price differential is highest in areas where flood insurance is not commonly purchased, this price effect is more than double the 6.6% decline observed in SFHA homes exposed to coastal flooding.

3.3 Robustness Checks

3.3.1 Covariate Balance & Manipulation of the Running Variable

In our RD analyzes, the drop in log price is sharp and observed very close to the cutoff point, so it is unlikely to be driven by other factors, such as an increasing salience of climate risk over our time sample. As a robustness check, however, Table 7 tests whether other key features of houses sold are unbalanced across the cutoff point. To run this check, we plot expected depth (x-axis) against each of our hedonic variables (y-axis) as well as a dummy for coastal source of flooding (y-axis), which is a key source of variation for our price estimates. To match our difference in RD specification, we take the difference between the average bin value in the post-period and the average value in the pre-period. Reassuringly, around the cutoff, we do not observe any other sudden changes in size (i.e., lot size or living square footage), building age, or number of bedrooms. Thus, even though we include hedonic controls, the price difference estimated around the cutoffs is unlikely to be driven by a confounding feature of the house itself when we use the difference in RD design.

The other standard concern in an RD design is bunching or manipulating one's position on either side of the discontinuity. In our case, geographical features determine the running variable, and thus a property owner is unable to change their corresponding score. Confirming this, Figure 7 shows a seamless histogram of houses across the 8 versus 9 cutoff, which shows no evidence of bunching. Ultimately, we recognize the concern that properties with omitted flood-prone attributes (e.g., elevation, distance to coast, etc) may still be over-represented on one side of the cutoff based on how the scores are set. This is one of the main justifications for the difference-in-RD design, as it allows us to control for these pre-existing differences in our estimates.

3.3.2 No Evidence of Pre-trend

One may be worried that the observed price drop shown in Section 3.1 is not due to disclosure but rather to time trends in climate awareness that coincide with the First Street disclosure. To address this concern, we conduct a robustness test by dynamically plotting the price difference for extreme flood risk properties from Eq 3.1 year-by-year in Figure 6. Looking at Figure 6, we see that the coefficient on "extreme" (i.e., treated) trended down a little in 2016-2018 and rose again in 2019, but in general, it was stable before the release of First Street data.

3.3.3 Different Levels of Geographic Controls

Next, in Table 6, as a robustness test, we check whether our choice of fixed effects drives our findings. Ideally, we can rely on the RD design to produce a causal estimate without any fixed effects; however, since our sample covers all of the United States, there are many heterogeneous factors that create considerable variance in prices (e.g., location, size of house, age of house). Controlling for these factors can help us reduce noise and more precisely estimate the effect. Column 1 shows the results using no geographic fixed effects at all; column 2 adds coarse county-level fixed effects, column 3 adds census tract-level fixed effects, and column 4 adds the most granular block-group fixed effects. Across all columns with geographic controls, we see that the drop in price is very similar in magnitude, ranging from -2.6% with county-fixed effects to -3.3% with block group fixed effects. While the specification with no geographic controls has a lower average point estimate of 1.2%, when considering the very large standard error of 1.96%, we cannot reject the hypothesis that they are statistically the same as other columns. In addition, consistent with the explanation that more granular fixed effects reduce noise, the t-statistics monotonically increase from column 1 (t stat of 0.52) to column 4 (t stat of 2.32). Thus, while our findings do not hinge solely on using Block-group fixed effects, they allow us to obtain the sharpest estimates of the price effect. As a result, we focus on this specification for the remainder of the paper.

3.4 Is the Estimated RD Price Drop Actuarily Fair?

Given an average home sale price of $382,850 in our data (see Table 1), the observed 3.2% price drop for the full sample represents a $12,634.05 price decline. To put this magnitude into context, we utilize First Street's data on expected annual flood costs, which are estimated for each property in our dataset based on the type of expected flooding and historical cost estimates from the United States Army Corps of Engineers (USACE). After factoring in a building's characteristics (e.g., number of stories, presence of a basement, etc), the value of the building, the value of the land, and the expected depth of flooding, the USACE data provides an estimate of the amount of damage expected based on historical records from past realized insurance claims.

Using First Street's proprietary cost estimates, we find that the average annualized flood-related costs of owning a home are $2,934.00 and $7,434.50 for risk ratings of 8 and 9 in our data, respectively. A back-of-envelope calculation assuming a 5% discount rate and staying in the house for 10 years implies that these costs have net present values (NPVs) of $22,655.57 and $57,407.24, respectively. Taking the difference between these two NPVs suggests that a reasonable estimate of the difference in value (i.e., increased cost) between owning a home with a risk rating of 8 versus 9 is $34,752.

Overall, relative to the estimated cost difference of $34,752, the observed price decline of $12,634.05 is significantly smaller—just about 36% of the actuarial value. Thus, while equilibrium prices reflect a change in perceived risk, they do not appear to fully account for a reasonable estimate of future costs. It is important to note that while First Street's cost estimates are not provided directly or freely to the consumer, a homebuyer can click on the risk profile of any home on Redfin, Realtor.com, or Zillow and request a free 7-day trial that provides access to this detailed cost data for an unlimited number of homes. The fact that the price decline does not reflect these expected costs suggests that homebuyers seldom click through to this detailed information, that they have trouble discounting future values, or that they simply do not internalize the expected costs. Although the presence of flood insurance could alleviate costs, this is unlikely to explain the disparity. because Table 5 shows that the price drop is especially strong and similar in magnitude for homes outside SFHAs. For these homes, there is no subsidized public flood insurance market, and private flood insurance is both rare and often overpriced. Yet, we continue to find that equilibrium price declines do not appear to be close to actuarial costs.

3.5 Turnover Effects

To get a complete picture of the effect of flood disclosure on the housing market, we conclude by also looking at the total days on market of each home. To do this, we run the same difference in RD specification shown in Eq 3.1 but replace the left-hand-side variable with total days on market. The estimated results are shown in Table 8 Panel A and visually depicted in Figure 8 Panel A. Consistent with a story of reduced demand for the riskiest homes, we find that when homes that are disclosed to have "extreme" flood risk have significantly longer turnover times—roughly a week in the unconditional sample. Similar to the price effect, this delay in sale time is significantly higher for homes exposed to coastal source of flooding (Figure 8 Panel B), which take 13 days on average to sell. Together with our pricing results, this suggests that the First Street disclosure significantly moved buyers away from the riskiest homes, on net, which subsequently created a shallower market and made matching harder overall.

4 Distributional Effects

In this Section, we start to unpack the equilibrium effect of disclosure on *who* holds the riskiest properties following disclosure. First, in Section 4.1, we examine the effect of disclosure on two proxies of equilibrium homebuyers' financial constraints: (1) their average income and (2) whether they took out an FHA loan, which is geared towards first-time and credit-constrained borrowers. Next, in Section 4.2, we examine how disclosure affects two other household characteristics, both on their own as well as after controlling for income: average age and a dummy for a registered Republican. Overall, we find that flood disclosure shifts the distribution of homebuyers—after disclosure, the buyers of the riskiest homes are more financially constrained, older, and more likely to be Republicans.

4.1 Financial Constraints

In this section, we estimate the model given by Equation 3.1, but we replace the dependent variable (previously log price) with two proxies for homebuyers' financial constraints from HMDA: income and a dummy for FHA loans.[15] Since our demographic information comes from the HMDA dataset, we note that we must further restrict our sample to only home purchases that required a mortgage.

First, in Panel A of Table 8, we find that after disclosure, the most risky ("above severe") homes are being purchased by individuals with roughly 2.5% (column 2) to 5.3% (column 1) lower income.[16] This is notable because lower income individuals are less likely to be able to bear the costs of a large climate event, and they are also less likely to buy insurance to help cover these tail events. This is further exacerbated by the fact that the price discount for holding a risky home does not seem to be actuarially fair, based on First Street's cost models. In equilibrium, this suggests that disclosing flood risks directly to consumers may have climate justice implications, because the heavy costs of climate change may disproportionately fall on the shoulders of individuals who can least bear them. This discontinuity is also visually shown in Panel a of Figure 9.

In the Appendix, Table A4 shows the result broken down by key subsamples. Column 2 versus 3 verifies that the shift in income of buyers is largely driven by areas exposed to coastal flooding, with an effect up to 13%. While similar in magnitude, within non-SFHA properties (column 5), we find a stronger and more significant coefficient compared to SFHA properties (column 4). Thus, while we lose sample size and thus power in the subsample analysis, the evidence is consistent with the hypothesis that the redistribution effect is strongest in nonSFHA and homes with salient (ie coastal) sources of flooding.

Finally, in Panel B of Table 8, we also examine FHA as an outcome and find that after disclosure,

[15] We also consider race from the HMDA data; however, since it is insignificant, we present the results in the appendix instead.

[16] We also test the effects of disclosure on race, but while there is some evidence that average age declined, we fail to find consistent effects. The results are instead presented in Appendix Table A8.

the most risky ("above severe") homes are purchased by individuals who are 1.5% (column 2) to 2.7% (column 3) more likely to have a FHA loan. This regression is visually depicted in Figure 9 Panel b. Since FHA loans are typically given to first-time and low-credit borrowers, this is further evidence that the expected costs of flooding and climate change may increasingly shift to the more financially constrained. In the Appendix, Table A5 further tests the findings across different subsamples. Here, we find especially pronounced effects in nonSFHA zones (5) and similar effects across areas exposed to coastal flooding (column 2) and otherwise (column 3).

Overall, our financial RD regressions using HMDA outcomes show a consistent story—following flood disclosure, the most risky homes are increasingly held by lower-income and low-credit borrowers. This effect exists even outside of SFHAs, and it is often strongest for homes exposed to coastal flooding.

4.2 Heterogenrous Preferences Unrelated to Financial Constraints

To test whether homebuyers may also sort on other demographic factors unrelated to financial constraints, we re-estimate Equation 3.1 but replace the left-hand side variable with buyer demographics. The buyer demographics that we consider are the average age of the household and a dummy variable for whether anyone in the household registered as a Democrat.[17]

In some specifications, we additionally control for the correlation between these outcomes and income (e.g., older individuals tend to work in higher-level positions and, thus, on average, have higher incomes) by adding both a control variable for income as well as interacting this income variable with the post-disclosure dummy (e.g., over time, higher-level positions may disproportionately pay more, and thus, on average, older individuals have a larger income gap over time).

Table 9 Panels A and B show our results for average age and the dummy variable for Democrat, respectively. Column 1 shows the estimation without income controls, while column 2 repeats the analysis using income controls. Columns 3 and 4 show the results with income controls but using alternative specifications; instead of a linear fit within an auto-estimated band, they use a polynomial of order 2 or 3, respectively. From Panel A of Table 9, we see that the average homebuyer for above-severely risky homes becomes older following disclosure. When we add income controls (columns 2-3), the effect stays strong and is generally significant. We find that the average homebuyer becomes 0.88 to 2.01 years older, depending on the exact specification. This suggests that older generations of homebuyers are less concerned about climate risks and, as a result, potential future flood risks in their homes.

Next, Panel B of Table 9 shows that the average homebuyer for above-severely risky homes is not significantly more or less likely to be Democrat following disclosure. Across all columns, we find that the effect is very small and statistically insignificant. On one hand, we might expect a negative and

[17] Results are reversed but similar if we regress on a dummy variable for Republican.

meaningful coefficient, since the Democratic party is generally concerned about climate change and meeting net-zero carbon emission targets. However, in our sample, registered Democrats also have less income and are older (compared to the average citizen). Thus, on net, the effect may be insignificant. To help tease out the competing effects on sorting, we will utilize a structural model in the next section.

Figure 9 visually depicts both regressions using a linear auto-band. We can see a small jump up at the threshold for average age and an insignificant jump down for the Democrat dummy.

Overall, the results in this section show that there are other demographic characteristics that may be correlated with a preference for climate-related risks, which drive resorting after information disclosure. These preferences are partially correlated with income, but even after controlling for income effects, we find that they continue to significantly change the landscape of who holds climate risk.

5 Discrete-Choice Model and Estimation

Last but not least, in order to characterize the extent of resorting based on flood-disclosure, in this section, we set up and calibrate a discrete-choice model of housing. After fitting the model, we will use a Shapley decomposition to measure how much of the documented resorting is due to prices versus preference heterogeneity.

5.1 Model and Score Moments

In our model, we assume that household i chooses among homes j with utility

$$u_{ij}(\theta) \;=\; -(\alpha_0 + \pi_{p,y} \ln y_i)\, \ln p_j \;+\; (\beta_r + \pi_{r,\text{age}} \cdot \text{age}_i + \pi_{r,\text{dem}} \cdot \text{is_dem}_i)\, \text{is-risky}_j \;+\; \beta_{\text{run}} \cdot \text{running}_j \;+\; \varepsilon_{ij},$$

The variable *running* (which we refer to as "expected flood depth") is the underlying predicted risk, which is a continuous measure. The variable *is-risky* is a binary variable indicating whether the property is disclosed to be of high risk (i.e., Flood Factor 9). The error ε_{ij} is i.i.d. Type-I extreme value, implying standard logit choice probabilities $P_{ij}(\theta)$.

