Flood risk and Insurance Take-up in the Flood Zone and its Periphery

Ivan Petkov *
Northeastern University

Francesc Ortega †
CUNY, Queens College

October 9, 2023

Abstract

Many studies have investigated flood risk and insurance coverage in the 100-year flood zone, but much less is known about the periphery of the flood zone. We present a new approach to estimate flood risk and insurance take-up in the vicinity of the flood zone based on building-level inundation data. We illustrate our approach using data for New York after hurricane Sandy. We show that flood risk falls rapidly as we move away from the flood zone, but remains fairly high for properties located within 250 meters of the flood zone. We also document substantial voluntary insurance take-up in this area prior to the storm, reflecting homeowners' perception of flood risk. Next, we show that experiencing flooding during Sandy led to large increases in flood insurance coverage in the flood zone and its periphery. But, while in the flood zone the increase vanished after 3 years, it was highly persistent in the periphery. By using information on the types of insurance policies purchased by homeowners, we provide evidence that strongly suggests that periphery residents who experienced flooding revised upwardly their beliefs about flood risk and adapted by purchasing (affordable) flood insurance.

JEL Classification: R11, R30, Q54, G22

Keywords: Flood risk, flood insurance, FEMA, NFIP, hurricane Sandy

*Department of Economics, E-mail: i.petkov@neu.edu and homepage (`https://petkov33.github.io`).
†Department of Economics, E-mail: francesc.ortega@qc.cuny.edu and homepage: (`http://qcpages.qc.cuny.edu/~fortega`).

1 Introduction

The 100-year flood zone (also known as the *Special Flood Hazard Area* or SFHA) is a cornerstone of the National Flood Insurance Program (NFIP) and, more generally, public policies aimed at mitigating flood damage (Highfield et al., 2013; Patterson and Doyle, 2009).[1] Not surprisingly, it has received a lot of attention by researchers across various disciplines. However, as evidenced by some recent hurricanes, many properties located in the vicinity of the 100-year flood zone are also at risk of flooding (Kousky et al., 2020b; Bates et al., 2021; Billings et al., 2022). Residents of these areas are not required to purchase flood insurance and may not be sufficiently informed of their exposure to flood risk. As a result, flood insurance coverage in these areas is believed to be low (Kousky et al., 2020a). To make matters worse, in many parts of the country, the periphery of many flood zones is growing fast (Patterson and Doyle, 2009; Galloway et al., 2006; Indaco and Ortega, 2023).

This paper presents a novel approach to estimate flood risk outside of the (100-year) flood zone on the basis of building-level inundation data following a major storm, distance to the flood zone and elevation. We illustrate our method using data for New York in the aftermath of hurricane Sandy, which hit the city in October of 2012. Our approach identifies the geographic scope of the periphery of the flood zone with significant flood risk in a transparent manner and relying exclusively on publicly available data. We also use data on the number and type of NFIP flood insurance policies to estimate the effects of the storm on insurance take-up, its persistence, and the role played by the experience of flooding in the decision to purchase flood insurance. As recognized by Bradt et al. (2021), because flood insurance is largely voluntary outside of the 100-year flood zone, the responses of residents of areas *outside* the flood zone with significant *exposure* to flood risk are highly informative regarding how these residents update their flood risk beliefs and undertake adaptation decisions.

Our investigation yields three key findings. First, we estimate the flood risk gradient outside of the flood zone on the basis of the flooding caused by the storm. The estimates reveal substantial flood risk within 250 meters from the edge of the 100-year flood zone. In this area, about 30% of all houses flooded, which is approximately 1/3 of the flooding rate in the flood zone. Our estimates also show that the flooding rate fell rapidly with distance from the flood zone and as a function of ground elevation. Second, prior to hurricane Sandy, take-up rates were fairly high in the 100-year flood zone: we estimate that homeowners in the average census tract with a majority of houses in the flood zone purchased around 78 policies annually for each 100 houses.[2] In comparison, take-up rates were about one third

[1] The 100-year flood zone, also known as the Special Flood Hazard Area (SFHA) is the land that, according to FEMA's flood maps, has a 1% (or higher) annual chance of flooding.

[2] Keep in mind that some houses contain multiple units but often insurance policies are contracted by

in the immediate vicinity (about 26 policies for each 100 houses), and became negligible beyond 250 meters from the flood zone. We also estimated that hurricane Sandy led to large increases in flood insurance take-up, both in the flood zone and its vicinity. But, while in the flood zone the increase vanished after 3 years (consistent with Gallagher (2014)), insurance take-up remained persistently high (for a decade or more) in the periphery.

We also show that the *types* of insurance policies purchased after Sandy in the two areas differed in eligibility and cost, providing valuable information in regards to homeowners' motivation for purchasing insurance. The temporary flare-up in insurance take-up in the flood zone was a mechanical response to the insurance requirements imposed by FEMA in order to obtain relief aid, confirming the findings in Kousky (2017). However, the largely permanent increase in insurance take-up in the *vicinity* of the flood zone is more consistent with an upward revision of flood risk beliefs by homeowners that experienced severe flooding, possibly for the first time, during Sandy. As a result, these homeowners appear to have adapted to the new information by purchasing flood insurance and maintaining coverage for an extended period of time.[3] In contrast, the flooding caused by the storm did not affect the flood risk beliefs of homeowners living within the flood zone. They might have already been aware of their exposure to flood risk or, even if they did revise their beliefs upwardly, they may not have been able to afford the insurance policies available to them and discontinued these policies after the (3-year) insurance mandate lapsed.[4]

Our study is related to several strands of literature. Our paper is closely related to the literature on flood insurance take-up after large-scale flooding events. In a very influential study, Gallagher (2014) finds large but short-lived increases in insurance policy purchases across flooded and nearby areas. The study concludes that both the direct experience of flooding and its saliency in the media lead to revisions of flood risk beliefs. However, beliefs appear to revert back to pre-event levels fairly quickly. Subsequently, Kousky (2017) also documented sharp increases in flood insurance take-up following major storms (by 7% to 14% one year later) that were also short-lived, dying out three years after the storm. She argues that most of the (temporary) increase in policy purchases is due to the FEMA 3-year minimum insurance mandate imposed on disaster aid recipients. More recently, Hu (2022)

individual households. It is not unusual that several households living in a 3-family house purchase more insurance policies.

[3]Importantly, flood insurance policies were available to homeowners residing outside of the 100-year flood zone at low cost. Since the introduction of the new insurance pricing system (known as *Risk Rating 2.0*), these policies are no longer widely available.

[4]It is worth noting that our approach can easily be extended to incorporate multiple flooding events, assuming that accurate data for the corresponding inundation zones is available. Presumably, the owners of properties experiencing flooding for the first time (or higher levels of flooding than in the past), which will typically be located outside of the flood zone, will update their flood risk beliefs and react to the new information.

has shown that the decision to purchase flood insurance is also influenced by the experience of flooding by one's far-away friends.

Many studies aim at quantifying flood risk outside of the 100-year flood zone. Often, researchers rely on comparisons between properties inside the flood zone and *all* properties outside of it, including vast numbers for which flood risk is virtually zero. As a result, very little is learned about flood risk in the vicinity of the flood zone. Recognizing this problem, some studies restrict attention to either insured or damaged properties, but this approach raises concerns of potential bias due to adverse selection (Wagner, 2022). Additionally, studies that rely on insurance claims submitted following flooding events do not provide a clean measure of flood risk because they also reflect insurance take-up and, thus, are not helpful in areas with non-existing or low flood insurance penetration (Galloway et al., 2006).

Our work is also related to the literature examining the informational content of flood maps. Kousky et al. (2020a) and (Gourevitch et al., 2023) argue that one factor contributing to explaining the low insurance take-up in flood-prone areas adjacent to the 100-year flood zone is that homeowners wrongly believe they are safe from flooding because their house is not part of the local flood map. Along similar lines, Weill (2022) examines FEMA flood map updates and shows that the adoption of new flood maps affects insurance take-up and argues that this is partly a response to the information regarding flood risk conveyed by the map updates.

A better understanding of flood risk in the periphery of the flood zone is also relevant for an emerging literature on mortgage finance (Kousky et al., 2020b; Sastry, 2023). In their analysis of the effects of hurricane Harvey, Billings et al. (2022) found that flooding led to large increases in bankruptcy and loan delinquency rates among homeowners with low credit scores located *outside* of the flood zone, but found no such effects within it. Similarly, Ouazad and Kahn (2022) analyze coastal risk in the mortgage market and document a growing volume of mortgages and securitizations in flood-prone areas lying *outside* of the 100-year flood zone. Our work is also related to studies analyzing the economic effects of hurricane Sandy. The majority of these studies have focused on the housing market (Ortega and Taspinar, 2018; McCoy and Zhao, 2018; Gibson and Mullins, 2020; Cohen et al., 2021), but a few have also examined the effects of the storm on business establishments (Indaco et al., 2021; Meltzer et al., 2021).

Relative to the existing literature, our paper makes four contributions. First, our analysis considers *all* buildings located in the flood zone and its surrounding areas, as opposed to focusing only on a pre-selected subset (such as those submitting damage claims after a storm). Second, our risk estimates are based on *actual* flooding data based on a large-scale flooding event. The resulting risk gradient is helpful to identify the geographic scope of the periphery of the flood zone with non-negligible flood risk, which probably varies widely across cities located in flood-prone areas. Such information is important for homeowners, policymakers and financial institutions holding loans collateralized with uninsured properties exposed to flood risk.[5] Third, we estimate the persistence of the increase in insurance take-up following a large storm, allowing for this persistence to differ in the flood zone and its periphery. Last, our exploration of the *types* of insurance policies purchased after the storm sheds light on the factors that determine homeowners' decisions to purchase (and maintain) flood insurance.

Our paper is not the first to focus on the flood risk and insurance take-up at the edge of the 100-year flood zone. A number of studies employ large-scale models to identify the flood risk gradient in the flood zone and its surrounding areas (e.g. Wing et al. (2018)). In highly cited study, (Gourevitch et al., 2023) quantify the extent of overvaluation in flood-prone housing markets and highlight that this is largely driven by properties located outside of the flood zone. An unappealing feature of large-scale models is their lack of transparency and the proprietary nature of their estimates (e.g. Flood Risk Foundation). In contrast, our modeling approach is simple, easily reproducible, and delivers intuitive estimates of the risk gradient and insurance demand in the vicinity of the flood zone as a function of distance to the flood zone and ground elevation.

We view our paper is highly complementary to Bradt et al. (2021). These authors analyze the determinants of flood insurance take-up rates at the census-tract level using panel data covering the whole United States. They report two main finding. First, the price-elasticity of the demand for insurance if fairly low (around 0.3) both in the SFHA and outside of it. They also provide evidence of intense adverse selection outside of the SFHA, where properties facing slightly higher flood risk are dramatically more likely to purchase insurance. The latter finding relies on proprietary flood risk measures provided by the *First Street Foundation*. In comparison, our approach relies solely on publicly available data and our semi-parametric specifications are easy to grasp. By focusing on a single city, we are able to accurately estimate the geographic scope of the periphery of the flood zone where flood risk is substantial. As a result, our estimates of insurance take-up outside the flood zone following a flooding event (and their persistence) are more precise than what one can achieve by estimating standard panel data models. In addition, we show how information on the *types* of policies purchased by homeowners is helpful to disentangle voluntary insurance take-up from insurance purchases required by FEMA in order to receive relief grants.

The structure of the paper is as follows. section 2 presents our data sources. section 3

[5] An alternative approach is to consider the 500-year flood zone as the relevant flood-prone periphery. However, our analysis of hurricane Sandy's inundation area makes clear that flooding extended well beyond the 500-year flood zone. Restricting to the latter would greatly underestimate the geographic scope of the flood-prone area outside of the 100-year flood zone.

presents summary statistics. section 4 estimates our empirical model to estimate the flood risk gradient. section 5 analyzes insurance take-up prior to Sandy and the effects of the storm on take-up. section 6 discusses the role of informational signals on insurance take-up. section 7 concludes. The Appendix contains additional tables.

2 Data sources and periphery bands

This section describes our data sources and our definition of the periphery of the flood zone.

2.1 Data sources

Footprints. We obtain shape files for building footprints from the *Microsoft building footprints* project. It combines *Bing Maps* imagery with deep neural networks to produce high-quality building footprints worldwide.[6] The data for the United States contains nearly 130 million footprints, but we restrict our analysis to New York City (NYC).

Flood maps. Our definition of the flood zone is primarily based on the *Special Flood Hazard Area* (SFHA), also known as the 100-year flood zone, at the time hurricane Sandy arrived in New York (October 29, 2012).[7] Importantly, properties within the SFHA with federally backed mortgages are required to maintain flood insurance, although this mandate is not always enforced.

Sandy's inundation zone. Our flood risk analysis in section 4 relies on building-level flooding data at the time of hurricane Sandy. This FEMA data product characterizes the geographic boundary of the area that flooded during Sandy and provides estimates of the level of flooding at each building.

Elevation. We also make use of building-level elevation data, which is an important determinant of flooding. Our elevation data is a geo-referenced raster (assembled by the *U.S. Geological Survey*) derived from the *National Elevation Dataset* released in October, 2012. The raster reports ground elevation (in meters) and has a resolution of 100 meters.[8]

Flood insurance take-up. Our flood insurance take-up analysis relies on the universe of individual policies issued by the *National Flood Insurance Program* (NFIP) for the period 2009-2022. The dataset contains information on individual policies (such as premiums and

[6] For further details, visit `https://www.microsoft.com/en-us/maps/building-footprints`.

[7] The FEMA flood maps (FIRMS) that were effective at the time Sandy reached New York had been adopted in 1983 and slightly revised in 2007. According to FEMA, buildings located in the SFHA have an annual flooding probability of 1 percent or higher. Specifically, there were approximately 21,000 properties in the SFHA, belonging to classes A, AE or VE (Dixon et al. (2017)). The flood maps also identify properties with a 0.2 percent annual flood probability, collectively known as the 500-year flood zone.

[8] Data retrieved from `https://earthworks.stanford.edu/catalog/stanford-zz186ss2071`.

policy type). However, it lacks any policy-specific geographical identifier (such as address) except for the building's census tract. Accordingly, we shall aggregate individual policies to the census tract by year level. Following Bradt et al. (2021), we discard policies with a negative premium, those with coverage for contents only (as opposed to building) and any policy with discounts above the CRS limit for the type of occupancy (though these adjustments have no effect on our findings). We exclude year 2022 from our insurance take-up analysis because of the introduction of the new pricing system known as *Risk Rating 2.0*, which we discuss in section 7.

2.2 Periphery bands

The literature has almost exclusively focused on the SFHA (100-year flood zone). However, we suspect that there are unexplored interesting dynamics in the periphery of the flood zone. First, homeowners residing outside of the SFHA, but not far from it, may have assumed that their properties were not at risk of flooding. Unexpectedly experiencing flooding for the first time may change their beliefs about flood risk and induce them to purchase flood insurance and perhaps adopt other adaptation measures. It is worth emphasizing that homeowners outside of the SFHA are not subject to the flood insurance mandate. Hence, the voluntary nature of purchasing a flood insurance policy in the periphery speaks more directly to the updating of flood risk beliefs. Lastly, the experience of hurricane Harvey in 2017 in the area around Houston showed that flooding can be extensive even outside of the SFHA. Hence, investigating insurance policy take-up outside of the SFHA is important from a public policy perspective.

To explore flood insurance dynamics in the periphery of the SFHA, we constructed a series of concentric bands around it. More specifically, let b be a building in the city, with latitude-longitude coordinates x_b, and let B denote the set of all buildings in the city that lie either in SFHA or its periphery P, defined as the set of buildings within 1,000 meters of the SFHA.[9] We then partition periphery set P into 4 non-overlapping bands defined in terms of the distance between each building and the (nearest building within the) SFHA, that is. Accordingly, $band1$ contains all buildings within 250 meters of the SFHA and $band2$ contains buildings within 250 and 500 meters of the SFHA. Similarly, $band3$ and $band4$ contain buildings within 500-750 meters and 750-1,000 meters from the SFHA, respectively.

