Sorting Over Flood Risk and Implications for Policy Reform

Laura A. Bakkensen Lala Ma*

September 2019

Do individuals sort across flood risk? This paper applies a boundary discontinuity design to a residential sorting model to provide novel estimates of heterogeneous sorting across flood risk by race, ethnicity, and income. We find clear evidence that low income and minority residents are more likely to move into high risk flood zones. We then highlight the overall and distributional implications of proposed price and information reforms to the U.S. National Flood Insurance Program. While such reforms are likely welfare increasing overall, heterogeneous behavioral responses yield significant distributive effects that also alter the composition of residents in harm's way.

*Bakkensen: University of Arizona, laurabakkensen@email.arizona.edu. Ma: University of Kentucky, lala.ma@uky.edu. We thank Glenn Blomquist, Trudy Ann Cameron, Tatyana Deryugina, Brian Gerber, Caroline Hopkins, Ben Miller, Daniel Phaneuf, Kathy Segerson, Leonard Shabman, Chris Timmins, Crystal Zhan for their insightful feedback and suggestions. Research conducted in this article by Laura Bakkensen was supported by an Early-Career Research Fellowship from the Gulf Research Program of the National Academies of Sciences, Engineering, and Medicine. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Gulf Research Program of the National Academies of Sciences, Engineering, and Medicine.

1 Introduction

Scholars and policy makers have long been concerned with how individuals locate relative to environmental hazards such as natural disasters. Residential patterns surrounding disaster risk can have critical consequences for many economic outcomes including household finance, economic growth, and migration (Strobl, 2011; Hornbeck, 2012; Cavallo et al., 2013; Gallagher and Hartley, 2017) as well as public programs for emergency management, welfare, and insurance (Michel-Kerjan, 2010; Deryugina, 2017). While a rich literature has estimated household preferences to avoid such risks, a longstanding challenge to empirical identification is the correlation between disaster risk and spatial amenities. In addition, an open question surrounds the potential for heterogeneous sorting that, if present, can lead to unintended and unwanted distributional consequences including from benevolently-intentioned public policies.

This paper provides novel empirical evidence on sorting across disaster risk and highlights implications for policy reform. Using the case of flood risk in South Florida, we first estimate a discrete choice residential sorting model with three innovations on the flood risk literature. Compared with the hedonic price model predominately employed in existing studies, our approach (i) allows for heterogeneous sorting over flood risk by homebuyer race, ethnicity, and income, (ii) accounts for property-specific insurance pricing that could otherwise confound analysis, and (iii) employs a boundary discontinuity identification strategy (Black, 1999) within our sorting model (Bayer et al., 2007) to deal with the endogeneity of disaster risk and spatial attributes. Our results provide the first estimates of heterogeneous sorting over flood risk.

Second, we investigate the potential consequences of heterogeneous sorting for policy reform. We estimate the welfare and distributional consequences of changes in prices and flood risk information faced by households under the U.S. National Flood Insurance Program. In particular, using the structural parameters from our sorting model, we estimate the compensating variation for different race, ethnicity, and income groups from a (counterfactual) removal of the program's three largest insurance price discount schemes, and predict the resulting reallocation of household types across flood risk zones. In addition, we assess the value of risk information using new flood risk maps released by the National Flood Insurance Program. We then compare these benefits to the costs of map revisions.

We find clear evidence that individuals avoid flood risk, as homes located just inside a high risk flood zone sell at a 6.3% discount relative to those just outside. Ignoring correlated amenities and insurance price discounts implies that high risk homes sell at a premium. Second, low income and minority residents are more likely to sort into high flood risk areas. This sorting takes place, even though high income white residents tend to be concentrated in high risk coastal zones, likely driven by the amenity value associated with flood risk (e.g., Kahn and Smith (2017)).[1] In addition to furthering our understanding of residential location choice around environmental risk, the presence of heterogeneous sorting reaffirms the established result that housing price capitalization effects, estimated from hedonic price models, should be interpreted with care as they may combine preferences to avoid flood risk with changes in the implicit prices of flood risk and other co-existing amenities due to sorting (Kuminoff et al., 2010; Kuminoff and Pope, 2014; Bakkensen and Barrage, 2017).