We estimate the parameters via GMM. Let c_{ij} be the indicator for the observed choice and define

the per-alternative derivative vector

$$z_{ij} \equiv \frac{\partial u_{ij}}{\partial \theta} \;=\; \begin{bmatrix} -\ln p_j \\ -\ln y_i \;\ln p_j \\ \text{is-risky}_j \\ \text{age}_i \;\times \text{is-risky}_j \\ \text{is_dem}_i \;\times \text{is-risky}_j \\ \text{running}_j \end{bmatrix}.$$

The household score contribution is $g_i(\theta) = \sum_j z_{ij}\,(c_{ij} - P_{ij}(\theta))$ and the sample moments are $G(\theta) = \frac{1}{N}\sum_i g_i(\theta)$. The GMM objective is

$$Q(\theta) \;=\; G(\theta)'\,W\,G(\theta).$$

5.2 Estimation Results

In terms of the sample, we focus on the set of homes within the bandwidth of homes near the 8 vs. 9 cutoff in 2024 with price and income residualized to block-group fixed effects and hedonic characteristics.

Since we have detailed micro-data on household choices, the coefficients capturing preference heterogeneity are well-identified. With N = 1,146 households in the post-disclosure period, Table 10 reports the estimated heterogeneity parameters.

Both heterogeneity parameters $\pi_{p,y}$ and $\pi_{r,\text{age}}$ are precisely estimated and statistically significant. The parameter $\pi_{p,y} = -2.14$ (s.e. $= 0.28$, $p < 0.001$) indicates that price sensitivity decreases substantially as household income increases. The parameter $\pi_{r,\text{age}} = 0.075$ (s.e. $= 0.044$, $p = 0.087$) suggests that tolerance for homes disclosed as risky (i.e., Flood Factor of 9) increases with age. The parameter $\pi_{r,\text{dem}} = -0.082$ (s.e. $= 0.123$, $p = 0.506$) is not statistically significant, consistent with our earlier RD results showing weak effects of political affiliation.

Additional Calibration of Mean Parameters

While the heterogeneity parameters $(\pi_{p,y}, \pi_{r,\text{age}}, \pi_{r,\text{dem}})$ are directly estimated from the choice data, we calibrate the mean-level parameters (α_0, β_r) to match two external targets.

In other words, while our model and data can identify the heterogeneity in preferences over housing characteristics, it does not identify the mean coefficients, and thus cannot be used to re-compute counterfactual equilibrium. Importantly, since the decomposition of resorting is primarily driven by the heterogeneous coefficients, we do this calibration exercise to simply ensure that the mean

coefficients used are of a comparable order of magnitude to existing estimates in the literature, and we ensure that the main findings of our decomposition analysis are not sensitive to the calibration exercise.

Concretely, we first calibrate the mean price coefficient α_0 to match an own-price elasticity of -0.67, consistent with Albouy et al. (2016). Then, we also calibrate the mean coefficient on risky properties β_r to match the estimated -3.26% price differential between risky and safe properties in the post-disclosure period. To properly caveat, these calibrations acknowledge that the "average" household in our choice model naturally differs from the marginal household determining equilibrium prices.

5.3 Three-Way Shapley Decomposition of Resorting

With the estimated model, we use the parameter estimates to assess how much observable redistribution of flood risk is due to prices versus preference heterogeneity. We employ a three-way Shapley decomposition that separately accounts for (1) price changes, (2) age-based preference heterogeneity, and (3) political preference heterogeneity.

5.3.1 Decomposition Method

We compare eight scenarios representing all combinations of three binary switches: the estimated price wedge due to disclosure (PRE vs. POST), age heterogeneity (OFF vs. ON), and political heterogeneity (OFF vs. ON):

M_{000} : (price=PRE, age=OFF, dem=OFF), M_{001} : (price=PRE, age=OFF, dem=ON),
M_{010} : (price=PRE, age=ON, dem=OFF), M_{011} : (price=PRE, age=ON, dem=ON),
M_{100} : (price=POST, age=OFF, dem=OFF), M_{101} : (price=POST, age=OFF, dem=ON),
M_{110} : (price=POST, age=ON, dem=OFF), M_{111} : (price=POST, age=ON, dem=ON).

For any metric M (here we use $\mathrm{Cov}(\ln \text{income}, \text{expected flood depth})$), the total change is $\Delta M = M_{111} - M_{000}$. The three-way Shapley decomposition allocates this total change to the three channels

by averaging over all possible orderings in which the switches are turned on. Specifically:

$$\Delta M_{\text{price}} = \tfrac{1}{6}\Big[(M_{100} - M_{000}) + (M_{101} - M_{001}) + (M_{110} - M_{010}) + (M_{111} - M_{011}) + (M_{110} - M_{010}) + (M_{111} - M_{011})\Big],$$

$$\Delta M_{\text{age}} = \tfrac{1}{6}\Big[(M_{010} - M_{000}) + (M_{011} - M_{001}) + (M_{110} - M_{100}) + (M_{111} - M_{101}) + (M_{110} - M_{100}) + (M_{111} - M_{101})\Big],$$

$$\Delta M_{\text{dem}} = \tfrac{1}{6}\Big[(M_{001} - M_{000}) + (M_{011} - M_{010}) + (M_{101} - M_{100}) + (M_{111} - M_{110}) + (M_{101} - M_{100}) + (M_{111} - M_{110})\Big].$$

In words, each channel's contribution is computed by averaging its marginal effect across all possible states of the other two channels. This ensures that the decomposition is symmetric and the three contributions sum exactly to the total change: $\Delta M = \Delta M_{\text{price}} + \Delta M_{\text{age}} + \Delta M_{\text{dem}}$.

5.3.2 Results and Interpretation

Table 11 presents the three-way decomposition results. We focus on the covariance between log household income and expected flood depth (the continuous running variable). In the calibrated specification, the total change in this covariance is -0.0067, indicating that after disclosure, lower-income households are more exposed to flood risk.

The decomposition reveals that price changes account for 91.8% of this resorting effect, with a contribution of -0.0061. This dominant role of prices reflects the fact that risky properties became relatively cheaper after disclosure, and lower-income households, who are more price-sensitive, disproportionately selected into these properties.

Age-based preference heterogeneity contributes 10.7% (-0.0007) to the total resorting. This positive contribution arises because older households exhibit less aversion to properties labeled as risky ($\pi_{r,\text{age}} > 0$), and older households tend to have slightly lower income in our sample. This age-based sorting thus reinforces the concentration of risk among lower-income households.

On the other hand, preference heterogeneity by political affiliation contributes -2.5% ($+0.0002$) to the total resorting. The negative sign indicates that this channel actually *weakens* the resorting of lower-income households into riskier properties. This occurs because Democrats exhibit stronger aversion to risky properties ($\pi_{r,\text{dem}} < 0$), and Democrats also have lower average income in our sample. This countervailing force slightly offsets the income-based resorting driven by prices.

Overall, the decomposition demonstrates that the redistributional impact of flood-risk disclosure

primarily operates through the price channel, with preference heterogeneity playing a much smaller role. The uncalibrated estimates (Table 11) yield nearly identical conclusions, with price changes explaining 93.1% of the resorting, confirming that our results are not sensitive to the mean coefficients.

6 Conclusion

Flood risk is one of the largest risks that homeowners face today, and yet, it remains unclear whether consumers are aware of and incorporate information on flood risk when they are choosing to buy a house. In this paper, we take advantage of a novel information release (ie the Flood Factor) by First Street on several major search platforms as a natural experiment to test how homebuyers respond to flood disclosure across the United States.

Following disclosure, preliminary findings suggest that homebuyers discount homes with flood risk by 3% on average and up to 15% for coastal properties outside of FEMA floodsones. Notably, this discount is largest for homes with the highest (ie "above severe" or "extreme") flood risk, and no price discount is observed for homes with lower and less salient levels of risk. Interestingly, we also find that flood information affects the demographics of homebuyers in equilibrium. Following disclosure, we find that homebuyers of the riskiest homes are lower income and more likely to hold an FHA loan. This suggests the costs of climate change may be borne disproportionately by the more financially constrained. Even controlling for income, we find some evidence that resorting occurs based age and political affiliation as well. Finally, using a discrete-choice housing model, we find that prices are the first-order driver of income-based resorting and preference heterogeneity is the first-order driver of age-based resorting.

Going forward, more work needs to be done in order to disentangle the various factors that drive resorting after flood disclosure. One potential path forward is to create a structural model that incorporates both an income effect as well as orthogonal taste-based preferences, as proxied by age and political affiliation. This will also allow us to simulate counterfactual future scenarios, such as an aging homebuyer population or further polarization of political beliefs. Our findings have potential policy implications. As technology improves and new data on climate change becomes more readily available, our findings may help predict the equilibrium impact of new flood data on consumers of different types as well as consumers living in different regions of the United States.

References

Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel (2015) "Regulating Consumer Financial Products: Evidence from Credit Cards *," *The Quarterly Journal of Economics*, 130 (1), 111–164, 10.1093/qje/qju037.

Albouy, David, Gabriel Ehrlich, and Yingyi Liu (2016) "Housing Demand, Cost-of-Living Inequality, and the Affordability Crisis,"Technical Report w22816, National Bureau of Economic Research, Cambridge, MA, 10.3386/w22816.

Bailey, Michael, Ruiqing Cao, Theresa Kuchler, and Johannes Stroebel (2018) "The economic effects of social networks: Evidence from the housing market," *Journal of Political Economy*, 126 (6), 2224–2276.

Baldauf, Markus, Lorenzo Garlappi, and Constantine Yannelis (2020) "Does Climate Change Affect Real Estate Prices? Only If You Believe In It," *The Review of Financial Studies*, 33 (3), 1256–1295, 10.1093/rfs/hhz073.

Bayer, Patrick, Marcus Casey, W Ben McCartney, John Orellana-Li, and Calvin Zhang (2025) "Distinguishing Causes of Neighborhood Racial Change: A Nearest-Neighbor Design," *American Economic Review*, 115 (11), 3999–4039.

Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins (2016) "A dynamic model of demand for houses and neighborhoods," *Econometrica*, 84 (3), 893–942.

Bell, Alex, Stephen B Billings, Sophie Calder-Wang, and Shusheng Zhong (2024) "An Anti-IV Approach for Pricing Residential Amenities: Applications to Flood Risk," *Available at SSRN*.

Beltrán, Allan, David Maddison, and Robert Elliott (2019) "The impact of flooding on property prices: A repeat-sales approach," *Journal of Environmental Economics and Management*, 95, 62–86.

Bernstein, Asaf, Stephen B. Billings, Matthew T. Gustafson, and Ryan Lewis (2022) "Partisan residential sorting on climate change risk," *Journal of Financial Economics*, 146 (3), 989–1015, 10.1016/j.jfineco.2022.03.004.

Bin, Okmyung and Craig E Landry (2013) "Changes in implicit flood risk premiums: Empirical evidence from the housing market," *Journal of Environmental Economics and management*, 65 (3), 361–376.

Buchak, Greg, Gregor Matvos, Tomasz Piskorski, and Amit Seru (2022) "Why is Intermediating Houses so Difficult? Evidence from iBuyers," July, 10.2139/ssrn.3616555.

Choi, Darwin, Zhenyu Gao, and Wenxi Jiang (2020) "Attention to global warming," *The Review of Financial Studies*, 33 (3), 1112–1145.

Cookson, J Anthony, Emily Gallagher, and Philip Mulder (2025) "Coverage Neglect in Homeowner's Insurance," *Working Paper*.

Dessaint, Olivier and Adrien Matray (2017) "Do managers overreact to salient risks? Evidence from hurricane strikes," *Journal of Financial Economics*, 126 (1), 97–121.

Eby, Matthew (2019) "Understanding FEMA Maps and Limitations,"Technical report, First Street.

Fairweather, Daryl, Matthew E. Kahn, Robert D. Metcalfe, and Sebastian Sandoval Olascoaga (2024) "Expecting Climate Change: A Nationwide Field Experiment in the Housing Market," November, 10.3386/w33119.

Fu, Runshan, Ginger Zhe Jin, and Meng Liu (2022) "Does Human-algorithm Feedback Loop Lead to Error Propagation? Evidence from Zillow's Zestimate," March, 10.3386/w29880.

Fuster, Andreas, Paul Goldsmith-Pinkham, Tarun Ramadorai, and Ansgar Walther (2022) "Predictably unequal? The effects of machine learning on credit markets," *The Journal of Finance*, 77 (1), 5–47.

Gallagher, Justin (2014) "Learning about an infrequent event: Evidence from flood insurance take-up in the United States," *American Economic Journal: Applied Economics*, 206–233.

Ge, Shan, Stephanie Johnson, and Nitzan Tzur-Ilan (2025) "Climate Risk, Insurance Premiums and the Effects on Mortgage and Credit Outcomes,"Technical Report No. 2505, Federal Reserve Bank of Dallas Working Paper, Available at RePEc: RePEc:fip:feddwp:99475.

Ge, Shan, Ammon Lam, and Ryan Lewis (2022) "The Effect of Insurance Premiums on the Housing Market and Climate Risk Pricing,"Technical report, SSRN.

Gibson, Matthew, Jamie Mullins, and Alison Hill (2017) "Climate change, flood risk, and property values: Evidence from New York City," *Working Paper*.

Harrison, David M, Michael J Seiler, and Liuming Yang (2024) "The impact of iBuyers on housing market dynamics," *The Journal of Real Estate Finance and Economics*, 68 (3), 425–461.