Figure 1 depicts New York City's SFHA and the 4 surrounding periphery bands. Buildings colored in blue are located in the SFHA and, not surprisingly, are located very close to the ocean or other water bodies. Buildings colored in orange are buildings located in pe-

[9]Casual examination of the surge data shows that the relevant periphery falls within this area.

riphery band 1 and tend to be further away from water bodies but, because of differences in elevation, some band-1 buildings are as close (or closer) to the coastline than some buildings within the SFHA. Similarly, the figure also depicts bands 2 (in light green), and 3 and 4 (in grey).

3 Summary statistics

Our focus in this study is on New York City's SFHA and the buildings located within 1,000 meters of it, which we refer to as the *periphery* of the flood zone. Together, these areas contain slightly over 175,000 buildings (footprints), 84% of which are residential 1-family to 4-family buildings.

Table 1 describes the key variables that will be used in our analysis. The top panel refers to all buildings: 8.8% are located in the SFHA (100-year flood zone) and 7.6% in the 500-year flood zone. By construction, the 4 periphery bands have similar counts of properties and, as a result, the shares of buildings in each of those bands is similar (ranging from 19% to 26% of the total).[10] The average elevation of the properties in our data is 13 meters above sea level (with a maximum of 101 meters). The last two variables in the table refer to the impact of hurricane Sandy. Specifically, 20% of all buildings (in our area of interest) flooded during Sandy and the average inundation depth was 0.6 feet (with a maximum of 15.1 feet). The bottom panel of the table focuses on houses (i.e. excluding apartment buildings), defined as 1-family to 4-family residential buildings. The share of houses in the SFHA (8.6%) and the share of those that flooded during Sandy (20.9%) are similar to those corresponding to all buildings in our area of interest.[11]

To gain familiarity with the periphery bands, it is helpful to examine the distance to the flood zone for the average building in each band (which will be increasing by construction), the average elevation, building type (share of houses) and flooding rates and levels at the time of hurricane Sandy.

Table 2 collects the results for the city as a whole (top panel) and also separately for each of the 5 boroughs (MN Manhattan, BK Brooklyn, QN Queens, SI Staten Island, and BX the Bronx). New York City's SFHA contains approximately 15,500 buildings whereas our periphery bands range from 34,000 to 45,000 buildings each. In all cases the share of

[10] The SFHA (100-year flood zone) and the periphery bands are mutually exclusive. Instead, the 500-year flood zone overlaps with the periphery bands.

[11] Though not reported in Table 1, it is worth noting that Manhattan has many fewer buildings (but much taller) in our area of interest (about 5,000) and only 7 percent are (1-family to 4-family) houses (about 350). We also note that the borough with the highest flooding rate during Sandy (for our area of interest) was Brooklyn (43%), followed by Queens (27%).

1-family to 4-family buildings is over 80%. The average distance from buildings located in the periphery to the (nearest building in the) SFHA is 124 meters for band 1, 373 meters for band 2, 622 meters for band 3, and 872 meters for band 4. The table also shows that, as we move away from the flood zone, average elevation increases gradually, which provides additional protection from flooding. Specifically, while average ground elevation is slightly less than 3 meters in the SFHA, it rises to 7.5 meters in band 1 and reaches almost 20 meters in band 4.[12]

The next section will examine in detail the effects of hurricane Sandy on the buildings located in our area of interest (in terms of flooding), but columns 5 and 6 in Table 2 anticipate some summary statistics. In the SFHA, 82% of properties flooded during Sandy and the average depth of flooding was 3.7 feet. Naturally, as we move away from the flood zone, the flooding rates and average depth both fall. It is also worth noting that flooding during Sandy was also widespread in band 1, where 30% of the properties flooded by an average of 0.75 feet.

The table also shows that there is variation across boroughs along several dimensions. First, in all boroughs the vast majority of buildings in our area of interest are 1-family to 4-family homes, with the exception of Manhattan where this building class accounts for fewer than 10% of all buildings. Average elevation in the SFHA is similar across the city's boroughs. However, average elevation in the periphery is much lower in Brooklyn than in the other boroughs, which explains why this borough experienced the highest flooding rates. In fact, 90% (or more) of the buildings in the flood zones of Brooklyn, Queens and Staten Island flooded during Sandy. In comparison, the flooding rates in the SFHA of Manhattan and the Bronx were much smaller (60% and 37%, respectively).

4 Flooding during Sandy

The widespread flooding caused by hurricane Sandy could have led to a surge in the demand for flood insurance in NYC by providing homeowners with building-specific signals of flood risk. Before turning to flood insurance take-up data, this section examines the spatial impact of Sandy in terms of flooding in our area of interest. In particular, the goal is to estimate the flood risk gradient using our periphery bands, which will describe the extent of flooding outside of the 100-year flood zone and how flooding rates decayed as a function of distance (and elevation) to the boundary of the flood zone.

[12]In addition, the 500-year flood zone overlaps with our definition of periphery, but only partially. Almost 1 in 4 buildings in band 1 belong to the 500-year flood zone (23%) and only 31% to the combined periphery bands.

4.1 Flood risk gradient and model specification

This section relies heavily on our building-level data describing Sandy's inundation zone. These data identify which buildings suffered from flooding (extensive margin) and with what depth (intensive margin). Our approach is to use the rings around $FZ100$ to compute flooding rates (and depth), which we expect to fall as we move away from the flood zone (and with elevation). Specifically, we shall estimate a cross-sectional model where the dependent variable is an indicator taking a value of one when the building flooded during Sandy:

$$Flooded_i = \beta_0 + \beta_1 Band1_i + ... + \beta_4 Band4_i + \gamma X_i + \varepsilon_i \tag{1}$$

The main explanatory variables are indicators for whether property i is located in Band1, Band2, and so on, and the excluded category is FZ100. In addition, we also include a polynomial in elevation and borough fixed-effects. We will also consider an analogous model where the dependent variable is continuous and measures the depth of flooding (in feet).

It is helpful to focus the analysis to the sample of houses (1-family to 4-family classes), excluding apartment buildings (and houses with more than 4 units).[13] Houses account for the majority in our area of interest (84% of all units) and are more vulnerable to flooding than apartment buildings.

4.2 Estimation Results

4.2.1 Extensive margin: the probability of flooding

The estimates for the extensive margin of flooding are reported in Table 3. The intercept in Column 1 shows that 87% of houses in FZ100 flooded during Sandy (compared to 21% in our overall area of interest). The flooding rates fell monotonically as we move to outer rings. We find that buildings in bands 1 and 2 suffered substantial flooding, even though they were located outside the SFHA: 32% and 11% of the houses in these bands flooded, respectively. Flooding also reached bands 3 and 4, but flooding rates fell to single digits (8% and 4%, respectively).

Column 2 includes borough fixed-effects, which allow for variation in flooding rates within the SFHA across boroughs. The estimates indicate that Brooklyn's flood zone suffered the highest flooding rate: 22 percentage-points higher than in the Queens's flood zone. As in the previous column, we also find that the probability of flooding decreased monotonically as we move away from the flood zone. In fact, flooding rates were 54 percentage-points

[13] Because Manhattan includes very few of these in our area of interest (only 191 houses), we drop it from the main estimation sample in this section.

lower in band 1 (than in the corresponding borough's flood zone), falling by an additional 20 percentage-points as we step into band 2, and only by an additional 3 percentage points each time as we move into periphery bands 3 and 4. Hence, the flood risk gradient points to a decreasing, and approximately *convex*, function of distance.

Column 3 includes a polynomial (of order 2) in each building's ground elevation. The results indicate that both elevation and SFHA distance are key determinants of flood risk. The probability of flooding falls monotonically with building elevation (up to 35 meters). In addition, controlling for elevation, modifies the relationship between distance from the FZ100 and flood risk. As we move from the SFHA into band 1, the probability of flooding falls by 46 percentage points. In turn, stepping into band 2 lowers the probability by an additional 12 percentage points, but we do not observe additional reductions in the flood probability from moving into bands 3 and 4.

Column 4 considers only the houses outside of the SFHA and imposes a quadratic function of distance from the flood zone (in 100m increments), as well as a quadratic function of elevation. At the edge of the flood zone the probability of flooding is estimated to be 57% (compared to 87% in the whole of the flood zone), confirming the prediction in (Galloway et al., 2006). The estimates also imply a convex relationship between flood risk and distance from the boundary, confirming the decreasing reductions in flooding rates found in the previous columns as we move from inner to outer bands.

Columns 5-8 are analogous to column 1, but estimated separately by borough, which allows the gradient of flooding probabilities to vary by borough. The estimates show flooding rates upward of 90% in the flood zones of Brooklyn, Queens and Staten Island, falling gradually as we move toward the outer periphery bands. Periphery band 1 exhibited high rates of flooding in these three boroughs, but the boundary of Sandy's surge differed across boroughs.[14] Only in the Bronx do we find that flooding was limited to the SFHA (with a 32% flooding rate). In the rest of the city, Sandy caused substantial flooding beyond the 100-year flood zone.

As a robustness check, we also estimate the model for the probability of flooding for all building types (Table A1), including apartment buildings (and the borough of Manhattan). The estimates are very similar to what we just discussed. Column 5 also shows that Manhattan experienced widespread flooding in the SFHA (with a flooding rate of 60%). Similar to what we found in Staten Island, flooding in Manhattan's periphery band 1 was also substantial (13%), but practically non-existing in the rest of the periphery.

In sum, our estimates for the city as a whole clearly show that Sandy impacted the

[14] The estimates show that flooding was substantial in Brooklyn's band 4 (around 22%) but practically non-existing in the same band of the other boroughs.

periphery of the SFHA, flooding a substantial share of buildings in that location. Our estimates also make clear that the probability (extent) of flooding fell in a roughly monotonic fashion as we move away from the flood zone and as a function of ground elevation.

4.2.2 Intensive margin: flooding levels

Let us now shift to our continuous measure of flooding depth. The estimates are reported in Table 4. Column 1 shows an average flooding of 3.9 feet during Sandy among houses in the SFHA. As was the case with the probability of flooding, flooding levels fell monotonically as we move to outer rings, but at a decreasing rate. The average flooding among buildings in band 1 was 0.8 feet (or 3.1 feet lower than in the flood zone). In turn, the estimated flooding levels were 0.25, 0.14 and 0.09 feet in bands 2, 3 and 4, respectively. It is important to keep in mind that buildings outside of the SFHA are typically more vulnerable to flood damage because they were not subject to building code regulations at the time of their construction.

Column 2 shows that flooding levels differed across the SFHA of the different boroughs. Not surprisingly, the highest depth was attained in Brooklyn's flood zone, where flooding levels were almost 0.5 feet higher than in Queens' flood zone. Column 3 introduces ground elevation as a quadratic polynomial, again showing that a building's elevation lowers flooding depth for a given distance from the flood zone. As we did before, column 4 focuses on the sample of buildings outside the SFHA and considers a polynomial for distance to the flood zone and elevation. The estimates show that buildings at the edge of the SFHA experienced 1.4 feet of flooding, almost half of the average surge in the flood zone, which probably led to substantial damage given that those buildings were not subject to minimum elevation requirements at the time of construction. Columns 5-8 present estimates by borough, which again depict the same patterns found in Table 3: substantial extent and intensity of flooding across most of the flood-zone periphery in Brooklyn and Queens, and to a lesser extent in Staten Island and the Bronx.

Overall, we observe a very similar pattern to what emerged from the analysis of the extensive margin of flooding.[15] In other words, both the risk and the intensity of flooding were highest in the SFHA, falling rapidly as we moved away. However, our estimates make clear that substantial flooding occurred in periphery band 1 and, to a lesser extent, band 2. In band 1, about 30% of all (1-family to 4-family) houses flooded with an average depth of 0.8 feet.

As Sandy was approaching the city, homeowners in New York's SFHA probably expected

[15] This was not unexpected given that we included all buildings in the estimation sample, rather than conditioning on having experienced flooding. Thus, non-flooded buildings are included with zero flooding levels.

flooding in their homes. In contrast, homeowners residing outside the flood zone, but close to it, were probably less aware of the risk of flooding. Thus, the substantial flooding experienced in the periphery of the flood zone in the aftermath of Sandy probably led to an important revision of flood risk beliefs among homeowners affected by flooding or that witnessed flooding in their neighborhood.

It is worth emphasizing that homeowners outside of the SFHA were much less prepared for flooding than residents of the flood zone. First of all, their houses were probably less resilient to flooding because they had not been subject to elevation requirements. And, additionally, probably very few of these houses were insured against flooding. The next section begins by describing flood-insurance take-up in NYC's flood zone and its periphery. As we shall see, insurance take-up rates were indeed very low in the periphery of the flood zone. We will then investigate whether experiencing flooding during Sandy led to increases in insurance take-up.

5 Flood insurance take-up before and after Sandy

The previous section documented that Sandy flooded the vast majority of building's in New York's SFHA. But, in addition, it also produced extensive flooding in the periphery of the 100-year flood zone.

This section has two main goals. First, we aim to estimate the flood-insurance take-up within and outside the SFHA. Importantly, many residents in the flood zone are subject to insurance mandates, while those located outside of the flood zone are largely free to purchase insurance or not. In particular, finding low take-up rates outside of the flood zone would suggest that residents were unaware that their homes were subject to flood risk (Bradt et al. (2021)).

Secondly, we also examine whether the experience of flooding led to increases in insurance take-up and whether these increases were temporary or permanent. This analysis is closely related to Gallagher (2014), but the novelty of our approach is the strong focus on insurance take-up outside of the SFHA.

5.1 Data and challenges

Empirical analyses of flood insurance take-up rates face an important challenge, which also applies to us. Our data on insurance policy purchases does not identify the exact location (or the address) of the house being insured. As a result, we cannot conduct a building-level analysis as in the previous section. Instead, we need to conduct the analysis at the census tract level (2010 definition) and overlay our periphery bands over census tracts.

One implication of this limitation is that estimating insurance coverage rates in the SFHA, or outside of it, is problematic. Previous studies have provided some estimates (Dixon et al. (2017), Bradt et al. (2021)) but considerable uncertainty remains, particularly in regards to coverage in the periphery of the SFHA. Another implication of this shift is that many census tracts are not completely contained within our area of interest (the SFHA and the 1km ring around it). To avoid chopping these census tracts (and generating censoring in our measures of insurance take-up), we have expanded our area of interest to include all buildings in those census tracts. Accordingly, we created an additional (fifth) periphery band, which contains buildings located further than 1km away from the SFHA that belong to census tracts that overlap with our original area of interest.

Table 5 presents summary statistics. The expanded area of interest contains 729 census tracts, which are observed over a 13-year period (from 2009 through 2021), giving rise to 9,477 census tract per year observations. The average census tract in our data contains 303 buildings (but the number ranges from 7 to 2,765) and has 4.8% of its buildings in the SFHA. It is also worth noting that some census tracts have no buildings in the SFHA whereas others are completely contained within it. The table also shows that almost half of the buildings (49%) in the average census tract in our expanded dataset are located in periphery band 5. Before turning to the flood insurance take-up variables in the table, it is helpful to introduce some notation. Let $Buildings_c$ denote the number of buildings (footprints) in census tract c. Similarly, define $Buildings_{c,b}$ as the number of buildings at the intersection between census tract c and flood-zone band b. Then the *share* of buildings in a census tract that are located within band b is given by

$$ShBand_{c,b} = \frac{Buildings_{c,b}}{Buildings_c}. \tag{2}$$

We focus on residential policies for occupancy types that exclude apartment buildings (i.e. 1-family to 4-family homes). Hence, all buildings in our analysis of insurance take-up can be referred to as houses. We also define $nPol_{c,t}$ as the number of new (annual) household-level policies in year t and census tract c. Lacking the address of individual policies, or its geographic coordinates, we are unable to aggregate policies at the building level. As a result, it is possible for, say, a 2-family house to have 2 insurance policies, if each household chooses to buy flood insurance. Thus, conceivably, the take-up rate in a census tract could be higher than one. Specifically, we define the take-up rate for census tract c in year t by

$$TR_{c,t} = \frac{nPol_{c,t}}{Buildings_c}, \tag{3}$$

which could attain values higher than one but, in practice, happens only in very rare instances.[16]

Table 5 shows that the average take-up in a census tract is 35.6 policies over our period of interest (or 10.6% of all buildings in the census tract). However, there are some census tracts in our data with zero new flood insurance policies in some years and there is one census tract with a record take-up of 2,194 policies and a maximum take-up rate of 1.55 (or 155%). The table also shows that the share of buildings that flooded during Sandy in the average census tract was 13.6% (with an average depth of 0.37 feet).