Policy changes can also have important distributional consequences in the presence of heterogeneous sorting. In our setting, the costs of insurance price reform fall more heavily on low income residents as a fraction of income. Resulting re-sorting would then lead to a greater concentration of low income and minority residents in harm's way. While policy reform may well be a desirable goal, these distributional impacts could have potentially long lasting implications for disaster vulnerability, recovery, and fiscal policy (Arrow et al., 1996; Robinson et al., 2016; Banzhaf et al., 2019).

Despite distributional costs, society may still realize large efficiency gains from reforms overall. We find that household welfare costs from insurance price reforms are significantly lower relative to costs estimated from an analysis that assumes no re-sorting, with expected welfare loss experienced by these households to be, on average, only 24 percent of the price

[1] We note, but cannot tease apart, mechanisms that could give rise to this heterogeneity including, e.g., tastes (Banzhaf and Walsh, 2008), beliefs (Bakkensen and Barrage, 2017), access to information (Hausman and Stolper, 2019), or housing discrimination (Christensen and Timmins, 2018)).

discount that was removed. Importantly for disaster resilience and recovery, we find that higher insurance prices would lead to fewer individuals living in high risk zones, highlighting that migration will likely be an important (albeit costly) channel to mitigate climate risks. In addition, we find that flood risk map updates are valuable sources of information and are appealing from both a distributional and efficiency perspective. Depending on the quality of old versus new maps, we estimate a conservative benefit cost ratio of 7.3 from new map revisions, and find benefits more greatly concentrated among low income individuals. Understanding sorting over flood risk and the implications for policy is critical as flooding remains one of the costliest and deadliest types of natural disasters around the world, and impacts are expected to increase significantly under a changing climate (Hallegatte et al., 2013; Smith and Katz, 2013).

The paper proceeds as follows. Section 2 reviews relevant literature, highlighting where our work contributes to existing knowledge. In Section 3, we describe our data, research setting, and empirically motivate some important sources of heterogeneity that we capture in our empirical model. We then present our residential sorting model in Section 4 and describe our estimation strategy in Section 5. Section 6 discusses our heterogeneous sorting results. Sections 7 and 8 present and discuss our policy counterfactuals, and Section 9 concludes.

2 Literature

Ever since Tiebout's observation that heterogeneous individuals sort across varied landscapes (Tiebout, 1956), rich literatures have emerged to understand how individuals locate relative to spatial (dis)amenities. First, an active residential sorting literature has developed to estimate preferences for spatial characteristics (Sieg et al., 2004; Walsh, 2007; Bayer et al., 2007; Klaiber and Phaneuf, 2010; Tra, 2010; Klaiber and Kuminoff, 2013; Fan and Davlasheridze, 2016; Bayer et al., 2016), including climate variables (Timmins, 2007; Albouy et al., 2016; Fan et al., 2018; Sinha et al., 2018; Ma, 2019).[2] More generally, a growing literature, known

[2]See comprehensive sorting overview by Kuminoff et al. (2013) and Klaiber and Kuminoff (2013).

as environmental justice, is concerned with understanding why environmental risk is often correlated with higher concentrations of lower income and minority residents (GAO, 1983; Taylor, 2000; Mohai et al., 2009).

Second, a large empirical literature utilizes hedonic property value models to estimate the capitalization of disamenities such as flood risk into home prices (Rosen, 1974). While results are mixed, the literature generally finds a price discount for residences in high risk flood zones, identified using both long run flood risk and also recent flood events (Hallstrom and Smith, 2005; Bin et al., 2008; Bernstein et al., 2019).[3] However, growing evidence surrounding the heterogeneous impacts of disasters on, for example, migration (Smith et al., 2006; Strobl, 2011) and income or debt (Gallagher and Hartley, 2017; Deryugina et al., 2018) highlights the potential for (ex-ante) sorting across underlying disaster risk, which has largely been overlooked in the hedonic literature. Moreover, the parameters from hedonic models, while aimed at estimating marginal willingness to pay, are typically not suitable for recovering the effects of counterfactual policy changes (Kuminoff et al., 2013).