Heinrich, Martin (2024) "Flooding Costs the U.S. Between $179.8 and $496.0 Billion Each Year,"Technical report, Joint Economic Committee.

Humphries, Samuel (2020) "A Flooded Future: New Flood Model Exposes a Grim Reality,"Technical report, American Security Project.

Keys, Benjamin J. and Philip Mulder (2024) "Property Insurance and Disaster Risk: New Evidence from Mortgage Escrow Data,"Technical Report No. 32579, National Bureau of Economic Research (NBER) Working Paper.

Krueger, Paul, Zacharias Sautner, and Laura Starks (2020) "The Importance of Climate Risks for Institutional Investors," *Review of Financial Studies*, 33 (3), 1067–1101.

Meyer, Robinson (2024) "Should You Trust Zillow's Climate Risk Data?"Technical report, Heatmap.

Mulder, Philip (2024) "Mismeasuring risk: The welfare effects of flood risk information," *Available at SSRN 4966795*.

Murfin, Justin and Matthew Spiegel (2020) "Is the Risk of Sea Level Rise Capitalized in Residential Real Estate?" *The Review of Financial Studies*, 33 (3), 1217–1255, 10.1093/rfs/hhz134.

Ortega, Francesc and Süleyman Taşpınar (2018) "Rising sea levels and sinking property values: Hurricane Sandy and New York's housing market," *Journal of Urban Economics*, 106, 81–100.

Raymond, Lindsey (2024) "The Market Effects of Algorithms."

Sastry, Parinitha, Ishita Sen, and Ana-Maria Tenekedjieva (2023) "When insurers exit: Climate losses, fragile insurers, and mortgage markets," *Fragile Insurers, and Mortgage Markets (December 23, 2023)*.

Sastry, Parinitha, Ishita Sen, Ana-Maria Tenekedjieva, and Therese C Scharlemann (2024) "Climate risk and the us insurance gap: Measurement, drivers and implications," *Drivers and Implications*(May 31, 2024)*.

Stroebel, Johannes and Jeffrey Wurgler (2021) "What do you think about climate finance?."

Zillow (2023) "More than 80% of home shoppers consider climate risks,"Technical report, Zillow.

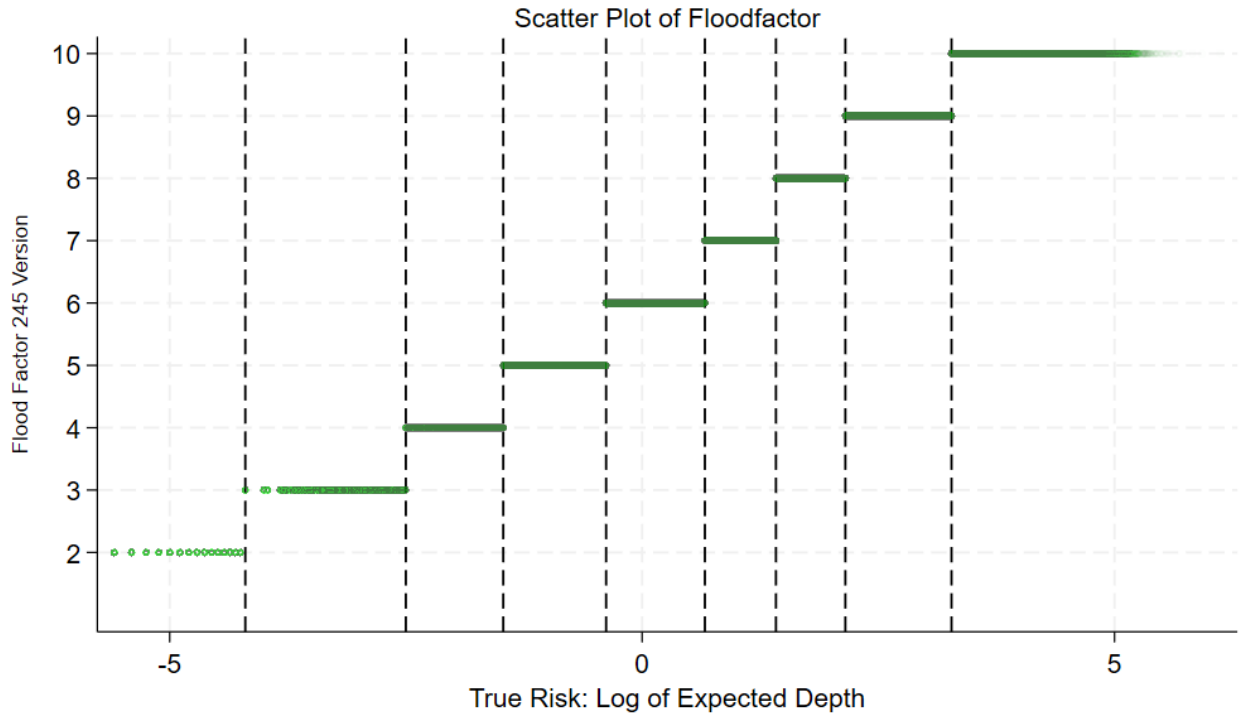
Figures

1. Background on First Street Flood Factor

(A) Mapping of Continuous Risk to Discrete Flood Scores

Figure 1: Cutoff Chart of Flood Factor

In order to estimate a regression discontinuity analysis, we back out the undisclosed running variable (shown on x-axis) that is then mapped into the discrete Flood Factor categories publicly disclosed by First Street (shown on y-axis). The running variable is a model-based proxy for continuous flood risk, estimated as the expected flood depth over the next 100 years.



(B) Disclosure of First Street Flood Factor

Figure 2: Disclosure of First Street Flood Factor on Housing Platforms

This figure shows the main timeline of when First Street flood scores were updated, as well as when they rolled out onto the Realtor.com, Redfin.com, and Zillow.com platforms. The dotted vertical lines and grey arrows further show the pre-period and post-period chosen for our empirical analyses.

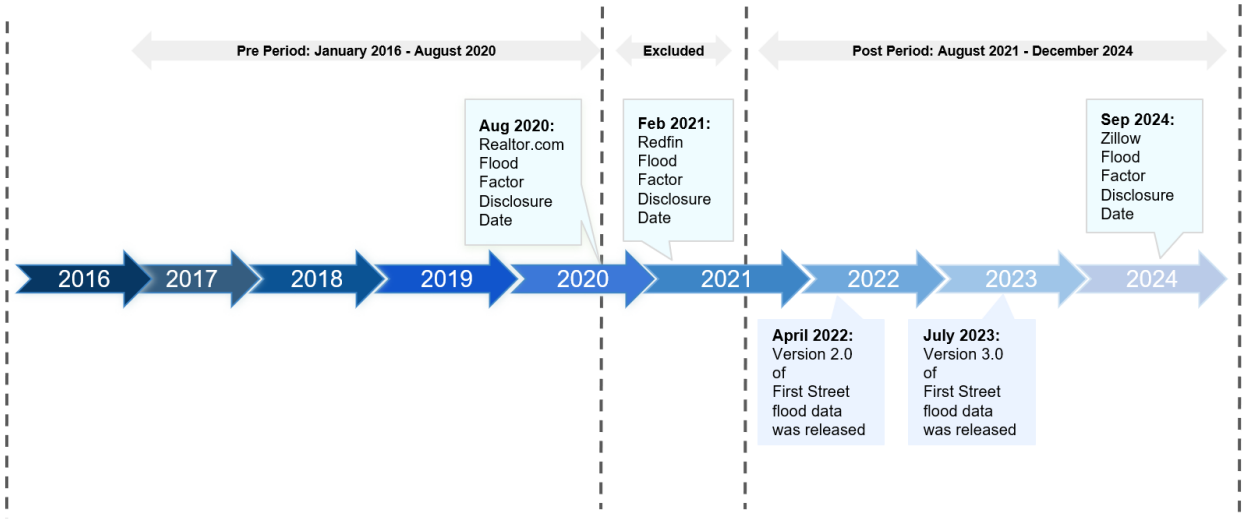


Figure 3: Example of Flood Score as Shown on Redfin

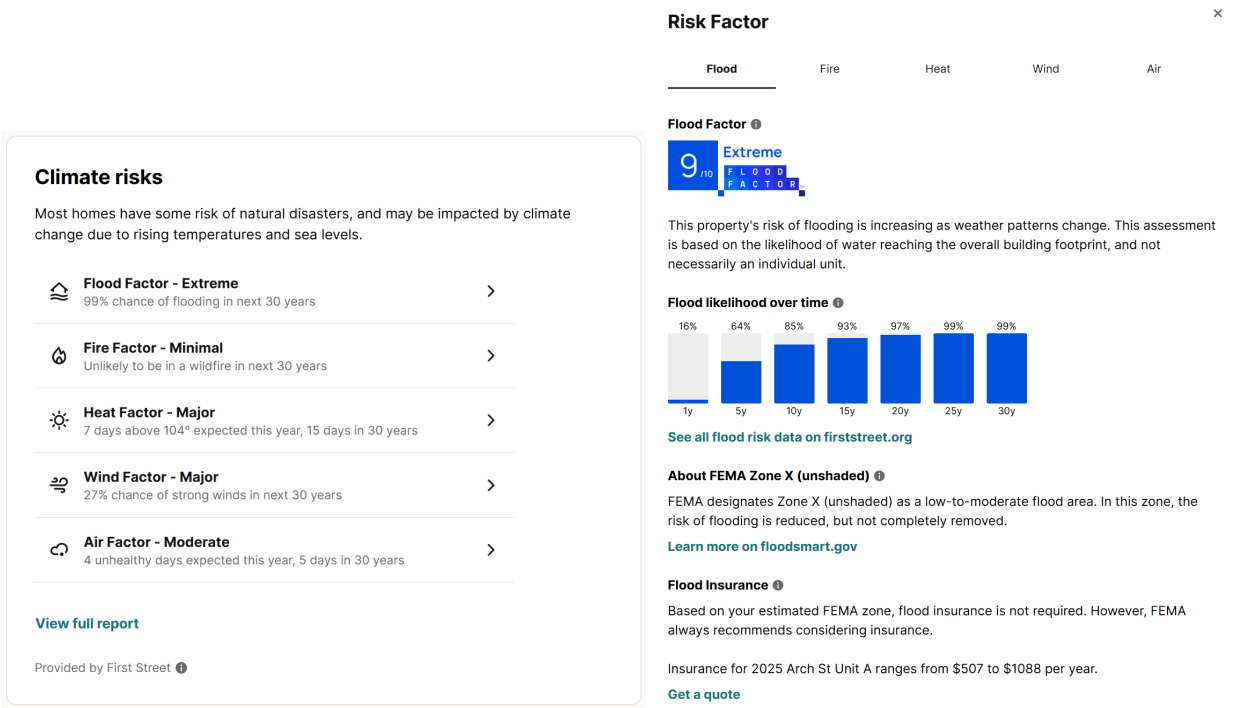
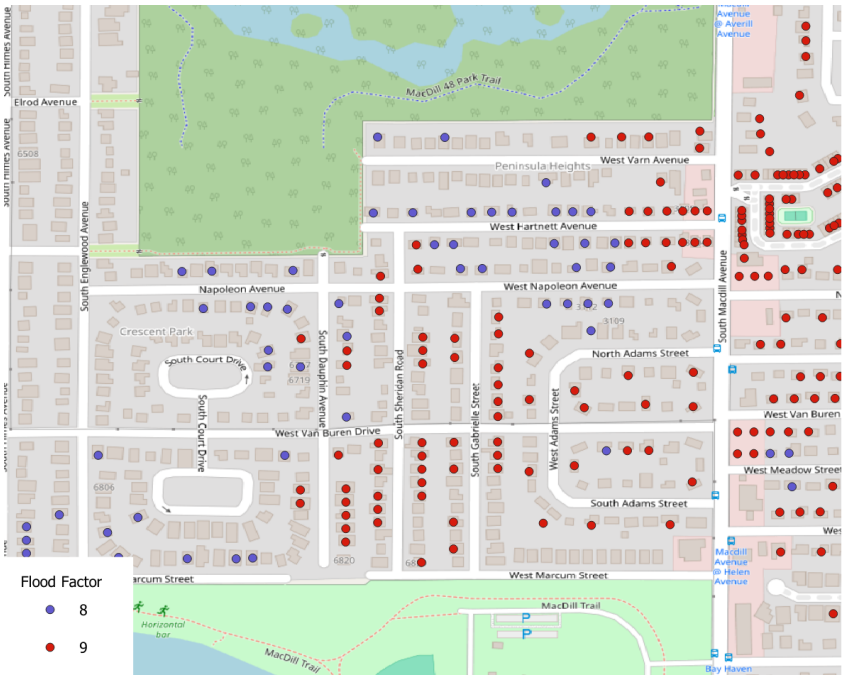
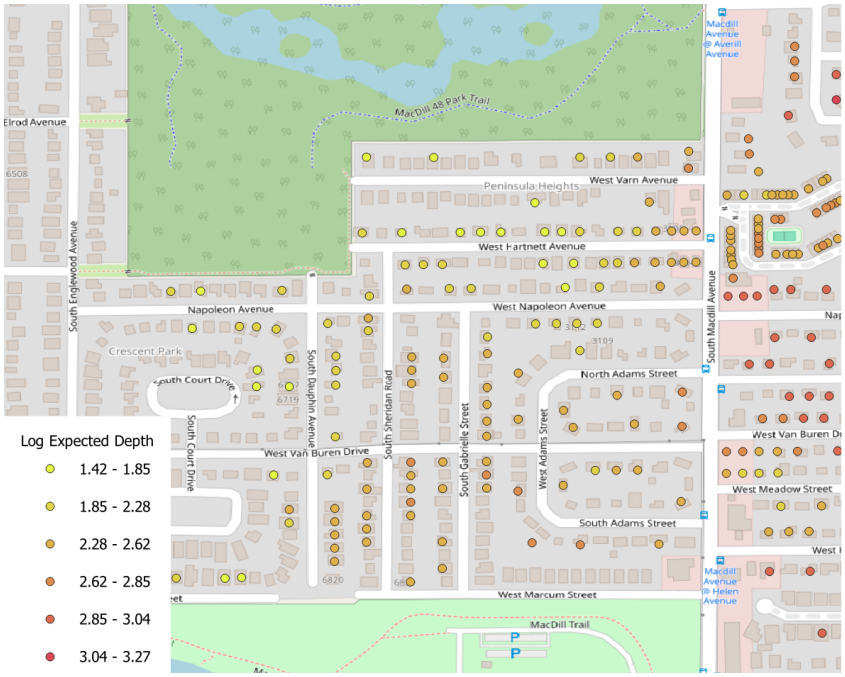


Figure 4: Close-up Map of Properties Based on Flood Risk in Hillsborough County

These maps display census tract-level shapefile for a zoomed in selection of community from Hillsborough County, Florida. Panel a shows all properties that are with extreme flood risk as dots colored based on the severity of their First Street Flood Factor: properties in risk factor 8(9) as a red (blue) dot, while Panel b shows the properties as dots colored based on the running variable(expected depth). We limit the sample of house transactions to the pre-period (January 2016–August 2020) and the post-period (August 2021–December 2024). Our data comes from First Street and Corelogic, and our data sample covers the years 2016 to 2024.