5.2 Flood insurance coverage prior to Sandy

Our analysis of the pre-Sandy insurance coverage across New York's flood zone and its periphery will be based on the following specification. The dependent variable is the number of flood insurance policies purchased in census tract c in a year prior to Sandy, such as 2010.[17] We postulate that the number of insured households in a census tract c is determined by

$$nPol_c^{2010} = \alpha + \beta_1 shBand1_c + ... + \beta_5 shBand5_c + \gamma avElev_c + \varepsilon_c, \tag{4}$$

where $shBand1_c$ is the share of buildings in census tract c located in periphery's band 1, and so on. It is important to note that the omitted category in the equation is the share of buildings in the 100-year flood zone ($shFZ100_c$). Last, $avElev_c$ is the average elevation of the buildings in the census tract. We note that the interpretation of intercept α is the number of insured households in the census tract in a hypothetical census tract with all buildings located within the flood zone. Similarly, β_1 is the number of insured households in a hypothetical census tract with all buildings located in band 1 of the periphery and the other coefficients are interpreted analogously. In addition, we will also estimate a model where the dependent variable is the take-up rate in the census tract.

The estimates are collected in Table 6. The dependent variables in columns 1-4 are

[16] Furthermore, we measure the number of buildings in a census tract on the basis of the existing footprints prior to Sandy. Hence, our count does not incorporate new buildings added between 2013 and 2022. This is another reason for why census-tract take-up rates are not bounded by one. Alternatively, we could have normalized the number of insurance policies by the overall number of households in a census tract and produced a measure of census-tract take-up rates that would have been bounded by one. Ultimately, we chose to use the number of buildings in our normalization to maintain consistency with the building-level flood risk analysis in section 4.

[17] Our preferred pre-Sandy baseline is year 2010, although we will report estimates for years 2009 through 2012. Keep in mind that hurricane Irene took place in 2011. This hurricane could have severely impact New York but, in the end, it weakened before reaching the city and caused very limited flooding but could have affected insurance take-up in 2011. Likewise, the 2012 insurance take-up was already affected by Sandy, which landed in New York on October 29, 2012.

the number of new (household-level) policies in a census tract in pre-Sandy years 2009-2012, respectively. Columns 1 and 2 show that a census tract fully within the SFHA was estimated to have around 645 annual policies in years 2009 and 2010. In years 2011 and 2012, insurance take-up for such a county would have increased to 671 and 681, respectively, possibly due to hurricanes Irene and Sandy, respectively. In turn, columns 5-8 conduct the analysis for take-up rates (as in Equation 3).[18] According to our estimates, the take-up rates for hypothetical census tracts fully contained in the flood zone would have ranged from 1.05 to 1.07 between 2009 and 2012. In other words, there would be slightly more than one insurance policy in the average building in these tracts, indicating a very high coverage in New York's SFHA prior to Sandy.

Let us now focus on insurance take-up in 2010, our preferred pre-Sandy baseline year. As shown in column 2, a census tract fully contained within the SFHA is predicted to have 644 flood insurance policies. Similarly, a census tract fully in band 1 is predicted to have only 121 policies (i.e. $644 - 523 = 121$), and tracts fully contained outside of band 1 would have effectively zero insurance take-up. Thus, insurance take-up in the SFHA in 2010 was 5.3 times as large as in band 1 (or 644/121), and take-up declines as we move away from the boundary of the SFHA.[19]

Because the density of buildings and residents could differ in the SFHA and its periphery, it is important to conduct the analysis in terms of take-up rates. As shown in column 6, the predicted 2010 take-up rate for a hypothetical census tract fully contained in the SFHA is predicted to be 1.04 (i.e. 104 policies for each 100 buildings), whereas the corresponding value for a census tract within band 1 is predicted to be 0.24.[20] Hence, our estimates imply that the 2010 take-up rate in the SFHA was 4.3 times as large as in band 1 (or 1.04/0.24). In sum, insurance take-up in the SFHA was 4 to 5 times larger than in band 1. Thus, the relative flood insurance take-up in the SHFA and band 1 is essentially the same regardless of whether we focus on policy counts or take-up rates. Table 6 provides yet another way to compare the take-up in band 1 in relation to the SFHA. At the bottom of column 6 (which

[18] It is also worth noting that insurance take-up was almost unaffected by hurricane Irene, which reached NYC on August 27, 2011. All mass transit in NYC was halted on 8/27/2011 and several hospitals and numerous nursing homes in flood-prone areas were evacuated. In the end, flooding in the city was limited but some upstate counties experienced large flooding and substantial damage. Thus, purely informational signals may have limited effect on insurance take-up.

[19] Our findings are consistent with Brody et al. (2018) who show that insurance claims outside the flood zone are inversely related to the distance from the boundary. Naturally, the number of claims submitted after a flooding event reflect both the severity of the event *and* the number of active insurance policies in the area.

[20] To put these estimates in perspective, at the bottom of Table 6 we report the expected number of policies for the average census tract (with the composition described in Table 5), which is predicted to be 0.08 (or 8 policies for each 100 buildings). In comparison, the 2010 take-up rate for the average census tract with a majority of buildings in the SFHA is predicted to be 0.76. For more details, see section 7.

refers to take-up in 2010), we find the predicted take-up rates for the average census tract with a majority of buildings located within the SFHA and for the average census tract with a majority of buildings in periphery band 1. By this yardstick, the take-up rate in SFHA-majority tracts is about 3 times as large as the corresponding value for band-1-majority tracts (predicted to be 75.6% and 25.8%, respectively).

The relative take-up rates in band 1 (vis-a-vis the SFHA) do not reveal whether homeowners in this area were, or not, under-insured. Under the premise that realized flooding during Sandy revealed flood risk in a particular location, we can compare the *relative* flooding rates in the SFHA and band 1 to the *relative* insurance take-up rates. As reported in Table 3 (column 1), during Sandy, 87 percent of the buildings in the SFHA and 32 percent of the buildings in band 1 flooded, respectively. Thus, flood risk in the latter was 2.7 times the corresponding value for band 1. Since we found that SFHA take-up relative to band 1 is between 3 and 5, depending on how we measure it, it seems that properties in band 1 may have been somewhat under-insured.

It is interesting to compare our estimates of insurance coverage in and out of the SFHA to existing estimates in the literature. Bradt et al. (2021) estimate nationwide insurance take-up rates close to 50% in the SFHA, and roughly 2% take-up rates outside of the flood zone. Our estimates uncover much higher insurance take-up rates in New York's flood zone prior to Sandy (around 75 policies for each 100 buildings for the average census tract with a majority of buildings in the SFHA in year 2010). Our estimates also reveal substantial take-up rates in band 1 (around 26 policies for each 100 buildings for the average tract with a majority of buildings in band 1) and negligible take-up in band 2 and beyond, which is consistent with the very low estimate by Bradt et al. (2021) for non-SFHA areas.[21]

In sum, prior to Sandy, take-up rates were very high and stable in New York's SFHA. Take-up was also substantial in periphery band 1 with roughly a 1-in-4 (or 1-to-5) ratio relative to the corresponding value in the SFHA. In the previous section we estimated that the odds of flooding during Sandy in band 1 were about 1-in-3 relative to the value in the flood zone. Thus, we conclude that homeowners in periphery band 1 had under-estimated the risk of flooding and, as a result, too few of them had purchased flood insurance prior to Sandy. The remainder of the section focuses on the effect of Sandy on flood-insurance take-up and flood risk beliefs.

[21] Dixon et al. (2017) surveyed 2,890 households in New York's flood-prone areas and estimated that 40% of the buildings in the SFHA had flood insurance at the time of Sandy (in 2012). Our analysis is based on the universe of policies and buildings in New York's SFHA and its surrounding area and, thus, not affected by survey response bias.

5.3 The effect of Sandy on flood-insurance take-up

Let us now turn to the analysis of the effects of Sandy on flood insurance take-up. We begin with a simple specification that allows us to explore the effects across the 100-year flood zone and the periphery over a 5-year period. Specifically, we estimate the census-tract changes in insurance take-up between years 2010 (preceding hurricanes Irene and Sandy) and 2015 as a function of the geography of the census tract:

$$\Delta nPol_c \quad = \quad \alpha + \beta_1 shBand1_c + ... + +\beta_5 shBand5_c + X_c'\gamma + u_{c,t}, \tag{5}$$

where the dependent variable is the 2010-2015 change in the number of households taking up flood insurance in census tract c (or the corresponding take-up rate as defined in Equation 3). We also estimate versions of these equations where we control for average elevation in the census tract and the share of flooded buildings during Sandy. Including the latter is useful to test the importance of the experience of flooding in the census tract as a factor that drives insurance demand.

The results are collected in Table 7. Column 1 reports within-tract changes in annual policies between 2010 and 2015. The estimates entail large increases in the SFHA and in band 1: the take-up in a tract fully within the flood zone would have increased by 144 policies. Similarly, for a tract fully contained in band 1 of the periphery, the number of new policies would have increased by $144 - 52 = 92$ policies. In comparison, very little response (or none) is found in bands 2 through 5 of the periphery. Thus, hurricane Sandy led to increased insurance take-up in the SFHA and in band 1 almost exclusively. We also note that the estimate for the coefficient corresponding to band 1 is not statistically different from zero. Thus, we cannot reject the hypothesis of an equal increase in take-up in the flood zone and in band 1.

Column 2 controls for average elevation, which has a negative and statistically significant effect on insurance take-up, but does not change the pattern. The negative effect of elevation is not surprising, given that elevation was found to lower the probability of flooding (Table 3). We can provide a more intuitive description of the magnitudes by referring to representative census tracts. As can be seen at the bottom of Table 7, the predicted 2010-2015 change in take-up in the average census tract is only 13 policies. However, the predicted change is much larger for majority-SFHA census tracts (at around 111 policies) and for majority-band-1 census tracts (71 policies). It is worth emphasizing that the take-up increase in band 1 was proportionally much larger than the increase in the flood zone, because of the much lower baseline take-up prior to Sandy (by a 1-to-5 ratio).

Column 3 also controls for the share of houses in the tract that flooded during Sandy. Not surprisingly, the coefficient for this variable is positive and highly significant. It is also worth noting that the estimate for the changes in take-up in the SFHA and band 1 are greatly diminished and, in fact, no longer statistically significant. These estimates suggest that the experience of flooding was responsible for the lion's share of the increase in insurance take-up in these areas. In other words, the geographic overlap of a census tract with the periphery of the flood zone and its average elevation are *proxies* for the tract's flood risk and, once we control for the actual extent of flooding in the tract, these proxies lose their explanatory power.

Columns 4-6 repeat the analysis but switching the dependent variable to be the 2010-2015 change in the *take-up rate* (or TR). Columns 4-5 indicate a 21-22 percentage point increase in the take-up rate for a hypothetical census tract completely contained within the SFHA and a very similar (or slightly lower) increase in band 1. As shown at the bottom of column 5, these estimates imply 18 and 15 percentage-point increases in the take-up rate in the average majority-SFHA and in the average majority-band-1 census tracts, respectively. In Table 6, we estimated that a pre-Sandy take-up rate of 26% in the average census tract with a majority of houses located in band 1. Hence, Sandy led approximately to a 57% increase in the take-up rate in periphery band 1 between 2010 and 2015. In comparison, the corresponding increase in the SFHA was around 15/76 = 24%. These figures are consistent with Bradt et al. (2021), who argued that the insurance take-up elasticity to perceived flood risk is higher for homeowners located in the periphery of the flood zone than for those within it.

Lastly, when we control for the share of Sandy flooded buildings in our specification (column 6), the estimated residual increase in take-up rates in the SFHA vanishes and is also substantially reduced in band 1, again suggesting that the experience of flooding during Sandy was the key force behind the observed increases in insurance take-up in the flood zone and its immediate periphery.

5.4 Dynamic effects of insurance take-up

Having established that the bulk of the insurance take-up response to Sandy was concentrated in the SFHA and periphery band1, we analyze in greater detail the dynamics of the response. In particular, we wish to determine how quickly the purchase of flood insurance peaked after Sandy's landfall and for how long the newly-insured residents continued to hold onto their policies. This analysis will provide important clues regarding homeowners' motivations to purchase flood insurance after Sandy, which could potentially differ for those located in the SFHA and those in its more immediate periphery.

We now estimate the following model:

$$nPol_{c,t} = \alpha_t + \beta_0^t shSFHA_c + \beta_1^t shBand1_c + X_c'\eta + \lambda^t shFlooded_c + u_{c,t}, \quad (6)$$

where the dependent variable is the number of new policies in census tract c and year $t = 2009, ..., 2021$ (or the corresponding take-up rate). A set of yearly dummy variables, denoted by α_t, captures the trajectory of the annual take-up for the aggregate of periphery bands 2 through 5. The main coefficients of interest are β_0^t and β_1^t, which correspond to the interaction between year dummies (2012-2021) and the share of buildings in the census tract within the SFHA and the share of buildings in band 1, respectively. These coefficients identify the potential divergence in the post-Sandy take-up trajectory relative to the corresponding value for the aggregate of periphery bands 2 to 5.

Vector X_c contains the share of buildings in the SFHA and the share in band 1 (without any interactions), which capture the baseline (average annual) take-up prior to 2012 in each census tract. Last, in some specifications we also include interaction terms between yearly dummy variables (for years 2012-2021) and the share of buildings in the census tract that flooded during Sandy, whose effects is captured by coefficients λ^t. In those specifications, vector X_c also includes the share of Sandy flooded buildings in the census tract (without any interaction terms). footnoteAn alternative specification with census-tract fixed-effects, which absorb vector ηX_c yields numerically identical estimates for α_t, β_0^t and β_1^t. Last, we cluster standard errors by zip code, which allows for contemporaneous correlated shocks across census tracts (but requires uncorrelated shocks across zip codes) and, hence, it is less restrictive than clustering at the level of census tracts.

The estimates of equation Equation 6 are reported in Table 8. Column 1 simply traces the dynamic evolution of the number of insured households in our (expanded) area of interest. The average take-up prior to Sandy (for years 2009-2011) was 29 policies per census tract. In 2012 (denoted by T0), we observe a small increase in take-up (of 2 policies) probably due to the immediate response to Sandy in 2012Q4. The increase was much more pronounced one year later (17 policies in T1) and remained elevated for the following 2 years (13 policies in T2 and T3). After that, we observe a reduction in the intensity of the response, but the average annual take-up remained about 7 policies higher than prior to Sandy for the period 2016-2021 (i.e. about 25% higher than the baseline take-up).

Column 2 separates out the differential response to the share of buildings in the SFHA from the share of buildings in the periphery (taken as a whole). Clearly, the baseline take-up was much higher in the SFHA than elsewhere: 9 policies in the periphery of the SFHA and $704+9 = 713$ policies in the SFHA. Already in 2012 (T0), take-up in the periphery increased by 1 policy whereas the increase in the SFHA was 34 policies (i.e. almost 5% of the baseline value). In turn, in the year after Sandy (T1), take-up increased by about 8 policies in the periphery but by 339 policies in the SFHA (i.e. about 48% of the baseline value). From that point on, the estimates indicate a rapid reversion toward the baseline value in the SFHA, but a higher degree of persistence *outside* of the SFHA. As shown in column 3, controlling for average elevation in the census tract barely affects the dynamic response to Sandy.

Column 4 estimates our preferred model, which allows for different take-up trajectories in the SFHA, band 1 and the rest of the periphery. The estimates clearly show that the take-up response was entirely concentrated in the SFHA and its adjacent area (band 1). Secondly, the take-up increase in the SFHA one year after Sandy (in T1) was very large (290 policies for a census tract completely contained in the SFHA) but the intensity of the response declined rapidly, vanishing in the 4th year after the storm. In contrast, while the take-up in band 1 also increased importantly one year after Sandy (by about 94 policies or 85% of the baseline in a hypothetical census tract completely contained within band 1), it remained elevated at similar values throughout our sample period (ending in 2021, 9 years after Sandy's landfall).