A long standing challenge to empirical estimation in both sorting and hedonic models are that the spatial attributes of interest are often correlated with other (unobserved) spatial characteristics. In our setting, flood risk is often highly correlated with access to desirable water amenities. Without an approach to disentangle these collinearities, the positive (and potentially incompletely observed) amenity value may cause an upward bias in the effects of flood risk on price (Bin et al., 2008). Also relevant for the U.S. flood risk context, prices for flood insurance under the National Flood Insurance Program, the nation's leading flood insurance provider accounting for more than 95% of all policies (Dixon et al., 2006), can be heavily discounted, where, in some cases, the subsidized premium can be more than 85% below the risk-based premium (Kousky et al., 2016). As only the non-subsidized portion of flood insurance premiums are expected to be capitalized into house prices by attentive homebuyers (Shilling et al., 1989; Harrison et al., 2001; Bin et al., 2008), preferences to avoid flood risk could be biased downward if insurance discounts are not accounted for.

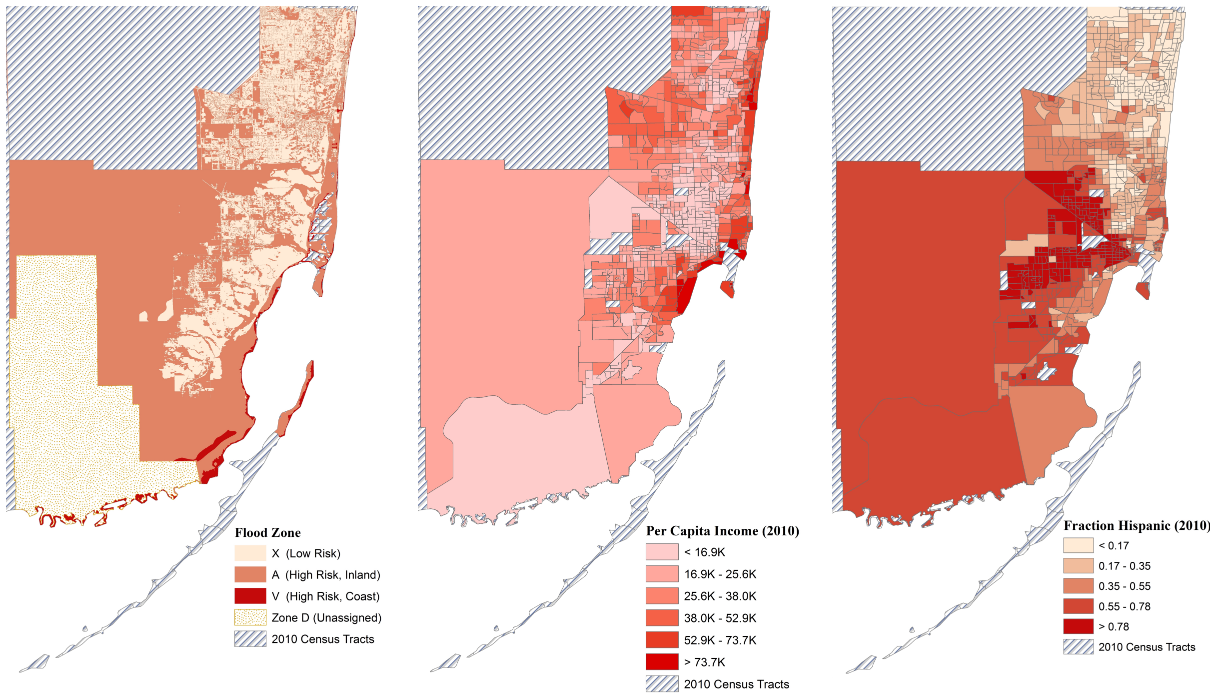
[3] See review by Beltrán et al. (2018).

We address the above concerns to examine sorting across flood risk zones. First, our use of a discrete choice residential sorting model allows for heterogeneous sorting over both flood risk zones and various amenities that are correlated with flood risks. Second, we account for relevant insurance-premium discounts in our calculation of housing costs. Of relevance, properties in high risk flood zones that carry a federally backed or regulated mortgage are required to purchase flood insurance (Flood Smart, 2016b), and compliance with this mandate is above 90% for recent home purchases (Dixon et al., 2006). This increases our confidence that we can recover salient net values of required flood insurance premiums from housing transactions data that include mortgage information. Importantly, this also improves our measurement of the portion of flood risk that is financially internalized by the homeowner. Lastly, we employ a boundary discontinuity design to deal with the endogeneity of disaster risk and spatial attributes (Black, 1999; Bayer et al., 2007). We show that observable factors do not change precipitously across floodplains within a certain distance of floodplain boundaries. Restricting our attention to homes on either side of a flood boundary, we apply boundary fixed effects to our sorting model. Combined, our heterogeneous sorting model, information on price discounts, and novel application of a boundary discontinuity design allow us to derive novel estimates of sorting over flood risk.