(a) Flood risk Factor



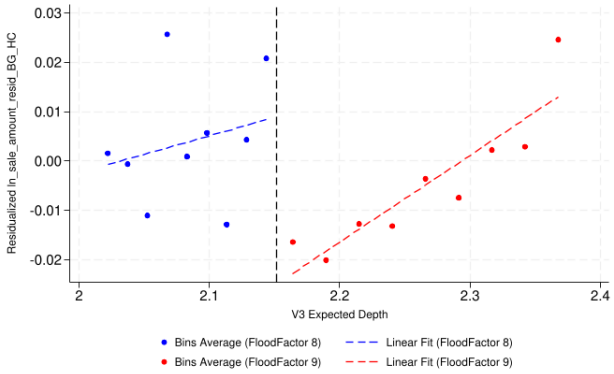
(b) Log Expected Depth

2. Regression Discontinuity

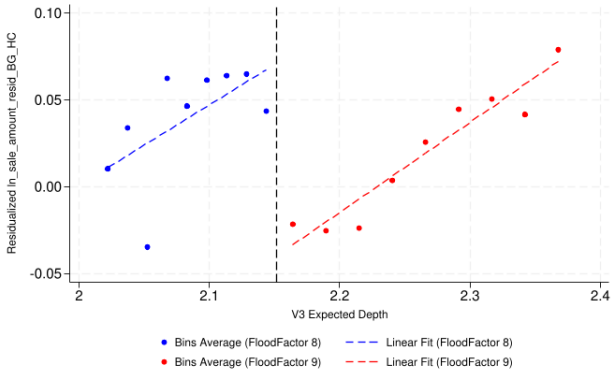
(A) Visualizing RD Regressions

Figure 5: Full Sample RD Plots for Price

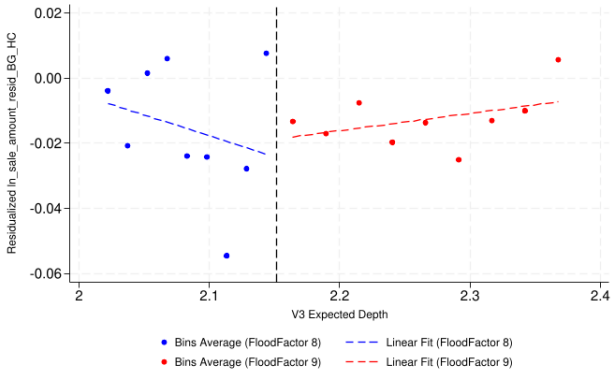
The figures below visually show a linear RD Analysis with auto-selected bandwidths using the (Panel a) full data sample,(Panel b) coastal Sample (Panel c) non coastal Sample explaining log price (Panel a). To control for omitted variable bias, the LHS variable is residualized with block-group fixed effects and basic hedonics. Each dot represents the difference between the average log home prices in the post-period minus the average log home prices in the pre-period. The point on the x-axis represents the midpoint of each equal-width bin, and the bars represent the standard errors. Our data comes from First Street, Corelogic, and HMDA, and our data sample covers the years 2016 to 2024.



(a) RD for Log Price: Full Sample



(b) RD for Log Price: Coastal Sample



(c) RD for Log Price: NonCoastal Sample

Figure 6: Pretend Test for Log Home Prices

To test whether there are notable pre-trends in the prices of flood-prone properties prior to disclosure, each point on the graph represents the coefficient β_1 and associated standard errors from the following cross sectional rd regressions: For each year t, we estimate the rd model on a single cross-sectional sample: $\ln(\textbf{SalePrice}_i) = \alpha + \beta_1 \mathbf{1}\{\textbf{FloodFactor}_i = 9\} + \beta_2 \textbf{Expected Depth} + \beta_3 \mathbf{1}\{\textbf{FloodFactor}_i = 9\} \times \textbf{Expected Depth} + \varepsilon_i$ Each point corresponds to a sample that starts at the date shown on the x-axis and ends one year later. Our data comes from First Street and Corelogic, and our data sample for this analysis covers the years 2016 to 2024. The excluded period is start from August 2020 to August 2021.

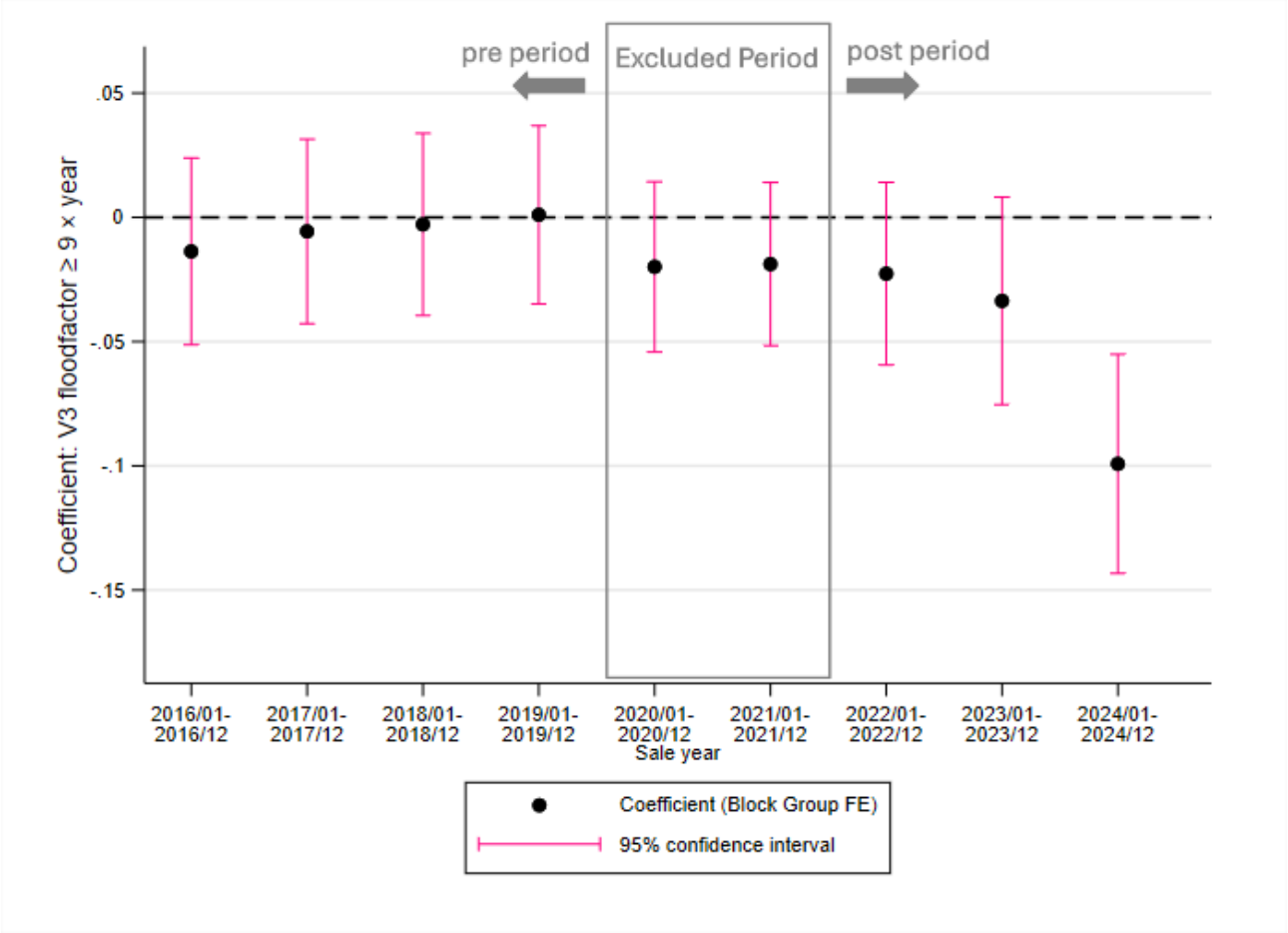


Figure 7: Distribution of Distance to the Discontinuity (Bunching Test)

The figure shows the distribution of the running variable (distance to the discontinuity) around the cutoff. Each bar represents the frequency of observations within equal-width bins. The vertical dashed line marks the cutoff at zero. This plot is used to assess whether the running variable is smoothly distributed around the cutoff, providing a visual diagnostic for potential sorting or manipulation. The sample includes observations with non-missing outcome variables used in the main regression analysis. Our data comes from First Street, Corelogic, and HMDA. The post-disclosure period spans 2021 to 2024.

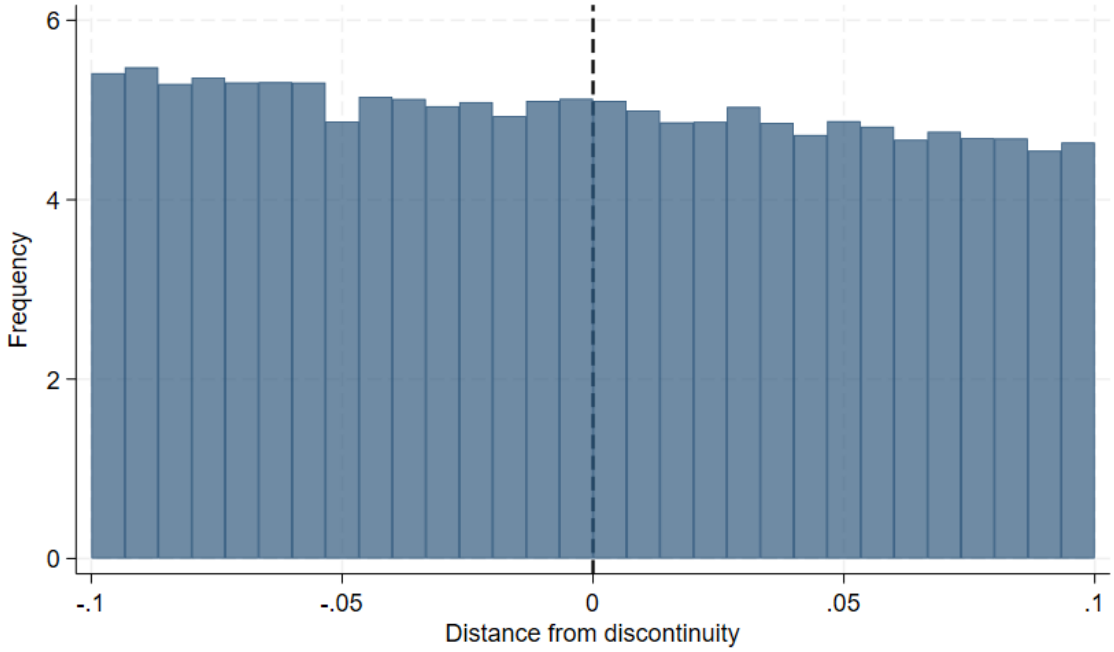
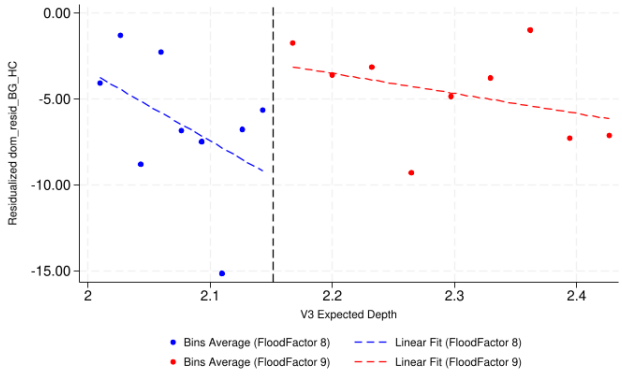
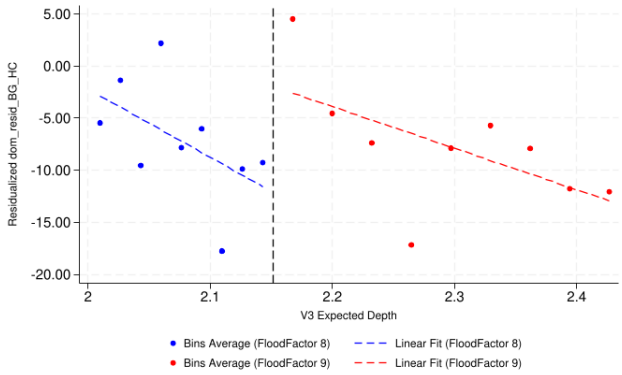