Column 5 in Table 8 is analogous to column 4 but the dependent variable is now the *take-up rate*. The pattern is essentially the same as in column 4, but the estimates provide a better sense of the magnitude of the effects. Prior to Sandy, the take-up rate in a hypothetical census tract completely contained in the SFHA would have been close to 100%.[22] This take-up rate remained unchanged in 2012 (T0) but jumped vigorously in 2013 (T1): a census tract fully contained in the SFHA would have experienced a 46 percentage-point increase in its take-up rate in 2013. In comparison, the 26% take-up rate (in 2010) in the representative majority-band-1 census tract (Table 6, column 7) would have increased by 17 percentage points in 2013 (and essentially remained at that level for the following 8 years and possibly beyond that).

Note that our results here are remarkably consistent with the those reported in the previous section regarding the take-up increase between 2010 and 2015. In Table 7 we estimated that the take-up rate increased by about 18 percentage points in the representative majority-SFHA census tract and by 15 percentage points in the majority-band-1 census tract between those two years. However, our analysis of the dynamic take-up response reveals that while the increase in take-up in the SFHA lasted only 3 years, the response in band 1 was much more persistent. These patterns are more clearly seen in Figure 2. The top figure illustrates the striking, but short-lived, flare-up in take-up in the SFHA immediately after

[22]More realistically, our earlier estimates (Table 6, column 7) implied a 76% take-up rate in the representative SFHA-majority census tract in year 2010.

Sandy. In contrast, the middle figure reveals a less intense but much more persistent take-up response in periphery band 1. Last, the bottom figure clearly shows the lack of response in the rest of the periphery. The following section will investigate the mechanisms that explain the diverging dynamic responses in the flood zone and its immediate periphery.

It is worth noting that Bradt et al. (2021) also documented increases in insurance take-up in census tracts that experienced a large flooding event and were included in a federal disaster declaration zone, with less persistent effects in the SFHA than outside of it. However, their analysis requires conditioning on the receipt of federal disaster aid, which may be correlated with the pre-storm insurance take-up in the area, and the persistence of the estimated take-up response is much shorter than we find.[23] Our analysis suggests that a precise delimitation of the area outside of the SFHA with significant flood risk is important in order to obtain accurate estimates of the take-up response to a storm and its persistence over time.

5.5 Mechanical response in the flood zone

Prior literature has documented short-lived increases in flood insurance take-up in the aftermath of large-scale flooding events. Gallagher (2014) found that insurance take-up in flooded areas spikes one year after large-scale flooding events and quickly returns to its baseline level, which he interpreted as a behavioral response. Similarly, Kousky (2017) estimated that, one year after a hurricane, insurance take-up rates in the affected areas increase by about 7% and the effect vanishes after 3 years. In contrast to the previous study, she argues that the bulk of the increase in take-up is purely mechanical. Homeowners with approved federal individual assistance grants are required to maintain flood insurance for a minimum of 3 years.

Next, we propose (and apply) an alternative approach to isolate the mechanical increase in insurance take-up driven by the 3-year insurance requirement tied to relief aid. Our approach relies on the information regarding the *types* of insurance policies purchased by the residents of a census tract. We partition all flood insurance policies into 3 groups: Group Flood Insurance Policies (or GFIP), Preferred Risk Policies (or PRP) and Other policies.

For the purpose of this section, the most relevant policy type is GFIP.[24] These policies are available to homeowners in the SFHA who experienced flooding-related damages (and were denied a subsidized SBA loan). Importantly, purchasing a GFIP and maintaining it for

[23]The comparison of Figures 4a and 4b in Bradt et al. (2021) shows that conditioning on the receipt of federal aid is needed to obtain the meaningful take-up responses to flooding events. The estimates depicted in Figure 4b suggest that insurance take-up rates peak one year and two years after the event in the SFHA and outside of it, respectively.

[24]We shall describe PRPs in the next section, where they will play an important role.

a 3-year period is required in order to receive any relief aid.[25] Hence, tracking the purchases of GFIPs immediately after Sandy helps isolate with high accuracy the *mechanical* increase in insurance take-up.

Next, now turn to the estimation of the model in Equation 6 but focusing exclusively on purchases of GFIPs. The results are collected in Table 9. Column 1 simply reproduces the earlier estimates of the event study referring to the overall number of insurance policies (corresponding to column 4 in Table 8), which shows the surge in take-up in the SFHA in 2013 (T1). The increase in 2013 in a hypothetical census tract contained within the SFHA was estimated at 290 policies (or about 45% of the baseline value in 2010). Importantly, the flare-up in insurance take-up was short-lived, lasting for about 3 years. The comparison of columns 1 and 2 of Table 9 indicates that the majority of the increase (between 50 to 70 percent) in insurance purchases in the SFHA corresponded to GFIPs. Column 2 also shows that the increase in GFIP purchases vanished also after 3 years, when homeowners wishing to maintain flood insurance were forced to buy regular (non-GFIP) policies, which entailed substantially higher premiums. The estimates indicate that the majority of homeowners chose to discontinue their coverage. Additionally, the estimates show no significant purchases of GFIPs outside of the SFHA, as one would expect.[26] All in all, our estimates strongly reinforce the findings in Kousky (2017), with the advantage of directly measuring purchases of the relevant type of insurance policy.

Given that a substantial part of the take-up response in the SFHA was mechanically driven by the 3-year insurance requirement that applied to many of the federal aid recipients, and these homeowners discontinued their coverage as soon as they could, it seems plausible that the flooding experienced by these homeowners during Sandy did not change their beliefs regarding flood risk. Or, if it did, the upward revision of flood risk beliefs was not large enough to justify purchasing the more expensive flood insurance available to them after GFIPs became unavailable.

5.6 Scarring effect in the periphery

Hurricane Sandy brought about extensive flooding in New York's flood zone and its periphery. As noted earlier, the share of flooded houses in the SFHA was 87% (Table 3). However, there

[25]GFIPs are also very affordable, with annual premia around $600 in 2022. Homeowners in the SFHA who did receive an SBA loan are also required to purchase insurance, and those policies fall in the category of *Other* policies.

[26]Column 4 shows that in the first year after hurricane Sandy (T1), there was also a significant increase in the *Other* flood insurance policies in the SFHA (in year T1), likely reflecting purchases by homeowners who received an SBA loan. This is also a mechanical effect since recipients of these loans need to show proof of insurance at the time of receiving the loan, but typically drop that coverage after one year.

was also widespread flooding in periphery band 1, where 32% of all houses flooded as well. Moreover, we showed in Table 8 (columns 4 and 5) that insurance take-up in band 1 increased substantially in 2013 and this increase lasted for 9 years (or longer), even though homeowners in this area were not subject to an insurance mandate.

The highly persistent increase in take-up in band 1 was in stark contrast with the short-lived increase in insurance purchases in the SFHA. Next, we explore to what extent it was tied to the (probably unexpected) flooding experienced by homeowners in that area. To examine this question we make use of our building-level flooding data to measure the fraction of properties in each census tract that were located in band 1 *and* experienced flooding during Sandy. We then go on to augment Equation 6 to include to the share of flooded buildings in band 1, allowing for time-varying effects on insurance take-up. The results are collected in columns 6 and 7 of Table 8. Clearly, the coefficients for the new terms in the model are highly significant and strongly suggest that the increase in insurance take-up from 2013 onward was intimately tied to the extent of flooding in periphery band 1. In other words, it appears that it was precisely the actual flooding of houses in the neighborhood that ignited and sustained the increase in insurance take-up. It is also possible that, for many homeowners in band 1, Sandy was the first storm causing substantial flooding in their homes, leading to an update in their beliefs regarding flood risk. As argued in Ortega and Taspinar (2018), in learning models, the experience of a rare event can trigger a persistent change in flood risk beliefs, sometimes referred to in the literature as *scarring*.

In order to explore this interpretation further, it is helpful to focus on another type of insurance policies. Preferred Risk Policies (PRPs) are affordable policies offering building and content coverage that are available only to properties *outside* of the SFHA with a *favorable loss history*.[27] The comparison of columns 1 and 3 in Table 9 shows that the vehicle for practically the full response to Sandy observed in periphery band 1 was due to PRPs. For instance, in year T1, out of the 94-unit predicted increase in take-up in a census tract fully within band 1, 86 of these would correspond to PRPs. Because PRPs are targeted to (non-SFHA) homeowners experiencing flooding for the first (or second) time, these estimates reinforce the interpretation that the increase in insurance take-up in band 1 was driven by upward revisions of flood risk beliefs by homeowners that were surprised by the flooding caused by Sandy in their neighborhood.

Summing up, our analysis of insurance take-up has documented a novel finding: a large and highly persistent increase in insurance take-up in the aftermath of hurricane Sandy in

[27]PRPs are very affordable, with annual premiums around $400. Only houses that flooded at most once before are eligible for this type of flood insurance policy. Kousky et al. (2020b) analyzed insurance take-up in Houston after hurricane Harvey and observed widespread purchases of PRPs outside the SFHA.

the areas adjacent to (but not part of) the SFHA. This increase was purely voluntary (i.e. not mandated by the government) and consistent with an upward revision of homeowners' flood risk beliefs in response to the extent of flooding in their neighborhood.

6 Informational signals and insurance take-up

The analysis in the previous section suggested that experiencing flooding in one's own neighborhood may be an important determinant of flood risk beliefs and insurance take-up. It is also possible that informational signals trigger belief updating and also affect the decision to purchase flood insurance.

Gallagher (2014) showed that insurance take-up spikes shortly after a flooding event, including in *non-flooded* locations neighboring the flooded area (for a limited period of time). In addition, several studies have investigated if the information conveyed through updates to FEMA flood maps has tangible effects. Indaco et al. (2019) show that the adoption of new flood maps (in Virginia Beach) had a large, detrimental effect on the value of the affected properties. More recently, Weill (2022) argues that the adoption of new flood maps affects insurance take-up. Specifically, he documents increases (decreases) in take-up when properties are mapped into (out of) the SFHA (but fails to find effects associated with the digitizing of previously available flood maps). In turn, Orellana-Li (2023) shows that homeowners (in Florida) who are drawn into high-risk flood zones are more likely to sell their homes (when the replacement cost exceeds the maximum flood insurance coverage).[28]

As it turns out, FEMA released updated preliminary flood maps for New York City during the first half of 2013, which were expected to become effective 2 or 3 years later. However, because of a successful appeal by New York City, a decade later, those maps are yet to be adopted.[29] As discussed in Dixon et al. (2017), the new maps substantially expanded New York's SFHA, including over 20,000 additional structures. Because they had previously not been considered at risk of flooding, many of these buildings had not been built according to the elevation requirements of the high-risk zone. The study also points out that these homeowners could expect large flood insurance premium increases at the time the preliminary maps became effective.

While the release of the 2013 preliminary flood maps could plausibly have had an immediate effect on values of the properties that would be included in the SFHA, it is unclear

[28]A related set of studies investigates the informational effects of recent disasters on home values (Hallstrom and Smith (2005),Ortega and Taspinar (2018), Gibson and Mullins (2020)), home improvements (McCoy and Zhao (2018)) and insurance take-up (Gallagher (2014)).

[29]It is worth noting that, as far as we can tell, no previous studies have found an effect of the *release* of preliminary flood maps, *prior to their adoption*, on any outcomes.

why it would affect their owners' decision to purchase flood insurance. After all, insurance mandates would not apply until the new maps were *adopted*, which was scheduled to happen a few years after from the release. Moreover, FEMA allows owners of homes newly mapped into the SFHA to purchase flood insurance at their pre-release annual premiums for a few years. Two previous studies have analyzed the effects of the release of the 2013 preliminary maps on housing values in New York's flood zone.Ortega and Taspinar (2018) did not find any effects, but Gibson and Mullins (2020) argue that the release of the maps negatively affected the values of the properties newly included in the flood zone.

Next, we analyze if the release of the preliminary maps (prior to their adoption) had effects on insurance take-up. Besides interesting in its own right, this question is relevant for our paper because of the temporal and geographic overlap of Sandy and the preliminary flood maps. As a result, the post-Sandy increases in take-up documented in the earlier sections could be partly due to the release of the preliminary flood maps.

6.1 Cohort analysis

The preliminary (base flood elevation) maps for New York were released piecemeal during the first half of 2013 and, in fact, most of these releases took place during the second quarter of 2013. It is thus interesting to compare the flood insurance purchases of two cohorts: the homeowners that purchased insurance *prior* to the release of the new maps (between 12/1/2012 and 3/1/2013) and those that purchased insurance *afterwards* (between 3/1/2013 and 6/1/2013). Presumably, the former would be responding to Sandy (which hit New York on 10/29/2012) whereas the latter would also be responding to the release of the new maps.[30] It is also important to note that the area of interest is periphery band 1 since most rezoned houses that will become part of the new flood zone are located in the immediate area outside of the (old) flood zone.

The findings are reported in Table A4. In regards to the *pre-map* cohort, column 1 shows an important increase in take-up in $T1 = 2013$ in band 1 of 41 policies (i.e. a 4.1 percent increase relative to baseline as seen in column 2). Interestingly, column 3 reveals a very similar behavior in band 1 for the *post-map* cohort: a 35-policy increase (although lower than a 1 percent increase relative to baseline). Had the release of the preliminary maps triggered an increase in take-up, loaded on top of the effect of Sandy, we would have expected a larger increase in take-up for the *post-map* cohort. Instead, we find that the

[30] We acknowledge this is not a perfect test since homeowners purchasing insurance immediately after Sandy may have been more motivated to do so (e.g. because of greater flood damage). However, bottlenecks in processing relief applications and flood insurance purchases likely introduced a great deal of randomness in determining which insurance purchase requests were processed earlier.

increase for this cohort is similar (in levels) and substantially lower (in rates) than what we estimated for the *pre-map* cohort.

6.2 Horse race: the effects of the storm versus the new flood map

Turning now to our second exercise, we extend our event study specification to run a horse race between the effects of Sandy flooding in band 1 and those of the rezoning of houses into the new flood zone. The estimates are collected in Table 10. Columns 1 and 2 simply report earlier estimates, showing a large and persistent increase in insurance take-up in band 1 from 2013 onward. Columns 3 and 4 extend the model to include the share of houses in the census tract that were located in band 1 *and* were flooded by Sandy, interacted with years-since-Sandy dummy variables, along with analogous terms for the share of buildings that were located in band 1 *and* included in the SFHA in the preliminary maps. In column 3 (policy counts), none of these interaction terms are statistically significant in the post-Sandy years, largely because of the high collinearity between the two sets of variables. However, the estimates in column 4 (take-up rates) are more informative and strongly suggest that the share of flooded buildings in band 1 is responsible for the post-Sandy increase in insurance take-up in this area (shown in columns 1 and 2).

As was the case in the cohort analysis just discussed, the evidence here is not conclusive. However, taken together, both exercises fail to find effects of the release of the preliminary version of the new flood maps on insurance take-up in band 1, the area most affected by the proposed expansion of the flood zone. The analysis in this section suggests that it is much more likely that the large increase in insurance take-up documented throughout the paper is a response to the flooding caused by hurricane Sandy than to the release of information regarding the new flood maps.

7 Conclusions

Our main goal was to develop a new approach to estimate flood risk in and *outside* of the 100-year flood zone and flood insurance take-up. We apply our approach to New York City in the aftermath of hurricane Sandy, for which we have access to building-level flooding data and data on the number (and type) of flood insurance policies purchased in each census tract between 2009 and 2021.

Our analysis of the impact of Sandy in terms of flooding in the flood zone and its periphery shows that, unsurprisingly, the extent of flooding was highest in the SFHA (just shy of 90% of all houses flooded), falling rapidly as we moved away. However, the flooding rates remained high within 250 meters from the edge of the periphery (band 1), where we estimate that around 32% flooded during Sandy. These estimates suggest that the demand for flood insurance surged in this area of the flood zone periphery, an area where homeowners were not subject to any flood insurance mandate and might even have been unaware of the risk of flooding.