3 Empirical Overview

We examine the question of sorting over flood risk by using property sales data from 2009 to 2012 across Florida's Miami-Dade-Ft. Lauderdale-Port St. Lucie Combined Statistical Area (CSA). This CSA area represents approximately 2.3 million households and total property valued at more than $1 trillion. Flood risk here is expected to increase over time, as Miami is one of the top twenty cities across the world at highest risk for future flood losses due to sea level rise (Hallegatte et al., 2013). This region also contains significant heterogeneity in terms of who is exposed to flood risk. Figure 1 displays (a) floodplains in south Florida, (b) 2010 Census tract-level average per capita income, and (c) fraction of residents who are Hispanic

Figure 1: Flood Zones and Neighborhood Demographics



(a) Flood Zone (b) Per Capita Income (c) Fraction Hispanic

Source. Generated by authors using NFIP Digitized Flood Insurance Rate Maps and 2010 Census data for south Florida.

in 2010. It provides suggestive evidence of the correlation between (coastal) flood risk and income as well as (inland) flood risk and ethnicity, motivating the potential for sorting. In addition, Figure 1 shows a high degree of granular variation in flood risk, which necessitates property location information at a fine geographic resolution.

Important to modeling environmental risk in this context is an understanding of the public institutions surrounding flood risk in the United States. In response to flood threats and due to a lack of private insurance, Congress enacted the National Flood Insurance Act of 1968 that created the National Flood Insurance Program (NFIP), a federal flood insurance program. The NFIP also produced publicly available flood risk maps, known as Flood Insurance Rate Maps (FIRMs), which are periodically updated. FIRMs assign locations to one of several flood risk categories including: Zone A, with a freshwater flood risk greater than 1 percent per year; Zone V, with a coastal saltwater flooding risk of more than 1 percent per year; and Zone X, with a flood risk of less than 1 percent per year. Zones A and V, with more than a 1 percent annual flood risk, are designated as Special Flood Hazard Areas (SFHA), and structures in these areas are required to purchase insurance if they have a federally backed mortgage. Program premiums are set according to the dollar value of coverage purchased, the specific property's structural attributes, as well as its location with respect to a FIRM flood zone, and, for a subset of locations (approximately nine percent in our sample), the Base Flood Elevation, which represents the level to which floodwater is anticipated to rise during a 100-year flood.

NFIP premiums are priced to reflect underlying flood risk, but price supports of several types reduce premium rates to below actuarially fair levels. The program's three largest price discount schemes include preferential rates to (i) pre-FIRM properties that were built before the first flood insurance rate map was released in their community; (ii) residents of locations in the Community Rating System, who receive a price reduction of up to 45% as determined by flood activities at the community level; and (iii) grandfathered properties with pre-existing flood insurance policies that can maintain preferential rates after new flood maps are released.[4] These discounts are intended to encourage uptake, ensure affordability, and eliminate some of the financial pressure on public post-disaster aid programs (Kousky and Shabman, 2014).

3.1 Data

We categorize our data into four main groups: 1) housing transactions in Florida from Dataquick, Inc., 2) digitized Flood Insurance Rate Maps and insurance premium rate tables, 3) mortgage applications collected under the Home Mortgage Disclosure Act, and 4) information on various other spatial attributes. We provide a brief overview of our process to construct the final data and refer readers to Appendix 9 for a detailed description of our data sources and the data construction process.

We begin with all arms-length sales of owner-occupied residential properties from the Miami-Dade, Port St. Lucie, Fort Lauderdale Combined Statistical Area from 2009 to 2012. The data include information on selling price, date of sale, numbers of bedrooms and

[4]Preferential rates can be passed on to future owners if a policy is continually held.

bathrooms, and mortgage information. We calculate a property's flood-insurance premium based on its flood zone (assigned using Flood Insurance Rate Maps (FIRM)), structural characteristics, and the year built. This information is sufficient to determine the effective insurance premium rate (per $100 of building coverage) for most properties.[5] We then multiply this rate by the amount of building coverage, set as either the recorded loan amount or $250,000 (whichever is lower) (NFIP, 2016). We note that the pre-FIRM discount is already embedded in the NFIP premium rate based on the year that a house was built. We then calculate a property's final insurance premium by incorporating the Community Rating System (CRS) program discount if a property belongs to a participating community as designated by the NFIP. We do not include the price for contents coverage as this type of coverage is not mandatory and would not likely impact home price. We also map each property to the closest flood zone boundary using Geographic Information Systems: we first split FIRM flood map polygon boundaries into segments, which are assigned a unique identifier, and then we find the closest segment (in terms of distance) to each house.