Figure 8: Full Sample RD Plots for Days on Market

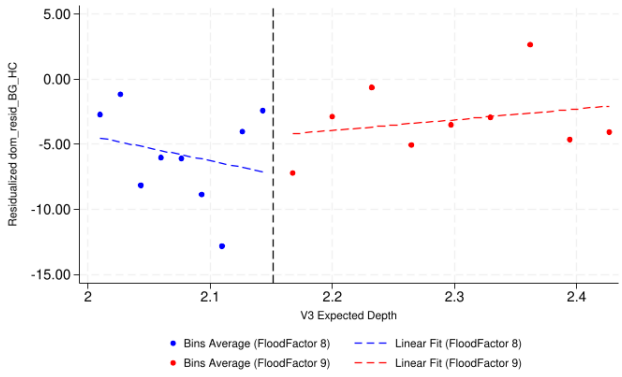
The figures below visually show a linear RD Analysis with auto-selected bandwidths using the (Panel a) full data sample,(Panel b) coastal Sample (Panel c) non coastal Sample explaining days on market. To control for omitted variable bias, the LHS variable is residualized with block-group fixed effects and basic hedonics. Each dot represents the difference between the average days on market in the post-period minus the average days on market in the pre-period. The point on the x-axis represents the midpoint of each equal-width bin, and the bars represent the standard errors. Our data comes from First Street and Corelogic, and our data sample covers the years 2016 to 2024.



(a) RD for Log Price: Full Sample



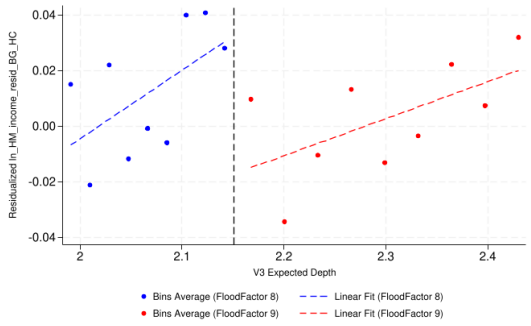
(b) RD for Log Price: Coastal Sample



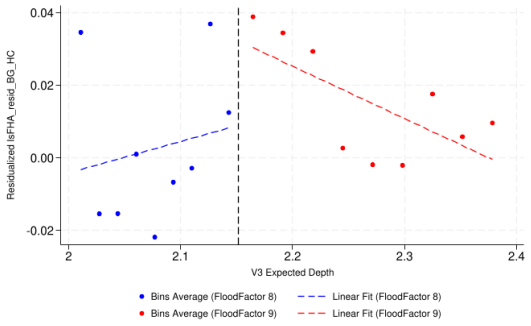
(c) RD for Log Price: NonCoastal Sample

Figure 9: Full Sample RD Plots for Income, Age, Democrat and FHA Status

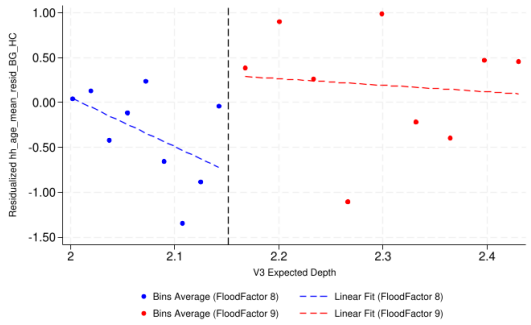
The figures below visually show a linear RD Analysis with auto-selected bandwidths using the full data sample, explaining: log income (Panel a) dummy for whether the applicant is FHA (Panel b), average household age (Panel c), and dummy for democrat (Panel d) . To control for omitted variable bias, the LHS variable is residualized with block-group fixed effects and basic hedonics. Each dot represents the difference between the average log home prices in the post-period minus the average log home prices in the pre-period. The point on the x-axis represents the midpoint of each equal-width bin, and the bars represent the standard errors. Our data comes from First Street, Corelogic, and HMDA, and our data sample covers the years 2016 to 2024.



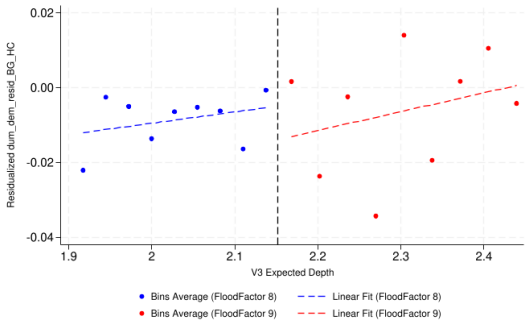
(a) RD for Log Income



(b) RD for Is FHA Dummy



(c) RD for Average Household Age



(d) RD for Democrat Dummy

Tables

Part 1. Summary Statistics

Table 1: Summary Statistics

This table summarizes the key variables used in our DID and RD Regressions. Flood risk data is sourced from First Street, while housing price and property characteristics data come from CoreLogic. Our demographic variables come from HMDA. Our data sample covers the years 2016 to 2024. We observe fewer observations for the expected-depth variable because properties classified as minimal flood risk have no expected-depth values.

	Count	Mean	SD	p25	p50	p75
CoreLogic:						
Sale Price ($k)	25,677,071	372.17	349.62	172.00	280.00	439.90
Living Area (Sqft)	25,677,071	1,785.64	992.81	1,232.00	1,680.00	2,289.00
Lot Size (Sqft)	25,677,071	20,063.94	37,422.48	6,325.00	9,400.00	16,117.00
Number of Bathrooms	25,677,071	1.97	1.23	1.00	2.00	3.00
Number of Bedrooms	25,677,071	2.27	1.66	0.00	3.00	3.00
Building Age	25,677,071	32.50	30.53	6.00	25.00	52.00
Sale Year	25,677,071	2,019.70	2.58	2,017.00	2,019.00	2,022.00
Days on Market	10,144,325	47.17	59.62	7.00	24.00	63.00
First Street:						
V3 Expected Depth (cm)	4,809,448	4.69	11.31	0.12	0.76	3.83
Flood Source include Coastal	24,484,572	0.05	0.21	0.00	0.00	0.00
Flood Source include Pluvial	24,484,572	0.13	0.33	0.00	0.00	0.00
HMDA Variables:						
Age Above 55	8,362,975	0.17	0.37	0.00	0.00	0.00
Applicant Race: NonWhite	8,362,975	0.39	0.49	0.00	0.00	1.00
Applicant Income ($)	8,271,936	128,936.07	143,332.12	65,000.00	98,000.00	150,000.00
Loan Type: Us FHA	8,335,645	0.17	0.38	0.00	0.00	0.00
Voter Registration (per household):						
Average Age	10,970,839	43.71	13.83	33.00	40.00	52.00
Dummy for a Registered Democrat	10,970,839	0.39	0.49	0.00	0.00	1.00
Dummy for a Registered Republican	10,970,839	0.39	0.49	0.00	0.00	1.00
Dummy for a College Graduate	10,970,839	0.43	0.49	0.00	0.00	1.00

Table 2: Summary Statistics - Flood Factor

This table tabulates homes' First Street risk factor against whether they are classified as being SFHA. Each observation is one house sale transaction. Flood risk data is sourced from First Street, while FEMA and home sale data comes from CoreLogic. Our data sample covers the years 2016 to 2024.

First Street Flood Factor	First Street Risk Category	Count	Fraction	Fraction in SFHA
1	Minimal	20,595,805	84.12%	0.82%
2	Minor	232,698	0.95%	7.03%
3	Moderate	520,158	2.12%	6.71%
4	Moderate	401,923	1.64%	7.85%
5	Major	455,635	1.86%	8.76%
6	Major	785,514	3.21%	11.29%
7	Severe	439,527	1.80%	16.29%
8	Severe	461,468	1.88%	24.39%
9	Extreme	424,407	1.73%	44.65%
10	Extreme	167,437	0.68%	62.01%
Total		24,484,572	100 %	3.51%

Table 3: Correlative Evidence based on DID Regression

This table presents progressive difference-in-difference regressions of the log of sale amount on FloodFactor. Regression data is constructed from flood factor(1–10), across the pre-period (January 2016 to August 2020) and post-period (August 2021 to December 2024). FloodFactor groups properties into six categories: Minimal (1), Minor (2), Moderate (3-4), Major (5-6), Severe (7-8), and Extreme (9-10). Observations with sale prices below $50,000 and properties with more than one building are excluded, and sale prices are winsorized at the 99th percentile. All hedonic controls are winsorized at the 1st and 99th percentile. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Flood risk data is sourced from First Street, while housing price and property characteristics data come from CoreLogic, covering transactions from 2007 onward. Significance levels are denoted by ***, **, and * for 1%, 5%, and 10%, respectively.

	Log Sale Price - Full V2					
	(1)	(2)	(3)	(4)	(5)	(6)
Risk is Above Severe × Post Disclosure	-0.00321	-0.0140***	-0.0118***	-0.0106***	-0.0103***	-0.00894***
	(-1.29)	(-7.10)	(-6.68)	(-6.04)	(-6.21)	(-5.37)
Risk is Above Major × Post Disclosure	0.00218	-0.00393**	-0.00159	-0.00143	-0.00153	-0.00107
	(1.06)	(-2.39)	(-1.12)	(-1.03)	(-1.18)	(-0.82)
Risk is Above Moderate × Post Disclosure	-0.00417**	-0.00137	0.00159	0.00163	0.00173	0.00194
	(-2.02)	(-0.85)	(1.17)	(1.22)	(1.38)	(1.55)
Risk is Above Minor × Post Disclosure	-0.0215***	-0.00575**	-0.00135	-0.00182	-0.00234	-0.00225
	(-6.19)	(-2.11)	(-0.60)	(-0.83)	(-1.14)	(-1.10)
Risk is Above Minimal × Post Disclosure	0.0458***	0.00396	-0.00152	-0.00138	-0.000593	-0.000433
	(14.64)	(1.60)	(-0.74)	(-0.68)	(-0.31)	(-0.23)
Obs	24484572	24484453	24480349	24464328	24464328	24464328
R2	0.0465	0.4192	0.6187	0.6569	0.6973	0.6973
R2-Adj	0.0465	0.4191	0.6164	0.6511	0.6921	0.6921
County-Period FE		✓				
Tract-Period FE			✓			
Block Group-Period FE				✓	✓	✓
Year-Month FE	✓	✓	✓	✓	✓	✓
SFHA FE						✓
Hedonic Controls					✓	✓

Part 2. Regression Discontinuity

Table 4: RD: Flood Risk Impact on Sale Prices

This table presents the regression discontinuity results examining the relationship between log sale price and expected flood depth after First Street. Each panel represents a different pair of flood risks, and for each pair, we show the RD estimate from three specifications: a linear fit within a auto-selected bandwidth around the cutoff (column 1), a polynomial fit of order 2 (column 2), and a polynomial fit of order 3 (column 3). Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Housing data comes from Corelogic while flood risk data comes from First Street. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Our time sample spans 2016 to 2024. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

Dep. Var: Residualized Log Sale Price	Linear (Auto Bandwidth) (1)	2nd Order (2)	3rd Order (3)
	Panel A: Risk Factor 3 vs 4		
FloodFactor = 4 × Post Disclosure	-0.0169	-0.0190**	-0.0271**
	(-1.59)	(-2.23)	(-2.44)
Obs	102503	246225	246225
	Panel B: Risk Factor 4 vs 5		
FloodFactor = 5 × Post Disclosure	-0.0106	-0.0075	-0.0049
	(-0.74)	(-0.73)	(-0.35)
Obs	60407	293224	293224
	Panel C: Risk Factor 5 vs 6		
FloodFactor = 6 × Post Disclosure	0.0067	-0.0036	-0.0022
	(0.64)	(-0.49)	(-0.22)
Obs	142019	487371	487371
	Panel D: Risk Factor 6 vs 7		
FloodFactor = 7 × Post Disclosure	0.0192	0.0169*	0.0184
	(1.57)	(1.86)	(1.49)
Obs	116132	481969	481969
	Panel E: Risk Factor 7 vs 8		
FloodFactor = 8 × Post Disclosure	-0.0266*	-0.0313***	-0.0320**
	(-1.95)	(-3.12)	(-2.35)
Obs	82580	358971	358971
	Panel F: Risk Factor 8 vs 9		
FloodFactor = 9 × Post Disclosure	-0.0330**	-0.0241***	-0.0302**
	(-2.32)	(-2.60)	(-2.39)
Obs	83027	481433	481433
	Panel G: Risk Factor 9 vs 10		
FloodFactor = 10 × Post Disclosure	-0.0335**	-0.0274***	-0.0443***
	(-2.06)	(-2.65)	(-3.21)
Obs	90382	409928	409928
Block Group FE	✓	✓	✓
Year-Month FE	✓	✓	✓
Hedonic Controls	✓	✓	✓