We also found that, prior to Sandy, insurance take-up rates were fairly high and stable in the SFHA (around 78 policies for each 100 houses in the average census tract with a majority of houses in the flood zone). In comparison, take-up rates were much lower in periphery band1 (estimated to be 1/4 or 1/5 of the value in the SFHA) and negligible in the rest of the periphery of the SFHA. In fact, comparing take-up rates and flood rates in band 1 and in the SFHA, it appears that band-1 homeowners may have been somewhat underinsured. Thus, it is highly likely that Sandy revealed information that might have led to an upward revision of flood risk beliefs.

Our analysis of the response to Sandy in terms of insurance take-up reveals a differential dynamic response in the flood zone and its periphery. First, we find a large and short-lived surge in take-up in response to Sandy in the SFHA, which practically vanished 3 years after the storm. This finding, and our analysis of the specific type of insurance policies purchased by SFHA homeowners during this period, confirms the conclusions drawn by Kousky (2017) who argues that the take-up response in the SFHA after a flooding event is largely a mechanical response due to insurance requirements to receive relief aid. More interestingly, our estimates provide new evidence of a largely *permanent* increase in insurance take-up in band 1 of the periphery, lasting more than 10 years. Importantly, our estimation of the risk gradient in the vicinity of the flood zone provided an accurate characterization of the area outside of the flood zone exposed to significant flood risk, and this was crucial in order to estimate the persistence of the insurance take-up response to the storm. Thus, a careful analysis of the relevant flood zone periphery for each major coastal urban area may lead to a better understanding of insurance take-up dynamics in the United States.

The persistent increase in insurance take-up in the periphery of the flood zone is consistent with voluntary purchases by homeowners that had experienced severe flooding (possibly for the first time) during Sandy and updated their beliefs regarding flood risk. This finding is based both on the timing of the response to Sandy as well as the *type of policy* purchased by these homeowners, which was available only to homeowners that had claimed flood relief at most once in the past. Last, our findings suggest that experiencing substantial flooding for the first time is a powerful determinant of the decision to purchase flood insurance and *maintain* it. Our findings also have implications for the new insurance pricing system, known as *Risk Rating 2.0*, introduced by FEMA at the end of 2021.[31] The distinctive feature of the new system is that it tailors annual premiums to each property, on the basis of a wide array of flood risk predictors.[32]

Our analysis showed that properties located within 250 meters of the SFHA face substantial flood risk. In addition, we showed that many homeowners in this area chose to purchase flood insurance to protect their properties, relying on an affordable type of insurance policy (known as PRP) only available to properties outside of the SFHA (that had not requested flood relief more than once in the past). Because distance to the SFHA was a highly significant predictor of flood risk in our estimates, it is very likely that homeowners located within 250 meters of the SFHA will be experiencing a substantial increase in annual premiums under the new pricing system, which will likely reduce insurance take-up rates in this area.

In a recent study, Mulder and Kousky (2023) provide evidence of large premium increases outside of the SFHA for properties with characteristics usually tied to flood risk. The authors also point out that the new pricing system is opaque, that is, homeowners are unable to know how the specific characteristics of their property map into the annual premiums quoted to them by the new system. Until FEMA resolves this problem, our flood risk analysis based on each property's distance to the SFHA, ground elevation and recent flooding history can be used to produce building-specific flood risk estimates that rely exclusively on publicly available data.

[31] Initially, the new system applied to new policies only but, since April 2022, it applies to renewals as well.

[32] Importantly, while flood maps are no longer relevant for insurance pricing purposes, the insurance mandate for properties within the SFHA with federally backed mortgages remains in effect. For more details, see `https://www.fema.gov/flood-insurance/risk-rating`.

References

Bates, Paul D, Niall Quinn, Christopher Sampson, Andrew Smith, Oliver Wing, Jeison Sosa, James Savage, Gaia Olcese, Jeff Neal, Guy Schumann, et al. (2021), "Combined modeling of us fluvial, pluvial, and coastal flood hazard under current and future climates." *Water Resources Research*, 57, e2020WR028673.

Billings, Stephen B, Emily A Gallagher, and Lowell Ricketts (2022), "Let the rich be flooded: the distribution of financial aid and distress after hurricane harvey." *Journal of Financial Economics*, 146, 797–819.

Bradt, Jacob T, Carolyn Kousky, and Oliver EJ Wing (2021), "Voluntary purchases and adverse selection in the market for flood insurance." *Journal of Environmental Economics and Management*, 110, 102515.

Brody, SD, Antonia Sebastian, Russell Blessing, and PB Bedient (2018), "Case study results from southeast houston, texas: identifying the impacts of residential location on flood risk and loss." *Journal of Flood Risk Management*, 11, S110–S120.

Cohen, Jeffrey P., Jason Barr, and Eon Kim (2021), "Storm surges, informational shocks, and the price of urban real estate: An application to the case of Hurricane Sandy." *Regional Science and Urban Economics*, 90, URL `https://ideas.repec.org/a/eee/regeco/v90y2021ics0166046221000545.html`.

Dixon, Lloyd, Noreen Clancy, Benjamin M Miller, Sue Hoegberg, Michael M Lewis, Bruce Bender, Samara Ebinger, Mel Hodges, Gayle M Syck, Caroline Nagy, et al. (2017), "The cost and affordability of flood insurance in new york city." *RAND Corporation, Santa Monica, CA*.

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.

Galloway, Gerald E, Gregory B Baecher, Douglas Plasencia, Kevin G Coulton, Jerry Louthain, Mohamed Bagha, and Antonio R Levy (2006), "Assessing the adequacy of the national flood insurance program's 1 percent flood standard." *Water Policy Collaborative, University of Maryland. College Park, Maryland*, 2020–07.

Gibson, Matthew and Jamie T. Mullins (2020), "Climate risk and beliefs in new york floodplains." *Journal of the Association of Environmental and Resource Economists*, 7, 1069–1111, URL `https://doi.org/10.1086/710240`.

Gourevitch, Jesse D, Carolyn Kousky, Yanjun Liao, Christoph Nolte, Adam B Pollack, Jeremy R Porter, and Joakim A Weill (2023), "Unpriced climate risk and the potential consequences of overvaluation in us housing markets." *Nature Climate Change*, 13, 250–257.

Hallstrom, Daniel G. and V. Kerry Smith (2005), "Market responses to hurricanes." *Journal of Environmental Economics and Management*, 50, 541–561, URL `https://ideas.repec.org/a/eee/jeeman/v50y2005i3p541-561.html`.

Highfield, Wesley E, Sarah A Norman, and Samuel D Brody (2013), "Examining the 100-year floodplain as a metric of risk, loss, and household adjustment." *Risk Analysis: An International Journal*, 33, 186–191.

Hu, Zhongchen (2022), "Social interactions and households' flood insurance decisions." *Journal of Financial Economics*, 144, 414–432, URL `https://ideas.repec.org/a/eee/jfinec/v144y2022i2p414-432.html`.

Indaco, Agustín and Francesc Ortega (2023), "Adapting to climate risk? local population dynamics in the united states." *IZA DP 15982*.

Indaco, Agustín, Francesc Ortega, and Süleyman Taspinar (2019), "The Effects of Flood Insurance on Housing Markets." *Cityscape*, 21.

Indaco, Agustín, Francesc Ortega, and Suleyman Taspınar (2021), "Hurricanes, flood risk and the economic adaptation of businesses." *Journal of Economic Geography*, 21, 557–591, URL `https://ideas.repec.org/a/oup/jecgeo/v21y2021i4p557-591..html`.

Kousky, Carolyn (2017), "Disasters as Learning Experiences or Disasters as Policy Opportunities? Examining Flood Insurance Purchases after Hurricanes." *Risk Analysis*, 37, 517–530, URL `https://ideas.repec.org/a/wly/riskan/v37y2017i3p517-530.html`.

Kousky, Carolyn, Howard Kunreuther, Michael LaCour-Little, and Susan Wachter (2020a), "Flood risk and the us housing market." *Journal of Housing Research*, 29, S3–S24.

Kousky, Carolyn, Mark Palim, and Ying Pan (2020b), "Flood damage and mortgage credit risk: A case study of hurricane harvey." *Journal of Housing Research*, 29, S86–S120, URL `https://doi.org/10.1080/10527001.2020.1840131`.

McCoy, Shawn J. and Xiaoxi Zhao (2018), "A city under water: A geospatial analysis of storm damage, changing risk perceptions, and investment in residential housing." *Journal of the Association of Environmental and Resource Economists*, 5, 301–330, URL `https://doi.org/10.1086/695611`.

Meltzer, Rachel, Ingrid Gould Ellen, and Xiaodi Li (2021), "Localized commercial effects from natural disasters: The case of hurricane sandy and new york city." *Regional Science and Urban Economics*, 86, S0166046220302933, URL `https://EconPapers.repec.org/RePEc:eee:regeco:v:86:y:2021:i:c:s0166046220302933`.

Mulder, Philip and Carolyn Kousky (2023), "Risk rating without information provision." *AEA Papers and Proceedings*, 113, 299–303, URL `https://www.aeaweb.org/articles?id=10.1257/pandp.20231102`.

Orellana-Li, John (2023), "Who Leaves the Floodplain? Homeowner Relocation on Information and Risk Transfer." *CUNY Mimeo*.

Ortega, Francesc and Suleyman Taspinar (2018), "Rising sea levels and sinking property values: Hurricane sandy and new york's housing market." *Journal of Urban Economics*, 106, 81–100.

Ouazad, Amine and Matthew E Kahn (2022), "Mortgage finance and climate change: Securitization dynamics in the aftermath of natural disasters." *The Review of Financial Studies*, 35, 3617–3665.

Patterson, Lauren A and Martin W Doyle (2009), "Assessing effectiveness of national flood policy through spatiotemporal monitoring of socioeconomic exposure 1." *JAWRA Journal of the American Water Resources Association*, 45, 237–252.

Sastry, Pari (2023), "Who Bears Flood Risk? Evidence from Mortgage Markets in Florida." *Columbia Business School mimeo.*

Wagner, Katherine RH (2022), "Adaptation and adverse selection in markets for natural disaster insurance." *American Economic Journal: Economic Policy*, 14, 380–421.

Weill, Joakim (2022), "Perilous flood risk assessments." *Available at SSRN 4143914*.

Wing, Oliver EJ, Paul D Bates, Andrew M Smith, Christopher C Sampson, Kris A Johnson, Joseph Fargione, and Philip Morefield (2018), "Estimates of present and future flood risk in the conterminous united states." *Environmental Research Letters*, 13, 034023.

Tables and Figures

Table 1: Summary statistics. NYC

	count	mean	std	min	max
NYC					
SFHA	175769	0.088	0.284	0.000	1.000
FZ500	175769	0.076	0.265	0.000	1.000
DistSFHA	175769	4.283	3.026	0.000	10.000
Band1	175769	0.256	0.436	0.000	1.000
Band2	175769	0.242	0.428	0.000	1.000
Band3	175769	0.220	0.414	0.000	1.000
Band4	175769	0.194	0.395	0.000	1.000
Elevation	175769	13.132	12.635	-1.000	101.000
Fam14	175769	0.841	0.365	0.000	1.000
Flooded	175769	0.201	0.401	0.000	1.000
Depth	175769	0.624	1.520	0.000	15.155
NYC only Fam14					
SFHA	147552	0.086	0.280	0.000	1.000
FZ500	147552	0.074	0.262	0.000	1.000
DistSFHA	147552	4.333	3.025	0.000	10.000
Band1	147552	0.250	0.433	0.000	1.000
Band2	147552	0.244	0.429	0.000	1.000
Band3	147552	0.222	0.416	0.000	1.000
Band4	147552	0.198	0.398	0.000	1.000
Elevation	147552	13.238	12.576	-1.000	88.000
Fam14	147552	1.000	0.000	1.000	1.000
Flooded	147552	0.209	0.407	0.000	1.000
Depth	147552	0.646	1.533	0.000	13.939

Notes: NYC (5 boroughs). Data includes buildings in the SFHA (100-year flood zone) or within 1,000m of the SFHA. $DistSFHA$ is the distance from each building to the SFHA (in hundreds of meters). Bands partition the properties outside of the SFHA as follows: Band1 (<250m), Band2 (250m-500m), Band3 (500m-750m) and Band4 (750m-1000m). $Fam14$ is an indicator for 1-family through 4-family buildings.

Table 2: Flood zone and bands NYC

	Elevation	DistSFHA	FZ500	Fam14	Flooded	depth	Obs
NYC							
Band1	7.538	1.243	0.225	0.823	0.302	0.745	45004
Band2	13.838	3.734	0.043	0.850	0.104	0.237	42452
Band3	17.262	6.221	0.024	0.848	0.071	0.131	38719
Band4	19.646	8.719	0.012	0.859	0.044	0.082	34051
SFHA	2.847	0.039	0.000	0.817	0.849	3.748	15543
MN							
Band1	6.645	1.208	0.270	0.054	0.132	0.280	1203
Band2	18.010	3.703	0.016	0.097	0.008	0.017	1201
Band3	21.580	6.206	0.003	0.075	0.001	0.014	957
Band4	20.246	8.672	0.000	0.100	0.000	0.000	623
SFHA	3.528	0.063	0.000	0.034	0.599	1.766	354
BK							
Band1	3.320	1.207	0.474	0.775	0.653	1.644	7695
Band2	5.861	3.696	0.159	0.763	0.376	0.899	5852
Band3	7.151	6.249	0.091	0.783	0.265	0.564	5984
Band4	7.626	8.719	0.043	0.813	0.193	0.370	6139
SFHA	2.302	0.023	0.000	0.642	0.940	3.742	2897
QN							
Band1	5.128	1.156	0.234	0.871	0.386	1.016	15099
Band2	9.947	3.737	0.039	0.910	0.151	0.337	12995
Band3	12.815	6.199	0.016	0.905	0.094	0.134	11052
Band4	14.550	8.714	0.001	0.877	0.018	0.025	8523
SFHA	2.317	0.036	0.000	0.908	0.926	4.038	7098
SI							
Band1	11.644	1.330	0.103	0.872	0.162	0.328	15653
Band2	18.011	3.735	0.007	0.912	0.013	0.022	17354
Band3	22.343	6.217	0.004	0.918	0.007	0.013	15430
Band4	26.157	8.707	0.007	0.937	0.011	0.023	13224
SFHA	3.814	0.018	0.000	0.882	0.895	4.922	3344
BX							
Band1	8.595	1.293	0.193	0.784	0.007	0.010	5354
Band2	17.760	3.773	0.052	0.756	0.002	0.003	5050
Band3	22.383	6.248	0.023	0.740	0.001	0.001	5296
Band4	25.193	8.761	0.010	0.783	0.000	0.000	5542
SFHA	3.858	0.107	0.000	0.774	0.370	0.902	1850

Notes: Data includes properties in the SFHA (FZ100) or within 1km form it. Bands partition the properties outside of SFHA as follows: Band1 (<250m), Band2 (250m-500m), Band3 (500m-750m) and Band4 (750m-1000m). Columns *Flooded* and *Depth* refer to flooding during hurricane Sandy.