Next, to characterize the neighborhoods in which houses are located, we map each house to nearby spatial amenities. These include (1) distances to the nearest park, river, and coast, (2) number of Institutional Controls Registry (ICR) sites within 3 kilometers (a proxy for local environmental quality), (3) test scores as a proxy for public school quality, and (4) tract-level per-capita income and race/ethnicity population shares from the 1990 census.

Lastly, we follow the procedure outlined in Bayer et al. (2016) to recover the race and income of buyers in our sales data using mortgage information from the Home Mortgage Disclosure Act. This is so that we can categorize households into different "types," defined by race and income, where income is categorized into bins based on quintiles of the observed income distribution.[6]

Our final sample includes 48,174 individual house sales between 2009 and 2012 across six

[5]See the appendix for assumptions used to calculate property-specific premiums for a small number of properties with missing data.

[6]Matching between sales and mortgage data is imperfect: we only recover information for 47% of our data. However, the resulting sample is representative compared to Census data. For details, see Appendix 9.

counties and 953 census tracts in Florida.[7] Each house is described by its structural attributes and neighborhood characteristics, such as the distance to various spatial (dis)amenities. At the time of each sale, we know the race/ethnicity (White, Black, Hispanic, or Asian) and income of the primary buyer involved, and the flood zone and premium that the buyer faces. The average sales price is $219,841, where prices are normalized to January 2010 dollars using the Consumer Price Index for All Urban Consumers in the South region for the expenditure category of "Housing" (BLS, 2012). Appendix Table B.1 provides summary statistics for property and household characteristics. The majority of the properties are either in an X or an A zone, with less than 1 percent of our sample belonging to the V zone.[8]

Regarding flood insurance premiums and discounts, we calculate an annual covered amount is $159,664 on average, with a median of $154,982 (see Appendix Table B.3). The full premium calculated prior to any discounts is, on average, $2,113 per year, with a median of $808. The pre-FIRM discounts, then provide an average discount of almost $1,000 relative to the full premium. Houses in our sample receive CRS discount rates of between 0 and 25 percent, with an average of 12.0 percent. Incorporating CRS discounts brings the average fully subsidized insurance premium to $984 (with a median of $714 per year). Large investments in flood mitigation, such as flood-proofing and elevating structures, can certainly distort the researcher's measurement of the flood risk that is borne by the household, but the CRS discounts that we observe in our study area are low enough that they are unlikely to alter the underlying flood risk in practice.[9] We calculate the total discount as the difference between the calculated insurance premium before and after the CRS and pre-FIRM discounts. The average total discount in place is $1,129, which represents about a 50 percent discount off

[7]The six counties include Miami-Dade, Broward, St. Lucie, Martin, Indian River, and Okeechobee. We lose Palm Beach County because no digitized flood map was available at the time of our analysis.

[8]This is consistent with the observed distribution of homes in the area. Using GIS data on all properties in Miami-Dade County, the authors estimate that 0.27% of properties are in the V zone compared with 0.23% observed in the V zone across our sample.

[9]The activities that communities can undertake to earn credit towards receiving a discount range from public information provision (earning the fewest credits) to flood mitigation measures (earning the most credits). The amount by which undertaken activities actually decrease household-level flood risk is potentially low. For example, across the United States, 93% of communities receive credits for outreach projects (credit type 330) whereas only 13% receive credits for flood protection activities (credit type 530) (FEMA, 2017).

the non-discounted nominal insurance premium.