Table 5: RD: Subsample Analysis of Flood Risk Impact on Sale Prices

This table presents the regression discontinuity results, broken down by key subsamples, examining the relationship between log sale price and expected flood depth after First Street. The key subsamples shown are: Full sample (column 1), properties that are exposed to coastal flood risk (column 2), properties that are not exposed to coastal flood risk (column 3), SFHA properties (column 3), and nonSFHA properties (column 4). The regression discontinuity specification shown represents a linear fit within a auto-selected bandwidth around the cutoff between Flood Scores 8 and 9. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Housing data comes from Corelogic while flood risk data comes from First Street. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Our time sample spans 2016 to 2024. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

	Sub-Samples				
Dep. Var: Residualized Log Sale Price	Full	Coastal Only	NonCoastal	SFHA Only	NonSFHA
	(1)	(2)	(3)	(4)	(5)
FloodFactor = 9 × Post Disclosure	-0.0330**	-0.110***	0.00762	-0.0347	-0.0335**
	(-2.32)	(-4.60)	(0.43)	(-1.22)	(-2.05)
Obs	83027	26352	56675	22790	60237
BlockGroup FE	✓	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓	✓

Table 6: Robustness Check: Different Levels of Fixed Effects for 8 vs 9 Homes

For homes that are rated 8 or 9, this table presents the difference in regression discontinuity results examining the relationship between log sale price and expected flood depth using an auto-selected bandwidth around the cutoff. In columns 1, 2, 3 and 4, we residualize the outcome variable with no geographic fixed effects, county fixed effects, tract fixed effects and block group fixed effects, respectively. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Housing data comes from Corelogic while flood risk data comes from First Street. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Our time sample spans 2016 to 2024. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

Dep. Var: Residualized Log Sale Price	(1)	(2)	(3)	(4)
FloodFactor = 9 × Post Disclosure	-0.0120	-0.0261	-0.0325**	-0.0330**
	(-0.52)	(-1.46)	(-2.20)	(-2.32)
Obs	83061	83061	83051	83027
BlockGroup FE	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓
Level of Controls	None	County	Tract	Block Group

Table 7: Balance Test for Covariate Variables

The tables below present regression discontinuity models examining the balance of covariate variables(which we use as hedonic controls in main regressions), including lot size (sqft) (column 1), living area (sqft)(column 2), number of bedroom (column 3), number of bathrooms (column 4), building age (column 5) of properties. In addition, we also examine a dummy equal to one if the property is exposed to a coastal source of flooding (6). All estimates are under specification of linear fit with auto-selected bandwidth around cutoff between Flood Scores 8 and 9. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Our time sample spans 2016 to 2024. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

Dep. Var:	Lot size	Living area	Num of bathrooms	Num of bedrooms	Building age	Coastal
	(1)	(2)	(3)	(4)	(5)	(6)
FloodFactor = 9 × Post Disclosure	4657.0	-23.82	-0.0380	-0.0232	-0.522	-0.000957
	(0.38)	(-1.14)	(-1.55)	(-0.80)	(-1.03)	(-0.61)
Obs	79382	69043	60816	50962	79382	79382
BlockGroup FE	✓	✓	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓	✓	✓

Table 8: Flood Risk Impact on Days on Market, Log Income, and FHA Status

The tables below present regression discontinuity models examining the relationship between the number of days on market (panel A), buyer's log income (panel B), or FHA status (panel c) and expected flood depth. The regression discontinuity specifications include: a linear fit within an auto-selected bandwidth around the cutoff, a polynomial fit of order 2, and a polynomial fit of order 3. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Housing data comes from Corelogic, and demographic information comes from HMDA. Flood risk data comes from First Street. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Our time sample spans 2016 to 2024. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

Panel A. Days on Market

Dep. Var: Residualized Days on Market	Linear(Auto Bandwidth) (1)	2nd Order (2)	3rd Order (3)
FloodFactor = 9 × Post Disclosure	6.756**	5.363***	9.532***
	(2.36)	(2.85)	(3.72)
Obs	32994	200193	200193
Block Group FE	✓	✓	✓
Year-Month FE	✓	✓	✓
Hedonic Controls	✓	✓	✓

Panel B. Log Income

Dep. Var: Residualized Log Income	Linear(Auto Bandwidth) (1)	2nd Order (2)	3rd Order (3)
FloodFactor = 9 × Post Disclosure	-0.0529**	-0.0247	-0.0438*
	(-2.07)	(-1.46)	(-1.91)
Obs	23600	132311	132311
BlockGroup FE	✓	✓	✓
Year-Month FE	✓	✓	✓
Hedonic Controls	✓	✓	✓

Panel C. FHA Status

Dep. Var: Residualized Is FHA	Linear(Auto Bandwidth) (1)	2nd Order (2)	3rd Order (3)
FloodFactor = 9 × Post Disclosure	0.0180	0.0152*	0.0266**
	(1.27)	(1.73)	(2.22)
Obs	41883	236227	236227
Block Group FE	✓	✓	✓
Year-Month FE	✓	✓	✓
Hedonic Controls	✓	✓	✓

Table 9: Flood Risk Impact on Age and Dummy for Democrat

The tables below present regression discontinuity results examining the relationship between the average household age (panel A) or dummy for Democrat in household (panel B) and expected flood depth for both the pre-period (January 2016 to August 2020) and the post-period (August 2021 to December 2024). Column 1 does not include any income controls, while columns 2 through 4 control for income and income interacted with a post-disclosure dummy. The alternate regression discontinuity specifications (columns 3 and 4) include: a linear fit within a auto-selected bandwidth around the cutoff, a polynomial fit of order 2, and a polynomial fit of order 3. The outcome variable is residualized using hedonic controls, block group fixed effects, and sale year-month fixed effects. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

Panel A. Average Household Age

Dep. Var: Residualized Household Age	Linear(Auto Bandwidth) (1)	Linear(Auto Bandwidth) (2)	2nd Order (3)	3rd Order (4)
FloodFactor = 9 × Post Disclosure	1.384**	2.014**	0.880	1.811**
	(2.21)	(2.38)	(1.63)	(2.47)
Obs	32381	14345	81436	81436
Control for Income		✓	✓	✓
County FE	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓

Panel B. Dummy for a Registered Democrat in Household

Dep. Var: Residualized Dummy for Democrat	Linear(Auto Bandwidth) (1)	Linear(Auto Bandwidth) (2)	2nd Order (3)	3rd Order (4)
FloodFactor = 9 × Post Disclosure	-0.00444	0.00777	0.00103	-0.0105
	(-0.44)	(0.35)	(0.07)	(-0.54)
Obs	83027	23600	132311	132311
Control for Income		✓	✓	✓
County FE	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓

Table 10: Model Output: Parameter Estimates

The table below shows the key estimation results of our discrete-choice housing model. For our estimation sample, based on empirical findings, we focus the set of homes within the bandwidth of homes near the 8 vs. 9 cutoff in 2024.

Parameter	Estimate	Std. Error	p-value
$\pi_{p,y}$	-2.1279	(0.2796)	0.0000
$\pi_{r,\text{age}}$	0.0749	(0.0438)	0.0876
$\pi_{r,\text{dem}}$	-0.0822	(0.1234)	0.5055

Table 11: Model Output: Shapely Decomposition

This table reports the three-way Shapley decomposition of Cov(ln income, expected flood depth) into contributions from price changes, age-based preference heterogeneity, and political preference heterogeneity. "Calibrated" refers to estimates where the mean price coefficient is calibrated to match an elasticity of -0.67 Albouy et al. (2016) and the mean risky-property coefficient is calibrated to match an estimated -3.26% price effect. "Uncalibrated" reports raw GMM estimates. Expected flood depth is the continuous running variable measuring underlying flood risk.

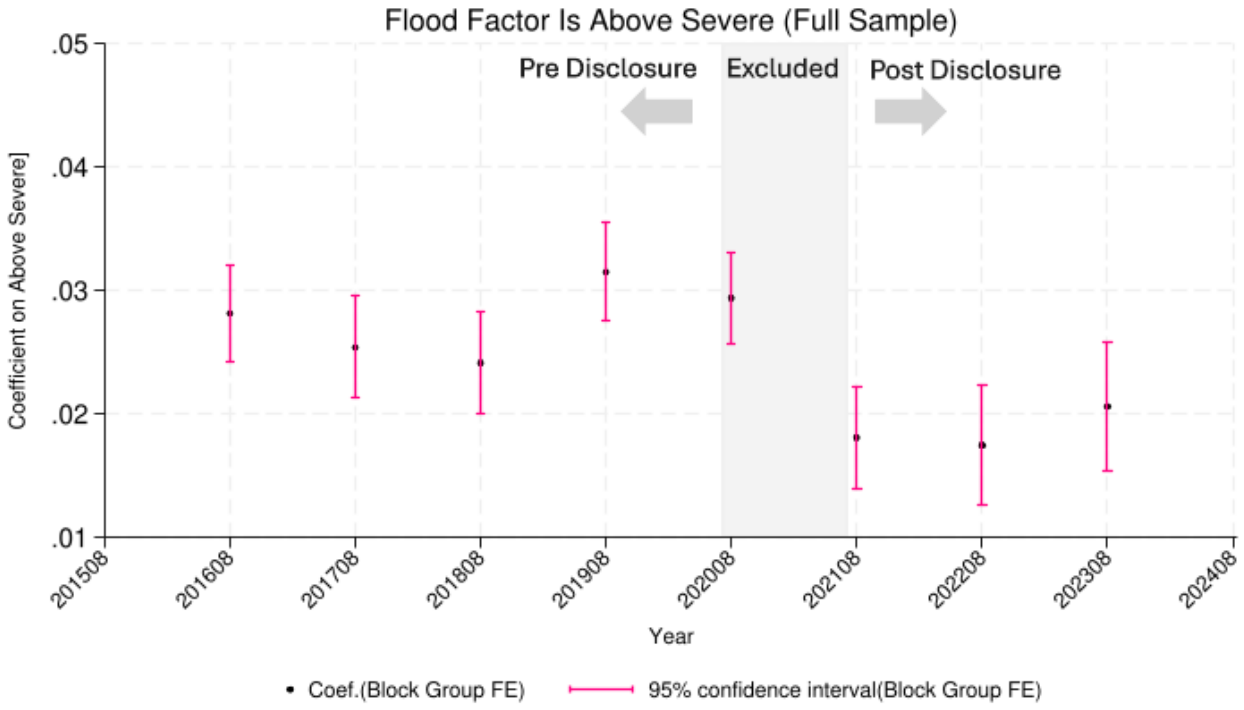
Specification	Total	Price (%)	Age Hetero. (%)	Dem. Hetero. (%)
Calibrated	-0.0066	91.9	10.7	-2.6
Uncalibrated	-0.0067	93.2	9.7	-2.9

Internet Appendix

A Part 1: Difference-in-Difference Analyses

Figure A1: Pretend Test for Log Home Prices

To test whether there are notable pre-trends in the prices of flood-prone properties prior to disclosure, each point on the graph represents the coefficient β_1 and associated standard errors from the following difference-in-difference regression: $\ln(\mathbf{SalePrice}_i) = \alpha + \beta_1 \mathbf{FloodCategory}_i + \beta_2 \mathbf{Treat}_i + \beta_3 \mathbf{HedonicControls}_i + \mathbf{BlockGroup\ FE} + \epsilon_i$. Each point corresponds to a sample that starts at the date shown on the x-axis and ends one year later. Note that January 2014-July 2014 and August 2024-December 2024 are included in our full sample but omitted from this pretend analysis, since they do not consist of a full year. Our data comes from First Street and Corelogic, and our data sample for this analysis covers the years 2013 to 2024.