Table 3: Flood risk gradient. Only houses

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DepVar	Flooded	Flooded	Flooded	Flooded	Flooded	Flooded	Flooded	Flooded
Constant	0.874***	0.896***	0.953***	0.566***	0.976***	0.945***	0.923***	0.316***
	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.005)	(0.012)
Band1	-0.551***	-0.537***	-0.459***		-0.249***	-0.541***	-0.759***	-0.309***
	(0.004)	(0.004)	(0.004)		(0.007)	(0.005)	(0.006)	(0.012)
Band2	-0.763***	-0.732***	-0.576***		-0.545***	-0.788***	-0.911***	-0.314***
	(0.003)	(0.003)	(0.004)		(0.008)	(0.004)	(0.005)	(0.012)
Band3	-0.797***	-0.768***	-0.581***		-0.676***	-0.843***	-0.917***	-0.316***
	(0.003)	(0.003)	(0.004)		(0.008)	(0.004)	(0.005)	(0.012)
Band4	-0.827***	-0.800***	-0.608***		-0.761***	-0.927***	-0.912***	-0.316***
	(0.003)	(0.003)	(0.004)		(0.007)	(0.003)	(0.005)	(0.012)
Elevation			-0.027***	-0.024***				
			(0.000)	(0.000)				
Elevation Sq			0.000***	0.000***				
			(0.000)	(0.000)				
BK		0.225***	0.155***	0.178***				
		(0.003)	(0.003)	(0.004)				
BX		-0.217***	-0.172***	-0.135***				
		(0.002)	(0.002)	(0.002)				
SI		-0.131***	-0.053***	-0.057***				
		(0.002)	(0.002)	(0.002)				
DistSFHA				-0.079***				
				(0.001)				
DistSFHA sq.				0.006***				
				(0.000)				
R-squared	0.319	0.427	0.493	0.346	0.244	0.399	0.491	0.274
N	147552	147552	147552	134865	21963	48903	58982	17704
Buildings	Fam14	Fam14	Fam14	Fam14	Fam14	Fam14	Fam14	Fam14
Boroughs	NYC4	NYC4	NYC4	NYC4 outFZ	BK	QN	SI	BX

Notes: The sample excludes apartment buildings (i.e. contains only building classes 1-family to 4-family homes). The dependent variable is a dummy variable taking a value of one when the building flooded during Sandy. The mean of the dependent variable (for the buildings in our area of interest) is 0.20 for NYC as a whole. Means by borough (for the area of interest) are: 0.43 (BK), 0.27 (QN), 0.09 (SI) and 0.03 (BX). We exclude the borough of Manhattan (because it only had 350 observations); omitted category in columns 2-4 is the Queens borough indicator (QN). Data includes properties in the SFHA (FZ100) or outside but within 1km (except in column 4 where FZ100 is removed). $DistSFHA$ stands for distance (in 100m) from the property to the SFHA and $DistSFHA$ sq. refers to its square. Bands partition the properties outside of SFHA as follows: band1 (<250m), band2 (250m-500m), band3 (500m-750m) and band4 (750m-1000m). *Elevation* is ground elevation in tens of meters. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 4: Surge depth. Only houses

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DepVar	Depth	Depth	Depth	Depth	Depth	Depth	Depth	Depth
Constant	3.952***	4.025***	4.168***	1.442***	3.893***	4.192***	5.109***	0.564***
	(0.020)	(0.019)	(0.019)	(0.012)	(0.030)	(0.021)	(0.050)	(0.028)
Band1	-3.160***	-3.127***	-2.937***		-2.078***	-3.124***	-4.788***	-0.555***
	(0.022)	(0.021)	(0.021)		(0.037)	(0.025)	(0.050)	(0.028)
Band2	-3.706***	-3.640***	-3.256***		-2.877***	-3.849***	-5.093***	-0.563***
	(0.021)	(0.021)	(0.020)		(0.038)	(0.023)	(0.050)	(0.028)
Band3	-3.812***	-3.749***	-3.288***		-3.258***	-4.046***	-5.098***	-0.564***
	(0.021)	(0.020)	(0.020)		(0.035)	(0.022)	(0.050)	(0.028)
Band4	-3.866***	-3.798***	-3.331***		-3.482***	-4.169***	-5.085***	-0.564***
	(0.020)	(0.020)	(0.020)		(0.034)	(0.022)	(0.050)	(0.028)
Elevation			-0.067***	-0.051***				
			(0.001)	(0.001)				
Elevation Sq			0.001***	0.001***				
			(0.000)	(0.000)				
BK		0.485***	0.313***	0.429***				
		(0.011)	(0.011)	(0.011)				
BX		-0.690***	-0.584***	-0.335***				
		(0.009)	(0.009)	(0.005)				
SI		-0.285***	-0.095***	-0.175***				
		(0.007)	(0.007)	(0.005)				
DistSFHA				-0.256***				
				(0.004)				
DistSFHA sq.				0.019***				
				(0.000)				
R-squared	0.470	0.516	0.544	0.272	0.321	0.566	0.679	0.196
N	147552	147552	147552	134865	21963	48903	58982	17704
Buildings	Fam14	Fam14	Fam14	Fam14	Fam14	Fam14	Fam14	Fam14
Boroughs	NYC4	NYC4	NYC4	NYC4 outFZ	BK	QN	SI	BX

Notes: The sample excludes apartment buildings (i.e. contains only building classes 1-family to 4-family homes). The dependent variable is the depth of flooding (in feet) based on Sandy's storm surge at the center of each parcel. The mean of the dependent variable (for the buildings in our area of interest) is 0.61ft for NYC as a whole. Means by borough (for the area of interest) are: 1.16ft (BK), 0.89ft (QN), 0.34ft (SI) and 0.07ft (BX). We exclude borough of Manhattan (because of small number of houses); omitted category is QN. Data includes properties in the SFHA (FZ100) or outside but within 1km (except in column 3 where SFHA is removed). $DistSFHA$ stands for distance from the property to the SFHA and $DistSFHA$ sq. is its square. Bands partition the properties outside of SFHA as follows: band1 ($<$250m), band2 (250m-500m), band3 (500m-750m) and band4 (750m-1000m). *Elevation* is ground elevation in tens of meters. Robust standard errors in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 5: Census tract panel data: summary statistics

	count	mean	std	min	max
NYC					
Buildings	9477	302.992	380.676	7	2,765
Sh SFHA	9477	0.048	0.168	0	1
Sh FZ500	9477	0.059	0.171	0	1
Sh Band1	9477	0.101	0.176	0	0.844
Sh Band2	9477	0.124	0.159	0	0.740
Sh Band3	9477	0.132	0.152	0	0.756
Sh Band4	9477	0.125	0.157	0	0.830
Sh Band5	9477	0.490	0.313	0	0.996
CT Elev	9477	14.084	12.123	1.319	82.395
nPol	9477	35.617	132.796	0	2,194
TR	9477	0.106	0.201	0	1.547
GFIP	9477	0.683	6.783	0	212
PRP	9477	16.348	63.497	0	1,401
Other	9477	18.587	91.234	0	1,713
Sh Flooded	9477	0.136	0.298	0	1
Flooding depth	9477	0.375	0.978	0	6.141

Notes: Observations are defined at the level of census tract by year. The data include all census tracts that overlap with New York City's SFHA (100-year flood zone) or the 1km-ring around it (i.e. 729 census tracts) over a 13-year period (2009-2022). Bands partition the properties outside of SFHA as follows: Band1 (<250m), Band2 (250m-500m), Band3 (500m-750m) and Band4 (750m-1,000m). Band 5 contains buildings located in the census tracts of interest but at more than 1,000m from the 100-year flood zone. $nPol$ is the number of annual policies purchased and TR is the take-up rate, defined as $nPol$ over the number of buildings in the census tract (at the time of Sandy). $ShFlooded$ is the share of buildings in the census tract that flooded during Sandy and $Floodingdepth$ is the average depth of flooding in the census tract (including houses with zero flooding).

Table 6: Flood insurance take-up before Sandy

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
DepVar	policyCount	policyCount	policyCount	policyCount	Takeup	Takeup	Takeup	Takeup
Year	2009	2010	2011	2012	2009	2010	2011	2012
Constant	648.29***	644.04***	670.52***	681.23***	1.05***	1.04***	1.07***	1.07***
	(147.46)	(144.65)	(152.60)	(156.61)	(0.09)	(0.09)	(0.09)	(0.09)
Sh Band1	-532.22***	-523.42***	-549.05***	-546.87***	-0.82***	-0.80***	-0.84***	-0.81***
	(180.65)	(177.31)	(187.00)	(192.31)	(0.12)	(0.11)	(0.12)	(0.11)
Sh Band2	-669.70***	-666.93***	-692.40***	-704.33***	-1.10***	-1.10***	-1.12***	-1.12***
	(134.74)	(132.29)	(139.52)	(143.29)	(0.09)	(0.08)	(0.09)	(0.09)
Sh Band3	-632.52***	-627.68***	-653.58***	-660.68***	-0.95***	-0.94***	-0.98***	-0.97***
	(152.47)	(149.58)	(157.79)	(161.98)	(0.10)	(0.09)	(0.10)	(0.10)
Sh Band4	-631.45***	-628.06***	-653.67***	-663.12***	-1.06***	-1.06***	-1.09***	-1.08***
	(145.59)	(142.83)	(150.67)	(154.62)	(0.09)	(0.09)	(0.09)	(0.09)
Sh Band5	-655.95***	-651.50***	-678.06***	-688.44***	-0.95***	-0.95***	-0.97***	-0.97***
	(149.18)	(146.34)	(154.37)	(158.43)	(0.09)	(0.09)	(0.09)	(0.09)
CT Elev	0.07	0.04	0.03	-0.02	-0.00***	-0.00***	-0.00***	-0.00***
	(0.07)	(0.07)	(0.07)	(0.08)	(0.00)	(0.00)	(0.00)	(0.00)
R-squared	0.56	0.56	0.56	0.55	0.61	0.59	0.61	0.62
N	729	729	729	729	729	729	729	729
Rep CT	28.651	28.432	29.458	30.880	0.084	0.083	0.086	0.090
	[2.555]	[2.543]	[2.668]	[2.796]	[0.004]	[0.004]	[0.004]	[0.004]
Maj SFHA	457.727	455.516	473.616	482.863	0.758	0.756	0.779	0.783
	[95.177]	[93.385]	[98.515]	[101.1]	[0.059]	[0.058]	[0.06]	[0.059]
Maj Band1	137.995	140.068	143.480	152.294	0.251	0.258	0.258	0.274
	[16.44]	[16.624]	[17.383]	[18.819]	[0.021]	[0.023]	[0.02]	[0.02]

Notes: Observations are defined at the level of census tract by year. The data include all census tracts that overlap with New York City's 100-year flood zone or the 1km-ring around it (i.e. 729 census tracts) over a 13-year period (2009-2022). The dependent variables are the number of new household-level policies purchased in the census tract in the corresponding year (columns 1-5) and the take-up rate (columns 6-7), where the denominator is the number of residential buildings (1-family to 4-family) in the census tract (prior to Sandy). The explanatory variables are the share of houses in the census tract located in band 1 ($shBand1$), in band 2 ($shBand2$), and so on. The omitted category is the share of houses in the census tract located in the 100-year flood zone. Bands partition the properties outside of the SFHA (FZ100) as follows: Band1 (<250m), Band2 (250m-500m), Band3 (500m-750m) and Band4 (750m-1,000m). Band 5 contains buildings located in the census tracts of interest but at more than 1,000m from the 100-year flood zone. The bottom panel reports the predicted value of the dependent variable for: (i) the average census tract (with at least one building in SFHA or band 1), (ii) the average majority-SFHA tract, (iii) the average majority-band-1 tract. The exact definitions can be found in the footnote to Table A2. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 7: Flood insurance take-up around Sandy (2010-2015 change)

2010-2015 DepVar	(1) ΔnPol	(2) ΔnPol	(3) ΔnPol	(4) ΔTR	(5) ΔTR	(6) ΔTR
Constant	144.21***	143.18***	48.19	0.22***	0.21***	-0.11*
	(40.02)	(40.09)	(48.19)	(0.05)	(0.05)	(0.06)
Sh Band1	-52.35	-50.33	3.91	-0.03	-0.02	0.17**
	(53.76)	(53.81)	(56.71)	(0.08)	(0.08)	(0.08)
Sh Band2	-150.88***	-143.22***	-56.51	-0.20***	-0.16***	0.14**
	(42.42)	(43.29)	(54.01)	(0.05)	(0.05)	(0.06)
Sh Band3	-123.68***	-115.96***	-36.28	-0.18***	-0.14**	0.13**
	(42.22)	(42.96)	(52.89)	(0.06)	(0.06)	(0.07)
Sh Band4	-140.60***	-131.12***	-53.99	-0.22***	-0.17***	0.09
	(39.72)	(39.82)	(44.99)	(0.06)	(0.06)	(0.06)
Sh Band5	-149.47***	-141.52***	-58.01	-0.20***	-0.16***	0.12**
	(40.40)	(40.60)	(47.26)	(0.05)	(0.05)	(0.06)
CT Elev		-0.44***	0.07		-0.00***	-0.00**
		(0.14)	(0.05)		(0.00)	(0.00)
Sh Flooded			77.12***			0.26***
			(21.50)			(0.03)
R-squared	0.16	0.16	0.24	0.12	0.16	0.40
N	729	729	729	729	729	729
Rep CT	13.070	13.106	13.040	0.041	0.042	0.041
	[2.172]	[2.171]	[2.049]	[0.004]	[0.004]	[0.004]
Maj SFHA	111.855	111.329	112.365	0.181	0.178	0.182
	[26.422]	[26.463]	[26.333]	[0.031]	[0.031]	[0.027]
Maj Band1	71.626	71.364	73.476	0.149	0.148	0.155
	[15.264]	[15.202]	[14.238]	[0.027]	[0.026]	[0.021]

Notes: Observations are defined at the level of census tract by year. The data include all census tracts that overlap with New York City's 100-year flood zone or the 1km-ring around it (i.e. 729 census tracts) over a 13-year period (2009-2022). In columns 1-3, the dependent variable is the change in the number of new household-level insurance policies (ΔnPol). In columns 4-6, the dependent variable is the change in the census tract take-up rate between 2010 and 2015 (ΔTR). The explanatory variables are the share of houses in the census tract located in band 1 ($shBand1$), in band 2 ($shBand2$), and so on. The omitted category is the share of houses in the census tract located in the 100-year flood zone. Bands partition the properties outside of the SFHA (FZ100) as follows: Band1 (<250m), Band2 (250m-500m), Band3 (500m-750m) and Band4 (750m-1,000m). Band 5 contains buildings located in the census tracts of interest but at more than 1,000m from the 100-year flood zone. The bottom panel reports the predicted value of the dependent variable for: (i) the average census tract (with at least one building in SFHA or band 1), (ii) the average majority-SFHA tract, (iii) the average majority-band-1 tract. The exact definitions can be found in the footnote to Table A2. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 8: Event Study: Annual Insurance Purchases

Dep. Variable	(1) policyCount	(2) policyCount	(3) policyCount	(4) policyCount	(5) TR	(6) policyCount	(7) TR
Constant	28.944***	9.0729***	19.557***	3.0292	0.0871***	0.7759	0.0822***
	(6.6955)	(2.8630)	(7.3101)	(2.4774)	(0.0125)	(2.1444)	(0.0122)
T0	2.0256***	1.0808**	1.0808**	-0.1736	0.0030**	0.2369	0.0039***
	(0.5870)	(0.4659)	(0.4659)	(0.2301)	(0.0014)	(0.2280)	(0.0015)
T1	17.294***	7.9600**	7.9600**	-0.3492	0.0185***	2.5558	0.0250***
	(4.7159)	(3.2826)	(3.2827)	(1.3648)	(0.0062)	(1.6128)	(0.0075)
T2	12.665***	7.9030**	7.9030**	0.3422	0.0197***	2.8565*	0.0256***
	(3.7480)	(3.1199)	(3.1201)	(1.3032)	(0.0065)	(1.6578)	(0.0079)
T3	12.639***	7.7803**	7.7803**	-0.2186	0.0158**	2.4211	0.0222***
	(3.8492)	(3.2305)	(3.2306)	(1.4030)	(0.0069)	(1.7378)	(0.0085)
T4_9	7.0206**	5.4262*	5.4262*	-0.5147	0.0102	1.6761	0.0151**
	(3.1685)	(2.8455)	(2.8457)	(1.2858)	(0.0064)	(1.5870)	(0.0077)
shSFHA		703.97***	686.59***	652.72***	0.9976***	638.33***	0.9807***
		(191.85)	(194.68)	(197.50)	(0.1061)	(193.42)	(0.1024)
T0 × shSFHA		33.469***	33.469***	27.388**	0.0183	22.627*	0.0079
		(12.284)	(12.284)	(12.558)	(0.0133)	(11.872)	(0.0125)
T1 × shSFHA		330.69***	330.69***	290.40***	0.4567***	256.71***	0.3823***
		(81.085)	(81.089)	(80.463)	(0.0664)	(74.191)	(0.0738)
T2 × shSFHA		168.69***	168.69***	132.04***	0.1911***	102.87***	0.1224*
		(38.056)	(38.058)	(37.011)	(0.0606)	(34.114)	(0.0704)
T3 × shSFHA		172.12***	172.12***	133.34***	0.1866***	102.72***	0.1129
		(40.949)	(40.951)	(41.230)	(0.0677)	(38.700)	(0.0771)
T4_9 × shSFHA		56.481	56.481	27.679	-0.0480	2.2650	-0.1051*
		(36.257)	(36.259)	(37.945)	(0.0565)	(39.034)	(0.0637)
CT Elev			-0.7095**	-0.2283	-0.0029***	0.0127	-0.0025***
			(0.3332)	(0.1601)	(0.0006)	(0.1125)	(0.0005)
shBand1				106.45***	0.1022***	49.014	0.0280
				(39.485)	(0.0375)	(32.311)	(0.0372)
T0 × shBand1				14.181**	0.0227**	-1.2697	-0.0108
				(5.7027)	(0.0106)	(3.1775)	(0.0095)
T1 × shBand1				93.932***	0.1691***	-15.419	-0.0724***
				(34.121)	(0.0575)	(18.218)	(0.0246)
T2 × shBand1				85.473***	0.1582***	-9.1730	-0.0646**
				(29.918)	(0.0583)	(13.023)	(0.0257)
T3 × shBand1				90.424***	0.1855***	-8.9393	-0.0537**
				(30.885)	(0.0664)	(13.658)	(0.0268)
T4_9 × shBand1				67.161**	0.1483***	-15.309	-0.0370
				(27.738)	(0.0538)	(11.524)	(0.0232)
shBand1Flooded						157.38	0.2079**
						(105.13)	(0.1040)
T0 × shBand1Flooded						39.887***	0.0865***
						(13.583)	(0.0199)
T1 × shBand1Flooded						282.30***	0.6234***
						(97.071)	(0.1145)
T2 × shBand1Flooded						244.33***	0.5753***
						(80.954)	(0.1296)
T3 × shBand1Flooded						256.51***	0.6175***
						(77.026)	(0.1524)
T4_9 × shBand1Flooded						212.90***	0.4782***
						(68.235)	(0.1300)
No. Observations	9477	9477	9477	9477	9477	9477	9477
R-squared	0.0015	0.4261	0.4301	0.4718	0.5213	0.5158	0.5873