3.2 Empirical Evidence

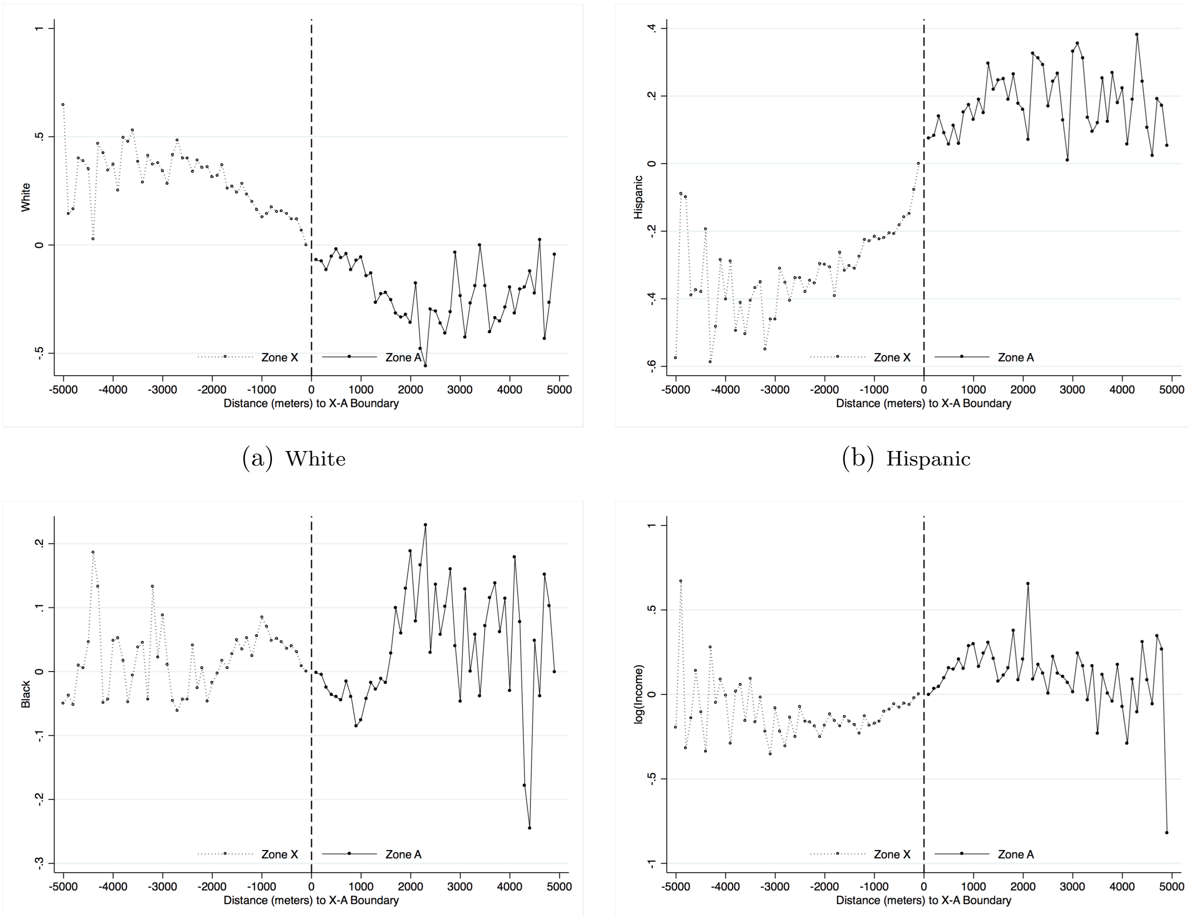
Before describing our model, we provide stylized evidence that heterogeneous sorting takes place according to observable household characteristics. We also assess the extent to which discounts on insurance premiums vary systematically by observable characteristics. These two considerations strongly suggest the use of a model that allows for heterogeneous sorting decisions with respect to flood risk exposure and other correlated amenities, and support the notion that NFIP reforms may have some potentially important distributional consequences.

To assess heterogeneous sorting across flood zones in our data, Figure 2 plots buyer characteristics against distance to the nearest X-A flood boundary (delineated by a vertical dashed line) for A and X zone houses within 5 kilometers of this boundary. Specifically, we plot the coefficients from regressions of buyer race, ethnicity, and income on a set of distance-to-boundary indicator variables at 100-meter intervals (plotted to the left of the dashed line) and the same set of dummy variables interacted with indicators for whether the house is in an A zone (plotted to the right of the dashed line). All averages are normalized to the 100-meter distance on the X side of the boundary. Roughly speaking, this means that moving from left to right on the x-axis increases flood risk.