B Part 2: Regression Discontinuity

B.1 Standard RD Analyses using the Post-Disclosure Period Only

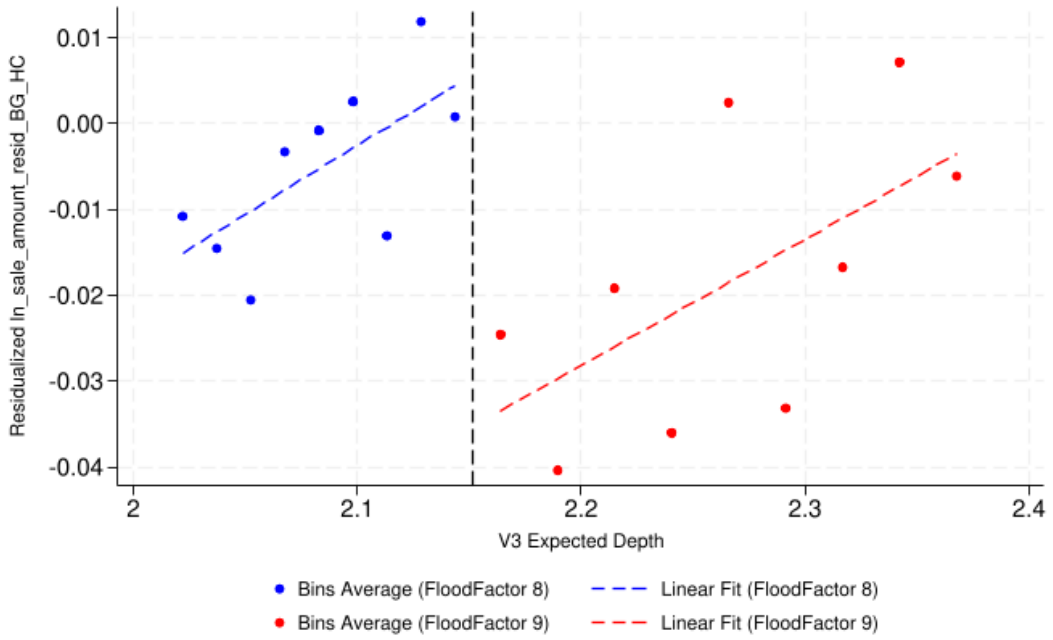
Table A1: Impact of Flood Risk Impact on Log Prices: Post-Period Only

This table presents the regression discontinuity results when we only examine around the cutoff in the post-disclosure period. The regression discontinuity specifications shown include a linear fit within an auto-selected bandwidth around the cutoff (column 1), fitted polynomials of order 2 (column 2) and fitted polynomials of order 3 (column 3). Flood risk data comes from First Street. Our time sample spans 2017 to 2024. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively. The post-disclosure period spans 2021 to 2024.

	Full Sample		
Dep. Var: Residualized LHS Variable	Linear(Auto Bandwidth)	2nd Order	3rd Order
	(1)	(2)	(3)
FloodFactor = 9	-0.0405***	-0.0292***	-0.0348***
	(-3.42)	(-3.81)	(-3.33)
Obs	32297	189458	189458
County FE	✓	✓	✓
Year-Month FE	✓	✓	✓
Hedonic Controls	✓	✓	✓

Figure A2: Plot of Log Price: Post-Period Only

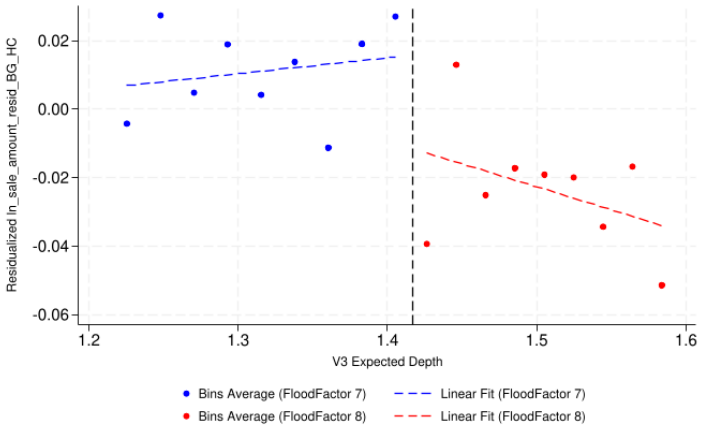
The figure below visually show a standard linear RD Analysis with auto-selected bandwidths using the full data sample for log price in the post-disclosure period only. To control for omitted variable bias, the LHS variable is residualized with block-group fixed effects and basic hedonics. Each dot represents the post-period average log home prices. The point on the x-axis represents the midpoint of each equal-width bin, and the bars represent the standard errors. Our data comes from First Street, Corelogic, and HMDA. The post-disclosure period spans 2021 to 2024.



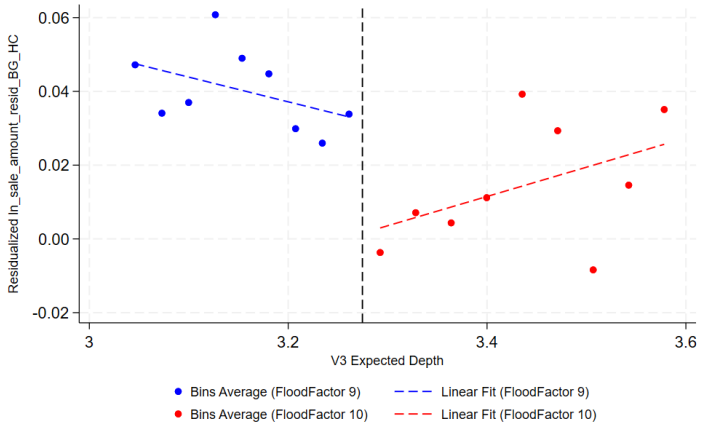
B.2 Analyses With Different Cutoffs

Figure A3: Analysis With Different Cutoffs

The figures below visually show a linear RD Analysis with auto-selected bandwidths using the different cutoffs



(a) RD for 7 vs 8 cutoff



(b) RD for 9 vs 10 cutoff

B.3 Analyses Across Different Subsamples

B.3.1 Alternative Sub-samples

Table A2: RD: Alternative Subsample Analysis of Flood Risk Impact on Log Sale Price

This table presents the regression discontinuity results, broken down by key subsamples, examining the relationship between the applicant's log income and expected flood depth after First Street. The key subsamples shown are: SFHA-Coastal sample (column 1), SFHA-NonCoastal properties (column 2), Non-SFHA Coastal properties that are exposed to coastal flood risk (column 3), Non-SFHA-Non Coastal properties that are *not* exposed to coastal flood risk (column 4). The regression discontinuity specification shown represents a linear fit within a auto-selected bandwidth around the cutoff between Flood Scores 8 and 9. Housing data comes from Corelogic, and demographic information comes from HMDA. Flood risk data comes from First Street. Our time sample spans 2016 to 2024. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively..

	Sub-Samples			
Dep. Var: Residualized Log Sale Price	SFHA-Coastal	SFHA-NonCoastal	NonSFHA-Coastal	NonSFHA-NonCoastal
	(1)	(2)	(3)	(4)
FloodFactor = 9 × Post Disclosure	-0.0662**	0.0378	-0.161***	0.000501
	(-1.98)	(0.72)	(-4.75)	(0.03)
Obs	14839	7951	11513	48724
Block Group FE	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓

B.3.2 Sub-sample Analyses for Other Variables

Table A3: RD: Subsample Analysis of Flood Risk Impact on Days on Market

This table presents the regression discontinuity results, broken down by key subsamples, examining the relationship between days on market and expected flood depth after First Street. The different subsamples include: Full sample (column 1), Coastal properties that are exposed to a coastal source of flooding (column 2), non Coastal properties (column 3), homes inside of a SFHA (ie 100 year floodplain) only (column 4), and Non SFHA properties (column 5). The regression discontinuity specifications shown use a linear fit within a auto-selected bandwidth around the cutoff between Flood Scores 8 and 9. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Housing data comes from Corelogic, and demographic information comes from HMDA. Flood risk data comes from First Street. Our time sample spans 2017 to 2024. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

	Sub-Samples				
Dep. Var: Residualized LHS Variable	Full	Coastal Only	NonCoastal	SFHA Only	NonSFHA
	(1)	(2)	(3)	(4)	(5)
FloodFactor = 9 × Post Disclosure	6.756**	13.07***	1.767	14.64***	3.037
	(2.36)	(2.84)	(0.49)	(2.67)	(0.91)
Obs	32994	13757	19237	10269	22725
Block Group FE	✓	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓	✓

Table A4: RD: Subsample Analysis of Flood Risk Impact on Log Income

This table presents the regression discontinuity results, broken down by key subsamples, examining the relationship between the applicant's log income and expected flood depth after First Street. The key subsamples shown are: Full sample (column 1), Coastal properties that are exposed to a coastal source of flooding (column 2), non Coastal properties (column 3), homes inside of a SFHA (ie 100 year floodplain) only (column 4), and Non SFHA properties (column 5).. The regression discontinuity specification shown represents a linear fit within a auto-selected bandwidth around the cutoff between Flood Scores 8 and 9. Housing data comes from Corelogic, and demographic information comes from HMDA. Flood risk data comes from First Street. Our time sample spans 2017 to 2024. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively..

	Sub-Samples				
Dep. Var: Residualized LHS Variable	Full	Coastal Only	NonCoastal	SFHA Only	NonSFHA
	(1)	(2)	(3)	(4)	(5)
FloodFactor = 9 × Post Disclosure	-0.0529**	-0.126***	-0.0233	-0.0491	-0.0528*
	(-2.07)	(-2.60)	(-0.78)	(-0.91)	(-1.82)
Obs	23600	6505	17095	5579	18021
Block Group FE	✓	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓	✓

Table A5: RD: Subsample Analysis of Flood Risk Impact on FHA Dummy

This table presents the regression discontinuity results examining the relationship between a dummy for whether an applicant is FHA and expected flood depth after First Street. The key subsamples shown are: Full sample (column 1), Coastal properties that are exposed to a coastal source of flooding (column 2), non Coastal properties (column 3), homes inside of a SFHA (ie 100 year floodplain) only (column 4), and Non SFHA properties (column 5). The regression discontinuity specification shown represents a linear fit within a auto-selected bandwidth around the cutoff between Flood Scores 8 and 9. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Housing data comes from Corelogic, and demographic information comes from HMDA. Flood risk data comes from First Street. Our time sample spans 2017 to 2024. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

	Sub-Samples				
Dep. Var: Residualized FHA dummy	Full	Coastal Only	NonCoastal	SFHA Only	NonSFHA
	(1)	(2)	(3)	(4)	(5)
FloodFactor = 9 × Post Disclosure	0.0180	0.0561**	-0.00116	0.0463*	0.00750
	(1.27)	(2.31)	(-0.07)	(1.77)	(0.45)
Obs	41883	12445	29438	10822	31061
Block Group FE	✓	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓	✓

Table A6: RD: Subsample Analysis of Flood Risk Impact on Average Age

This table presents the regression discontinuity results, broken down by key subsamples, examining the relationship between average household age and expected flood depth after First Street. The different subsamples include: Full sample (column 1), Coastal properties that are exposed to a coastal source of flooding (column 2), non Coastal properties (column 3), homes inside of a SFHA (ie 100 year floodplain) only (column 4), and Non SFHA properties (column 5). The regression discontinuity specifications shown use a linear fit within a auto-selected bandwidth around the cutoff between Flood Scores 8 and 9. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Housing data comes from Corelogic, and demographic information comes from HMDA. Flood risk data comes from First Street. Our time sample spans 2017 to 2024. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

	Sub-Samples				
Dep. Var: Residualized LHS Variable	Full	Coastal Only	NonCoastal	SFHA Only	NonSFHA
	(1)	(2)	(3)	(4)	(5)
FloodFactor = 9 × Post Disclosure	1.384**	1.348	1.401*	0.296	1.737**
	(2.21)	(1.20)	(1.86)	(0.23)	(2.42)
Obs	32381	9127	23254	7755	24626
Block Group FE	✓	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓	✓

Table A7: RD: Subsample Analysis of Flood Risk Impact on Dummy Democrat

This table presents the regression discontinuity results, broken down by key subsamples, examining the relationship between a dummy for whether a member of the household registered as democrat and expected flood depth after First Street. The different subsamples include: Full sample (column 1), Coastal properties that are exposed to a coastal source of flooding (column 2), non Coastal properties (column 3), homes inside of a SFHA (ie 100 year floodplain) only (column 4), and Non SFHA properties (column 5). The regression discontinuity specifications shown use a linear fit within a auto-selected bandwidth around the cutoff between Flood Scores 8 and 9. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Housing data comes from Corelogic, and demographic information comes from HMDA. Flood risk data comes from First Street. Our time sample spans 2017 to 2024. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

	Sub-Samples				
Dep. Var: Residualized LHS Variable	Full	Coastal Only	NonCoastal	SFHA Only	NonSFHA
	(1)	(2)	(3)	(4)	(5)
FloodFactor = 9 × Post Disclosure	-0.00444	0.0131	-0.0131	0.00141	-0.00777
	(-0.44)	(0.82)	(-1.03)	(0.08)	(-0.64)
Obs	83027	26352	56675	22790	60237
Block Group FE	✓	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓	✓

Table A8: RD: Subsample Analysis of Flood Risk Impact on Dummy for Non-White

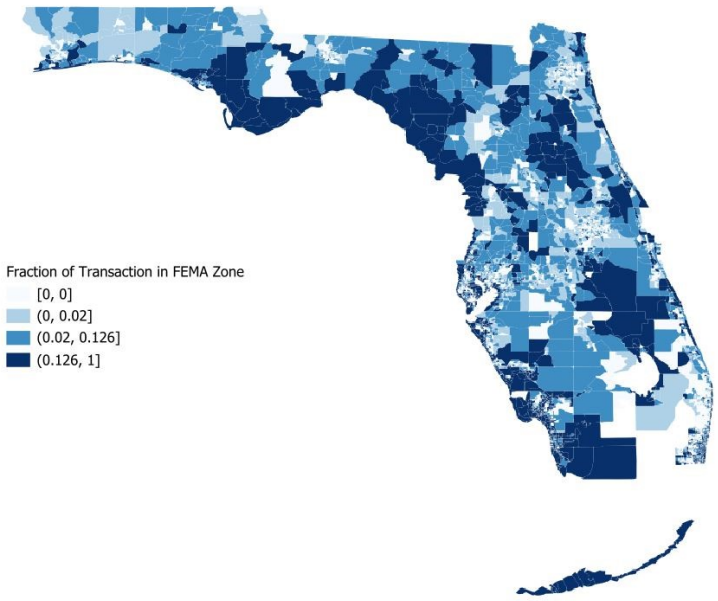
This table presents the regression discontinuity results, broken down by key subsamples, examining the relationship between a dummy for whether an applicant identifies as non-white and expected flood depth after First Street. The different subsamples include: Full sample (column 1), Coastal properties that are exposed to a coastal source of flooding (column 2), non Coastal properties (column 3), homes inside of a SFHA (ie 100 year floodplain) only (column 4), and Non SFHA properties (column 5). The regression discontinuity specifications shown use a linear fit within a auto-selected bandwidth around the cutoff between Flood Scores 8 and 9. Hedonic controls include building age bin, bedroom (1–6), bathroom (1–5), lot size (sqft), and living area (sqft). Housing data comes from Corelogic, and demographic information comes from HMDA. Flood risk data comes from First Street. Our time sample spans 2017 to 2024. Other coefficients not shown for brevity include: flood factor dummy, post disclosure dummy, and their interactions with expected flood depth. Significance levels are indicated by ***, **, and * for 1%, 5%, and 10%, respectively.