Notes: Observations are defined at the level of census tract by year. The data include all census tracts that overlap with the SFHA or the 1km-ring around it (i.e. 729 census tracts) over a 13-year period (2009-2021). The dependent variables are the number of new household-level policies purchased in the census tract in the corresponding year (columns 1-4 and 6-7) and the take-up rate (column 5). $T0$ denotes 2012, $T1$ denotes 2013, and so on. Dummy variable $T4\_9$ denotes the period 2016-2021. Bands partition the properties outside of the SFHA as follows: Band1 (<250m), Band2 (250m-500m), Band3 (500m-750m) and Band4 (750m-1,000m). Band 5 contains buildings located in the census tracts of interest but at more than 1,000m from the 100-year flood zone. Variables $shSFHA$ and $shBand1$ is the share of the houses in the census tract located in SFHA and in band 1, respectively. Similarly, $shBand1Flooded$ is the share of band 1 houses in the census tract that flooded during Sandy. Standard errors clustered by zip code in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 9: Event Study: Annual Insurance Purchases by Insurance Type

Dep. Variable	(1) policyCount	(2) GFIP	(3) PRP	(4) Other
Constant	3.0292	0.0049	7.5449***	-4.5206***
	(2.4774)	(0.0263)	(2.3250)	(1.3317)
T0	-0.1736	-0.0522	-0.0302	-0.0912
	(0.2301)	(0.0342)	(0.2060)	(0.0937)
T1	-0.3492	-0.3267**	0.6435	-0.6660
	(1.3648)	(0.1619)	(1.2851)	(0.4133)
T2	0.3422	-0.2213*	0.4827	0.0808
	(1.3032)	(0.1150)	(1.3128)	(0.3212)
T3	-0.2186	-0.2128**	-0.1073	0.1015
	(1.4030)	(0.1035)	(1.3963)	(0.3669)
T4_9	-0.5147	0.0031	-0.7901	0.2723
	(1.2858)	(0.0049)	(1.2526)	(0.4151)
shSFHA	652.72***	0.8916**	56.917	594.91***
	(197.50)	(0.3480)	(41.542)	(163.54)
T0 ×shSFHA	27.388**	8.5330***	1.7107	17.144*
	(12.558)	(2.9823)	(4.3643)	(9.3620)
T1 ×shSFHA	290.40***	122.99***	17.557	149.85***
	(80.463)	(23.197)	(14.302)	(57.558)
T2 ×shSFHA	132.04***	87.963***	14.052	30.025
	(37.011)	(16.936)	(14.474)	(30.127)
T3 ×shSFHA	133.34***	84.246***	22.190	26.903
	(41.230)	(16.262)	(16.522)	(31.709)
T4_9 ×shSFHA	27.679	-0.9006***	16.188	12.391
	(37.945)	(0.3369)	(13.947)	(35.245)
CT Elev	-0.2283	-0.0005	-0.3246**	0.0968**
	(0.1601)	(0.0016)	(0.1451)	(0.0433)
shBand1	106.45***	0.0389	58.417**	47.997**
	(39.485)	(0.0970)	(23.625)	(22.971)
T0 × shBand1	14.181**	0.8704	13.788***	-0.4774
	(5.7027)	(0.5815)	(3.9508)	(2.4174)
T1 × shBand1	93.932***	3.5273	86.802***	3.6031
	(34.121)	(3.5545)	(25.333)	(11.355)
T2 × shBand1	85.473***	2.4491	87.762***	-4.7387
	(29.918)	(2.5062)	(24.911)	(7.4536)
T3 × shBand1	90.424***	2.3440	92.899***	-4.8186
	(30.885)	(2.3637)	(25.939)	(7.9633)
T4_9 × shBand1	67.161**	-0.0461	79.186***	-11.978
	(27.738)	(0.0978)	(23.643)	(8.6014)
No. Observations	9477	9477	9477	9477
R-squared	0.4718	0.6625	0.1717	0.5878

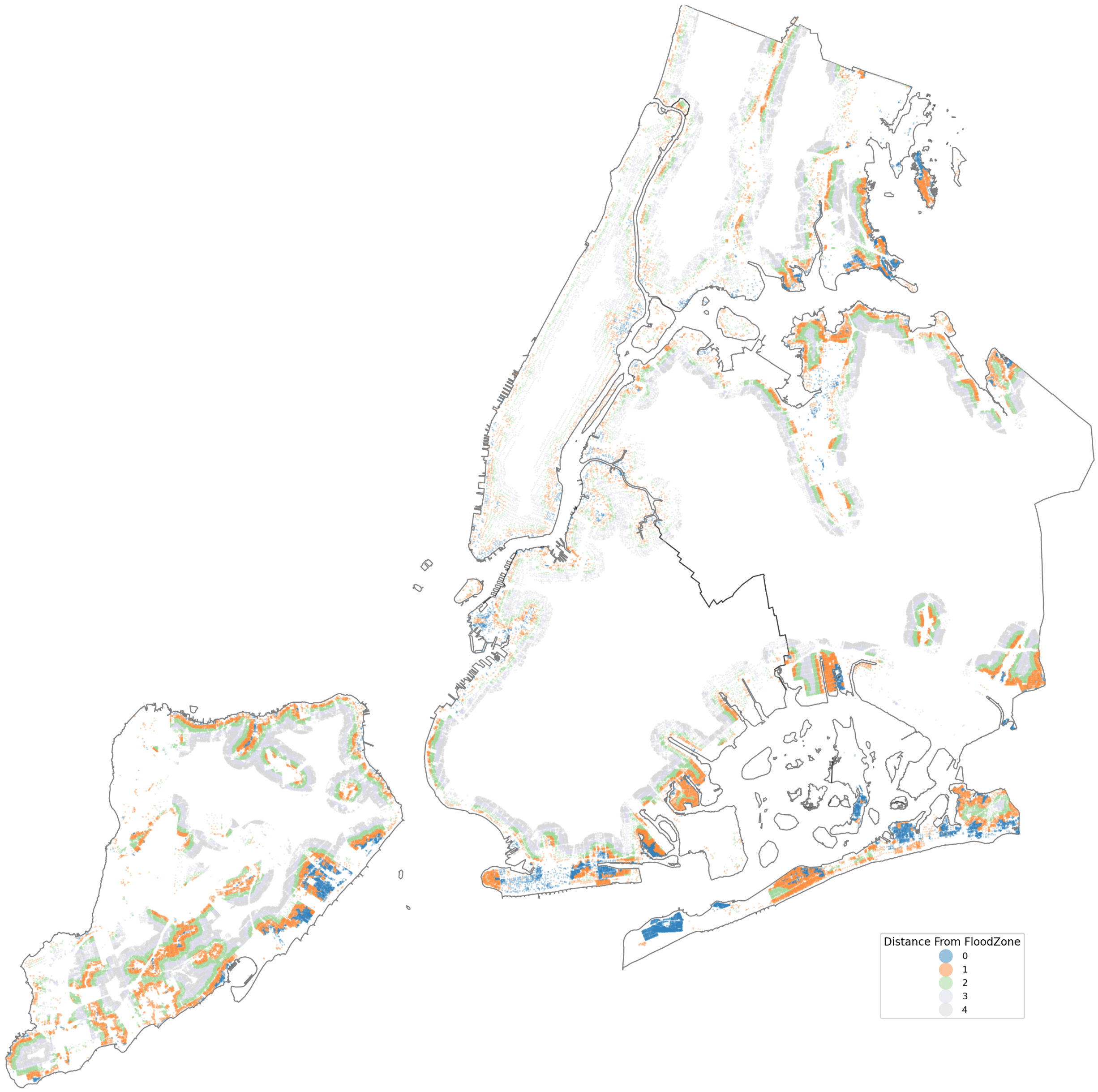
Notes: Observations are defined at the level of census tract by year. The data include all census tracts that overlap with the SFHA or the 1km-ring around it (i.e. 729 census tracts) over a 13-year period (2009-2021). The dependent variables are the number of new household-level policies purchased in the census tract in the corresponding year: column 1 refers to the overall policy count, columns 2-4 report GFIP policies, PRP policies and Other policies, respectively. The interpretation of the coefficients and the definition of the periphery bands is analogous to column 4 in Table 8. Variables $shSFHA$ and $shBand1$ is the share of the houses in the census tract located in SFHA and in band 1, respectively. Standard errors clustered by zip code in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table 10: Horse race between share of houses flooded by Sandy and share in band 1 reclassified into SFHA in 2013 preliminary flood maps

Dep. Variable	(1) policyCount	(2) Takeup	(3) policyCount	(4) Takeup
Constant	0.7759	0.0822***	0.9369	0.0824***
	(2.1444)	(0.0122)	(2.1536)	(0.0122)
T0	0.2369	0.0039***	0.2384	0.0040***
	(0.2280)	(0.0015)	(0.2245)	(0.0015)
T1	2.5558	0.0250***	2.5293	0.0250***
	(1.6128)	(0.0075)	(1.5722)	(0.0075)
T2	2.8565*	0.0256***	2.8168*	0.0256***
	(1.6578)	(0.0079)	(1.6151)	(0.0080)
T3	2.4211	0.0222***	2.3765	0.0222***
	(1.7378)	(0.0085)	(1.6946)	(0.0085)
T4_9	1.6761	0.0151**	1.6270	0.0152**
	(1.5870)	(0.0077)	(1.5469)	(0.0077)
shBand1	49.014	0.0280	51.015	0.0379
	(32.311)	(0.0372)	(31.526)	(0.0373)
T0 × shBand1	-1.2697	-0.0108	-1.3952	-0.0139
	(3.1775)	(0.0095)	(2.8890)	(0.0097)
T1 × shBand1	-15.419	-0.0724***	-13.279	-0.0787***
	(18.218)	(0.0246)	(16.279)	(0.0257)
T2 × shBand1	-9.1730	-0.0646**	-5.9656	-0.0700**
	(13.023)	(0.0257)	(10.822)	(0.0274)
T3 × shBand1	-8.9393	-0.0537**	-5.3424	-0.0597**
	(13.658)	(0.0268)	(11.723)	(0.0296)
T4_9 × shBand1	-15.309	-0.0370	-11.344	-0.0412
	(11.524)	(0.0232)	(10.107)	(0.0251)
shBand1Flooded	157.38	0.2079**	81.979	-0.1492
	(105.13)	(0.1040)	(218.90)	(0.2701)
T0 × shBand1Flooded	39.887***	0.0865***	44.372**	0.1973***
	(13.583)	(0.0199)	(22.038)	(0.0633)
T1 × shBand1Flooded	282.30***	0.6234***	205.89	0.8493***
	(97.071)	(0.1145)	(126.41)	(0.2716)
T2 × shBand1Flooded	244.33***	0.5753***	129.78	0.7679***
	(80.954)	(0.1296)	(99.672)	(0.2958)
T3 × shBand1Flooded	256.51***	0.6175***	128.05	0.8313**
	(77.026)	(0.1524)	(109.45)	(0.3620)
T4_9 × shBand1Flooded	212.90***	0.4782***	71.311	0.6295**
	(68.235)	(0.1300)	(122.65)	(0.3178)
shBand1Rezone			81.786	0.3879
			(281.70)	(0.3021)
T0 × shBand1Rezone			-4.8756	-0.1205*
			(29.673)	(0.0728)
T1 × shBand1Rezone			83.069	-0.2456
			(195.15)	(0.2822)
T2 × shBand1Rezone			124.53	-0.2094
			(155.97)	(0.2804)
T3 × shBand1Rezone			139.66	-0.2325
			(156.19)	(0.3202)
T4_9 × shBand1Rezone			153.93	-0.1645
			(153.58)	(0.2850)
No. Observations	9477	9477	9477	9477
R-squared	0.5158	0.5873	0.5168	0.5881

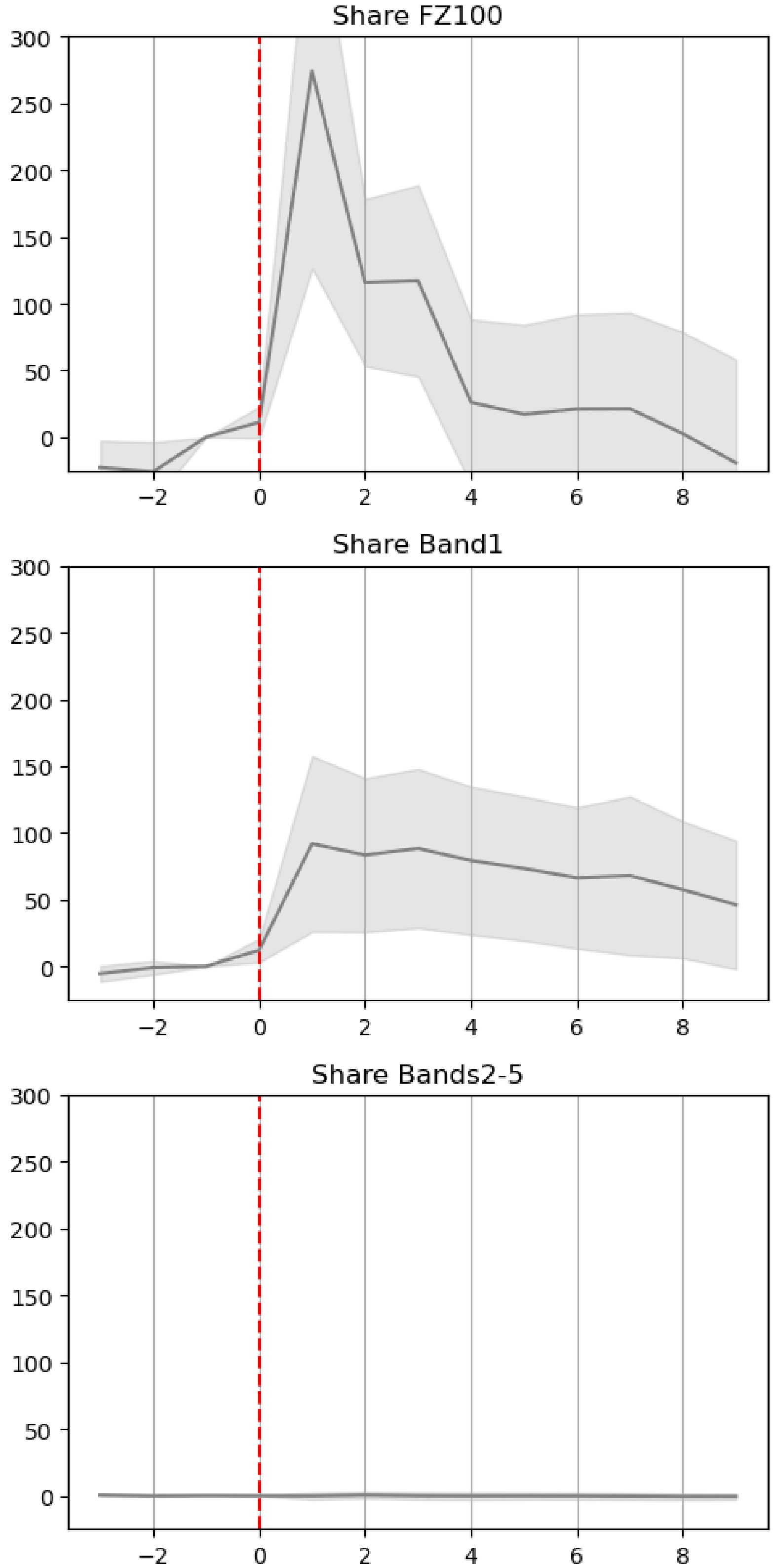
Notes: Observations are defined at the level of census tract by year. The data include all census tracts that overlap with New York City's 100-year flood zone or the 1km-ring around it (i.e. 729 census tracts) over a 13-year period (2009-2021). The variables are defined in Table 9. Columns 1-2 include interaction terms between period dummies and the share of buildings in band 1 that flooded during Sandy ($shBand1Flooded$). Columns 3-4 also include interaction terms between period dummies and the share of buildings in band 1 that were rezoned into the SFHA in the new (preliminary) flood maps ($shBand1Rezone$). Standard errors clustered by zip code in brackets. *** p<0.01, ** p<0.05, * p<0.1

Figure 1: NYC: 100-year flood zone and periphery bands



Note: Band 0 in the legend of the figure refers to the SFHA. Bands 1-4 partition the properties outside of the SFHA as follows: Band1 (<250m), Band2 (250m-500m), Band3 (500m-750m) and Band4 (750m-1,000m).