Figures 2(a) and 2(b) show that higher risk A zone areas are less white and Asian, and are primarily Hispanic. The share of black buyers in Figure 2(c) are mostly similar across X and A zones, although there is some evidence of fewer black buyers in the area immediately across the X-A zone boundary. Figure 2(d) plots the logarithm of income. As one crosses into the A zone, residents are higher income, though at about 3 kilometers from the boundary, incomes begin to fall.[10] From a revealed preference perspective, this would suggest that Hispanics and higher income households are more likely to sort towards flood risk. However, given flood

[10] We note that for all of these figures, estimates become noisy as one moves farther away from the boundary (the vertical dashed line) on the A zone side. This is because A zone represents inland flooding; as such, increasing the distance from this boundary could either mean being closer to the V zone or to a different, lower risk area.

Figure 2: Buyer Characteristics by Distance to Flood Boundary



(a) White

(b) Hispanic

(c) Black

(d) log(Income)

Note. Each figure plots the coefficients from a regression of some attribute against distance-to-flood boundary dummy variables at 100-meter increments from the X zone (on the left) to the A zone (on the right). All points are normalized to the 100-meter distance on the X side of the boundary.

risks' spatial correlation with water amenities, this could also be driven by heterogeneous preferences for coastal amenities. Nevertheless, it is apparent from these figures that there are systematic differences in the distribution of race and income across flood zones.

Disentangling heterogeneous sorting over flood risk from its correlated amenity value would be important to recover unbiased preference parameter estimates. Our boundary fixed effects model, combined with neighborhood demographic controls, is aimed at accomplishing this. Table 1 assesses mean differences between A and X zone characteristics using low flood risk properties where the area opposite its flood boundary is of high flood risk, and vice versa. We provide mean differences for the full sample and various distance-to-boundary samples (i.e. 5, 3, 1, 0.5, or 0.3 km), along with corresponding t-statistics that the mean difference is equal to 0. While we generally reject that mean differences are equal to 0, the unconditional differences decrease as we narrow the window of consideration around the flood boundary.[11] We thus restrict our sample to properties where the nearest flood boundary is at most 1 kilometer away in our boundary discontinuity design, limiting our comparison to houses near the same but opposite sides of a boundary through the use of boundary fixed effects. The boundary discontinuity design alone would be insufficient to deal with differences across flood boundaries due to sorting based on heterogeneous preferences to avoid flood risk and/or endogenous neighborhood differences (as noted by Bayer et al. (2007) in the context of sorting over school districts). Thus, we also control for endogenous neighborhood attributes (e.g. race and income), which we include from the 1990 Census at the tract level, in addition to allowing for heterogeneous preferences across homeowners.

To further assess our sample restriction, we demonstrate that different distance-to-boundary sample limitations within 1 kilometer do not materially affect the results of a hedonic model with boundary fixed effects. Table 2 presents hedonic regressions of the annual rental price on house, flood, and other spatial attributes. The annual insurance premium subsidy is subtracted from annual rental prices to adjust for flood-insurance discounts. Each column represents a separate regression. Our coefficients of interest are Special Flood Hazard Area (SFHA) indicator variables, denoted "SFHA", that designate high flood risk (zones A or V) status, where the omitted group is composed of X zone houses exposed to lower flood risk. All regressions include controls on house characteristics (house type indicators, number of bathroom, bedrooms, square foot, age and pre-FIRM status), neighborhood characteristics (local environmental quality, school quality, distance to the nearest park and river), year fixed effects, and county fixed effects.

In panel A, sales prices are approximately $2,203 lower for properties in the SFHA zone

[11] We also find graphical support for this in Appendix Figure B.2 that uses distance-to-boundary figures (similar to Figure 2) with various spatial characteristics as the dependent variable.

relative to those in the X zone (column (1)). Upon progressively adding controls for surface elevation, base flood elevation, and distance-to-coast bins,[12] houses in SFHA zones sell for \$1,120 lower than comparable houses in the X zone. Notably, there is a very steep price gradient with respect to distance to the coast in column (3). We next restrict our sample to houses within 1 kilometer of a flood boundary and re-estimate the model to include boundary fixed effects, following Black (1999), in column (4). Our MWTP estimate for SFHA zone houses becomes -\$659, which is 5.8 percent of average housing prices in our sample (assuming a 5 percent discount rate), which is comparable to previous work.[13] In panel B, we estimate the boundary fixed effects model with various distance-buffer sample restrictions. These estimates are economically similar and are not statistically different than the model using a 1-kilometer buffer; we therefore use the 1-kilometer sample restriction for all estimation. Our main estimation sample using the boundary discontinuity design consists of 32,027 sales across 784 tracts, where the average price is \$225,434.[14]