	Sub-Samples				
Dep. Var: Residualized LHS Variable	Full	Coastal Only	NonCoastal	SFHA Only	NonSFHA
	(1)	(2)	(3)	(4)	(5)
FloodFactor = 9 × Post Disclosure	0.00587	0.0234	-0.00399	0.00931	0.00440
	(0.60)	(1.46)	(-0.33)	(0.51)	(0.38)
Obs	81375	26321	55054	22530	58845
Block Group FE	✓	✓	✓	✓	✓
Year-Month FE	✓	✓	✓	✓	✓
Hedonic Controls	✓	✓	✓	✓	✓

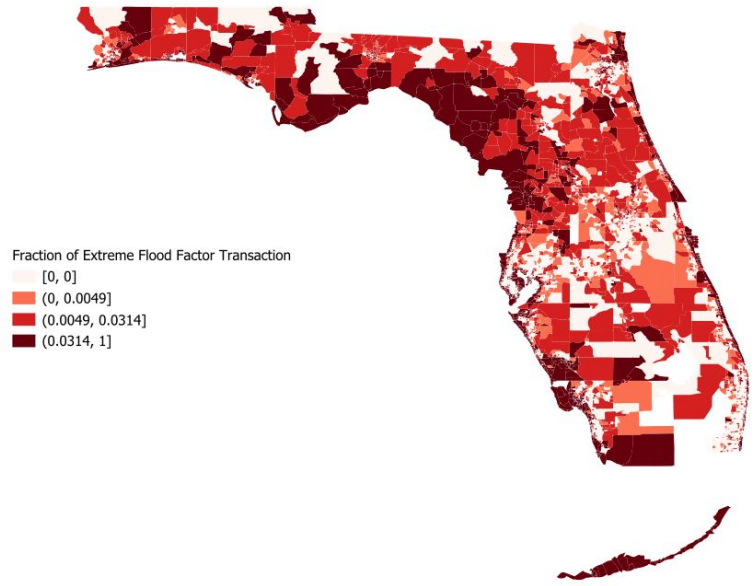
C Part 3: Geographic Maps

Figure A4: Census Tract Level Map of Flood-prone Properties in Florida

These maps display a census tract-level shapefile for Florida, where each polygon represents a tract. Panel a shows the fraction of transactions that are in a designated FEMA Flood Zone for each tract, while Panel b shows the fraction of transactions that are rated as having"extreme" flood risk by First Street. The shading reflects these fractions, with a darker shade representing a larger fraction. We limit the sample of house transactions to the pre-period (January 2017–August 2020) and the post-period (August 2021–December 2024). Our data comes from First Street and Corelogic, and our data sample covers the years 2017 to 2024.



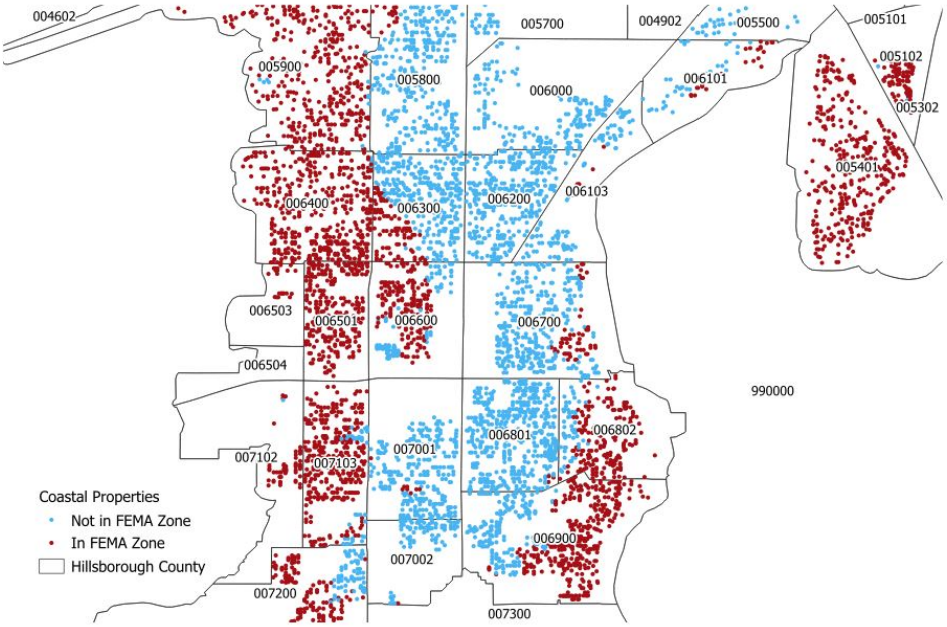
(a) FEMA Flood Zone



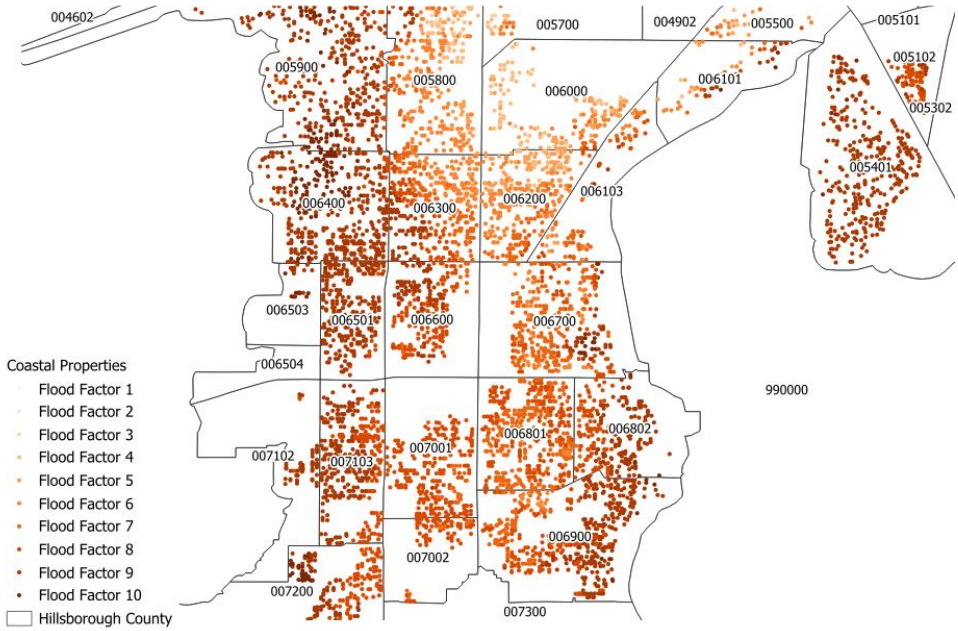
(b) Extreme Flood Risk (First Street)

Figure A5: Map of Properties Based on Flood Risk in Hillsborough County

These maps display census tract-level shapefile for a zoomed in selection from Hillsborough County, Florida. Panel a shows all properties that are (not) in a designated SHFA as a red (blue) dot, while Panel b shows the properties as dots colored based on the severity of their First Street Flood Factor. We limit the sample of house transactions to the pre-period (January 2016–August 2020) and the post-period (August 2021–December 2024). Our data comes from First Street and Corelogic, and our data sample covers the years 2016 to 2024.



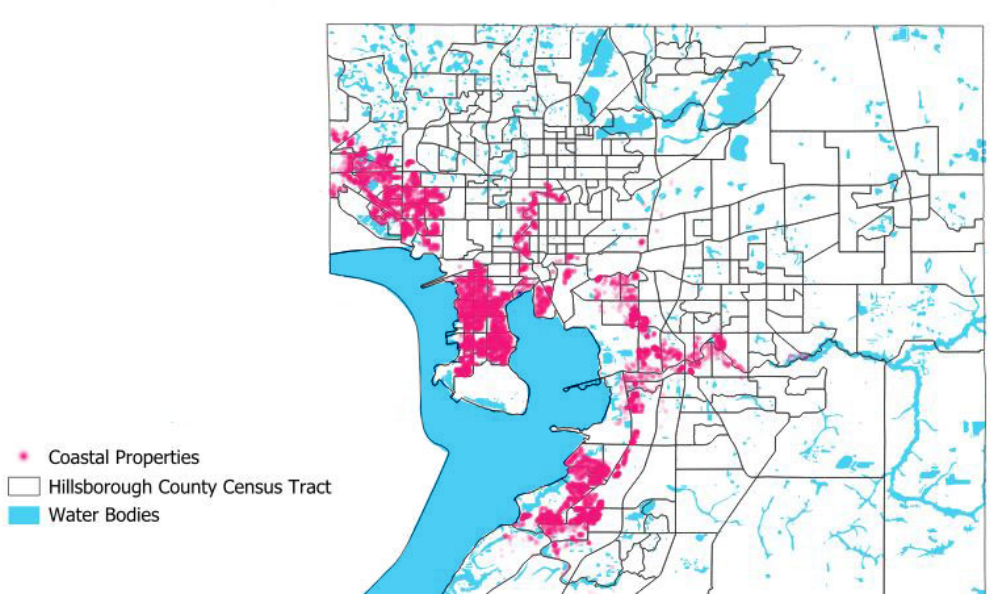
(a) FEMA Flood Zone



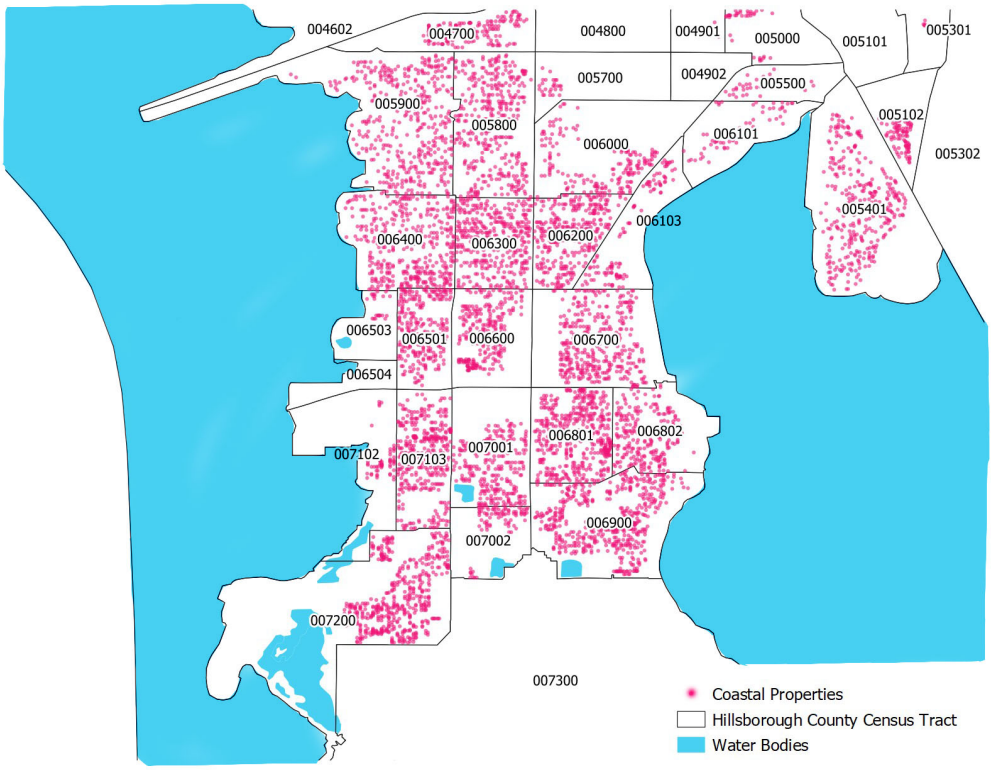
(b) Flood Risk (First Street)

Figure A6: Map of Coastal Properties and Bodies of Water for Hillsborough County

These maps display the location of individual coastal properties in Hillsborough County, Florida (panel a) and a select zoomed-in subsample of Hillsborough County, Florida (panel b). Each dot represents a property that is designated as "Coastal" by First Street, and we also overlay local bodies of water in blue. We limit the sample of house transactions to the pre-period (January 2016–August 2020) and the post-period (August 2021–December 2024). Our data comes from First Street and Corelogic, and our data sample covers the years 2016 to 2024.



(a) Hillsborough County, Florida



(b) Select Subsample of Hillsborough County