Figure 2: Dynamic take-up response to Sandy



Note: Estimates based on Equation 6 and reported in Table 8 (extended to allow for annual coefficients). The dependent variable is the count of new policies in the corresponding year. Year 0 refers to 2013 (and recall that hurricane Sandy reached NYC on October 29, 2012). Hurricane Irene reached NYC in year 2011 (i.e. $t = -1$). $FZ100$ stands for the 100-year flood zone, also referred to as the SFHA. Standard errors clustered by zip code.

Appendix

Appendix A: Representative census tracts

To get a better sense of the levels and changes in flood insurance coverage implied by the estimates in Table 6 and Table 7, it is helpful to construct representative census tracts.

We restrict the analysis to houses (family 1 to family 4) and to census tracts that overlap with the Special Flood Hazard Area (SFHA or 100-year flood zone) or periphery band 1. We partition these tracts in two groups: tracts with buildings primarily located in SFHA and tracts with buildings mostly in band 1. For each of these sets, we compute the average shares of buildings in SFHA, band 1, band 2, and so on, until band 5.

More specifically, we define a census tract c as being *majority SFHA* if

$$shSFHA_c > shBand1_c - (shBand2_c + ... + shBand5_c). \tag{7}$$

Note that if the census tracts we are considering were fully contained within the union of SFHA and band 1, the terms in parenthesis would be zero. Likewise, we define a census tract c as being *majority band 1* if

$$shBand1_c > shSFHA_c - (shBand2_c + ... + shBand5_c). \tag{8}$$

Table A2 reports the composition of each of the representative census tracts in terms of the shares of buildings located in SFHA and the 5 periphery bands. To set the stage, column 1 reports the composition of the average census tract overlapping with SFHA or periphery band 1. This representative tract has 6.4% of its buildings in SFHA, 22.6% in band 1, 20.9% in band 2, 11.6% in band 3, 5.5% in band 4, and 33% in band 5.

Column 2 displays the composition of the average *majority-SFHA* census tract. Not surprisingly, the majority of the buildings in this tract are located in SFHA (67.7%). In addition, 16.1% are in band 1 and the remaining 15.3% are in band 5. The table also reports the average elevation of the buildings in this type of census tract, which is less than 2.5m above sea level, and the share of buildings that flooded during Sandy (94.2%).

Let us now turn to the composition of the average *majority-band 1* census tract. The representative tract within this type only has 9.5% of its buildings in SFHA and 63.3% in band 1. In addition, it has 13.3% in band 2, slightly below 1% in band 3, and 13.1% in band 5. Note also that the average elevation of the buildings in this type of tracts is 5.8m. Partly because of the higher elevation and partly due to the greater distance from the ocean, it is not surprising that the share of buildings that flooded during Sandy is much lower (48.6%) than for the majority-SFHA tracts.

Appendix B: Additional Tables

Table A1: Flooding and distance to SFHA. NYC5, all building types

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
DepVar	Flooded	Flooded	Flooded	Flooded	Flooded	Flooded	Flooded	Flooded	Flooded
Constant	0.849***	0.859***	0.915***	0.524***	0.599***	0.940***	0.926***	0.895***	0.370***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.026)	(0.004)	(0.003)	(0.005)	(0.011)
Band1	-0.547***	-0.529***	-0.462***		-0.467***	-0.288***	-0.540***	-0.733***	-0.362***
	(0.004)	(0.003)	(0.003)		(0.028)	(0.007)	(0.005)	(0.006)	(0.011)
Band2	-0.745***	-0.713***	-0.572***		-0.591***	-0.565***	-0.775***	-0.882***	-0.368***
	(0.003)	(0.003)	(0.003)		(0.026)	(0.008)	(0.004)	(0.005)	(0.011)
Band3	-0.778***	-0.748***	-0.576***		-0.598***	-0.676***	-0.832***	-0.888***	-0.369***
	(0.003)	(0.003)	(0.003)		(0.026)	(0.007)	(0.004)	(0.005)	(0.011)
Band4	-0.805***	-0.776***	-0.597***		-0.599***	-0.747***	-0.909***	-0.884***	-0.370***
	(0.003)	(0.003)	(0.003)		(0.026)	(0.007)	(0.003)	(0.005)	(0.011)
Elevation			-0.025***	-0.022***					
			(0.000)	(0.000)					
Elevation Sq			0.000***	0.000***					
			(0.000)	(0.000)					
BK		0.198***	0.138***	0.160***					
		(0.003)	(0.003)	(0.003)					
BX		-0.191***	-0.144***	-0.112***					
		(0.002)	(0.002)	(0.002)					
SI		-0.113***	-0.043***	-0.046***					
		(0.002)	(0.002)	(0.002)					
DistSFHA				-0.077***					
				(0.001)					
DistSFHA sq.				0.006***					
				(0.000)					
R-squared	0.314	0.405	0.464	0.310	0.328	0.235	0.388	0.465	0.321
N	175769	175769	175769	160226	4338	28567	54767	65005	23092
Buildings	All	All	All	All	All	All	All	All	All
Boroughs	NYC5	NYC5	NYC5	NYC5 outFZ	MN	BK	QN	SI	BX

Notes: All 5 boroughs included and all buildings. Data includes properties in SFHA or outside but within 1km (except in column 3 where SFHA is removed). *DistSFHA* stands for distance from the property to the SFHA. Bands partition the properties outside of SFHA as follows: band1 (<250m), band2 (250m-500m), band3 (500m-750m) and band4 (750m-1000m). *Elevation* is ground elevation in tens of meters. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table A2: Composition representative census tracts

	Overlapping SFHA or band1	Majority SFHA	Majority Band1
sh SFHA	0.064	0.677	0.095
sh Band1	0.226	0.161	0.633
sh Band2	0.209	0.008	0.133
sh Band3	0.116	0.000	0.008
sh Band4	0.055	0.000	0.000
sh Band5	0.330	0.153	0.131
CT Elev	10.732	2.466	5.787
Sh Flooded	0.240	0.942	0.486

Notes: We restrict the analysis to houses (family 1 to family 4) and to census tracts that overlap with the 100-year flood zone (SFHA or FZ100) or periphery band 1. Column 1 reports the average shares of buildings in SFHA and each of the 5 periphery bands for these tracts. Column 2 (majority SFHA tracts) reports the shares for tracts with $shSFHA$ greater than $shBand1$ minus $(shBand2 + ... + shBand5)$. Similarly, column 3 (majority band 1 tracts) reports the shares for tracts with $shBand1$ greater than $shSFHA$ minus $(shBand2 + ... + shBand5)$. The table also reports the average elevation of the buildings in each type of census tract and the average share of buildings that flooded during Sandy. Bands partition the properties outside of SFHA as follows: Band1 (<250m), Band2 (250m-500m), Band3 (500m-750m) and Band4 (750m-1,000m). Band 5 contains buildings located in the census tracts of interest but at more than 1,000m from the 100-year flood zone. CT Elev is the average elevation in the census tract.

Table A3: Event Study: New annual insurance policies (by type), excluding renewals

Dep. Variable	(1) policyCount	(2) GFIP	(3) PRP	(4) Other
Constant	1.2417***	0.0110	1.6644***	-0.4337***
	(0.4083)	(0.0186)	(0.3890)	(0.1545)
T0	-0.1573	-0.0417*	-0.1253	0.0097
	(0.1537)	(0.0245)	(0.1586)	(0.0421)
T1	0.0107	-0.0957	0.4338	-0.3275*
	(1.2784)	(0.2192)	(1.2498)	(0.1906)
T2	-0.0453	0.0036	0.0570	-0.1059
	(0.4078)	(0.0053)	(0.3743)	(0.1093)
T3	-0.4062	0.0027	-0.4380	0.0292
	(0.3006)	(0.0057)	(0.3004)	(0.0645)
T4_9	-0.4563***	0.0044	-0.5662***	0.1055**
	(0.1002)	(0.0053)	(0.1055)	(0.0534)
shSFHA	51.780***	0.8145**	4.1687	46.797***
	(16.045)	(0.3250)	(3.5295)	(13.248)
T0 × shSFHA	-0.8099	6.0269***	0.6355	-7.4723**
	(3.2931)	(1.8881)	(1.6756)	(2.9813)
T1 × shSFHA	207.06***	120.29***	17.040	69.730***
	(44.681)	(21.111)	(11.265)	(22.700)
T2 × shSFHA	58.134**	-0.2107	6.5783	51.766**
	(22.654)	(0.1937)	(4.1131)	(20.422)
T3 × shSFHA	11.198	-0.7878**	5.9009**	6.0844
	(9.0460)	(0.3261)	(2.7592)	(7.6244)
T4_9 × shSFHA	-7.2788**	-0.8324***	-0.4104	-6.0359*
	(3.6224)	(0.3151)	(1.0697)	(3.4770)
CT Elev	-0.0522**	-0.0010	-0.0588**	0.0076
	(0.0264)	(0.0012)	(0.0247)	(0.0047)
shBand1	14.812***	0.0466	7.8199***	6.9457***
	(3.6650)	(0.1012)	(2.1077)	(1.9445)
T0 × shBand1	9.0335***	0.7244*	8.9051***	-0.5960
	(3.0553)	(0.3838)	(2.4577)	(0.8143)
T1 × shBand1	82.824***	1.1070	78.084***	3.6326
	(28.116)	(3.2643)	(23.481)	(5.3663)
T2 × shBand1	22.796***	-0.1016	22.990***	-0.0926
	(8.0040)	(0.1039)	(5.9154)	(4.0764)
T3 × shBand1	10.018**	-0.0285	12.496***	-2.4501
	(3.9583)	(0.1079)	(3.2362)	(1.6541)
T4_9 × shBand1	-3.5065**	-0.0606	0.7468	-4.1927***
	(1.5564)	(0.1029)	(0.8832)	(1.0326)
No. Observations	9204	9204	9204	9204
R-squared	0.3638	0.6848	0.1381	0.5295

Notes: Observations are defined at the level of census tract by year. The data include all census tracts that overlap with New York City's 100-year flood zone or the 1km-ring around it (i.e. 729 census tracts) over a 13-year period (2009-2022). The dependent variables are the number of new household-level policies purchased in the census tract in the corresponding year: column 1 refers to the overall policy count, columns 2-4 report GFIP policies, PRP policies and Other policies, respectively. The interpretation of the coefficients and the definition of the periphery bands is analogous to column 4 in Table 8. Respectively, $shSFHA$ and $shBand1$ are the share of buildings in the census tract that are located within the SFHA or Band1. Standard errors clustered by zip code in brackets. *** p<0.01, ** p<0.05, * p<0.1

Table A4: Event study: insurance take up by cohort before and after the release of the 2013 preliminary flood maps

	(1)	(2)	(3)	(4)
Dep. Variable	policyCount	Takeup	policyCount	Takeup
Cohort	Pre-Map	Pre-Map	Post-Map	Post-Map
Constant	4.3561**	0.0375***	3.4225	0.0403***
	(2.1615)	(0.0123)	(2.0884)	(0.0094)
T2	-0.5313*	-0.0085***	-0.4268	-0.0129***
	(0.3109)	(0.0027)	(0.4141)	(0.0029)
T3	-0.7528**	-0.0097***	-0.5764	-0.0151***
	(0.3705)	(0.0035)	(0.4758)	(0.0034)
T4_9	-1.0593*	-0.0148***	-0.7070	-0.0219***
	(0.5622)	(0.0050)	(0.7053)	(0.0051)
T1 × shSFHA	35.061**	0.0303	79.712***	0.1845**
	(16.217)	(0.0353)	(22.776)	(0.0901)
T2 × shSFHA	17.062*	0.0032	37.107***	0.0787*
	(9.1043)	(0.0247)	(11.750)	(0.0450)
T3 × shSFHA	16.888*	0.0041	35.771***	0.0802*
	(8.9259)	(0.0235)	(10.986)	(0.0438)
T4_9 × shSFHA	5.5332	-0.0167	4.8744	-0.0083
	(5.7546)	(0.0153)	(4.1225)	(0.0097)
CT Elev	-0.1903*	-0.0009**	-0.2027*	-0.0009***
	(0.1040)	(0.0004)	(0.1086)	(0.0003)
T1 × shBand1	41.443***	0.0408***	35.279***	0.0008
	(11.896)	(0.0145)	(13.285)	(0.0326)
T2 × shBand1	30.868***	0.0341***	23.842***	0.0077
	(8.9633)	(0.0109)	(8.9433)	(0.0176)
T3 × shBand1	29.076***	0.0307***	21.974***	0.0056
	(8.3234)	(0.0106)	(8.2993)	(0.0168)
T4_9 × shBand1	22.026***	0.0231***	14.981**	0.0106*
	(6.7427)	(0.0079)	(5.8180)	(0.0062)
No. Observations	3015	3015	2628	2628
R-squared	0.1084	0.1228	0.1893	0.2155

Notes: Observations are defined at the level of census tract by year. The data include all census tracts that overlap with the SFHA or the 1km-ring around it (i.e. 729 census tracts) over the period 2013-2021. Thus dummy variable $T1$ refers to year 2013 and it is omitted to avoid perfect collinearity with the intercept. The new preliminary ABFE (base flood elevation) maps were released piecemeal during the first half of 2013. Columns 1-2, labelled *pre-map*, track renewals of the cohort of customers who originally purchased a new insurance policy between 12/1/2012 and 2/28/2013. Columns 3-4, labelled *post-map*, track renewals of the cohort of customers who originally purchased a new insurance policy between 3/1/2013 and 6/1/2013. The dependent variables in columns 1 and 3 correspond to policy counts and in columns 2 and 4 to take-up rates. The interpretation of the coefficients and the definition of the periphery bands is analogous to column 4 in Table 8. Standard errors clustered by zip code in brackets. *** p<0.01, ** p<0.05, * p<0.1