Last, we can also use the hedonic model to assess the importance of accounting for premium subsidies. The last column of Table 2 re-estimates the model in column (3) without boundary fixed effects, but uses annual rents that ignore the price supports that we calculate for each house. The SFHA zone coefficient is -\$19. These differences point to the variation in discounts between zones and the extent to which ignoring price supports will matter for hedonic and sorting estimates.

4 Model

We estimate household willingness to pay to avoid flood risk using a residential sorting framework that we adapt to incorporate preferences to avoid flood risk.[15] In what follows, we describe the household's choice set, their preferences, and their optimization problem.

[12]The omitted category is houses farther than 5 kilometers from the coast.

[13]For example, Harrison et al. (2001), Bin et al. (2008), and Zhang (2016) find that houses in flood prone areas sell for a price discount ranging between 5 and 11 percent.

[14]Summary statistics for the boundary fixed effects sample are presented in Appendix Table B.2.

[15]Recent examples of work using residential sorting models that are most relevant to our paper include Bayer et al. (2007), Klaiber and Phaneuf (2010), Tra (2010), and Ma (2019).

Choice Set Beginning with the sample of houses in the Miami-Dade CSA that are near flood boundaries, a household chooses to live in one of several types of housing in these neighborhoods. In particular, it makes a discrete, residential location decision based on the attributes of each location it is facing and the costs of living there. A specific choice of housing is constructed as a combination of the following geographic and house characteristics: census tract, flood category (X, A, V), house structure,[16] building type, base flood elevation (BFE) if available, pre-FIRM status, and one of eight distance-to-coast bins ranging from less than 100 meters to more than 5 kilometers.[17] We incorporate pre-FIRM status, house and building type, and BFE into the residential choice because they determine the specific NFIP rate used to compute insurance premiums.[18] In addition, we include coastal distance bins in order to better control for the unobserved impact of water-related amenities later on. Because not all house types are available in each year, the number of available choices (J_t) will vary from year to year as well. Our categorization of choice results in approximately 2,150 alternatives to choose from in each year from 2009 to 2012.[19] For the remainder of the paper, we refer to each of these choices as a "residence."

Our data and choice framework imply several assumptions that we must make about how households make decisions with respect to residential location. First, our sample in south Florida and our boundary sample restriction places a limitation on the extent of the market. Previous work using sorting models have considered similarly sized markets.[20] Data from the Census for our study area and time frame also suggest that the extent of

[16] The housing structure types are assigned based on the NFIP rate structures, which are '1 to 4', '2 to 4', single, mobile, and residential. The overlap in categories (e.g. '1 to 4' versus single family) is due different housing categorizations being used for different flood zones. For example, pre-FIRM houses in zone 'AE' are categorized by mobile, single, '2 to 4', and other; on the other hand, the categories that are used if the houses are post-FIRM are '1 to 4', mobile, and other.

[17] One kilometer distance bins are used for houses located between 1 and 5 kilometers of the coast. Within 1 kilometer, we additional categorize houses to be within 100 meters, 100 to 500 meters, and 500 to 1000 meters. Houses located more than 5 kilometers away from the coast are considered to be in one category.

[18] For details, see Appendix 9.

[19] The number of choices, J_t, for $t = 2009, \ldots, 2012$ is respectively 2,408, 2,180, 2,137, and 1,893.

[20] Tra (2010) examines locational choices in the Los Angeles metropolitan area, Sieg et al. (2004) focuses on five counties in southern California, Bayer et al. (2007) and Bayer et al. (2016) examine moving within the San Francisco Bay Area (consisting of six counties), and Klaiber and Phaneuf (2010) model housing decisions in the Minnesota Twin Cities area.

the market considered here is appropriate.[21] Moreover, re-estimating our sorting model without the boundary sample restriction recovers a higher flood risk willingness to pay that is comparable to the hedonic estimate without boundary fixed effects (likely driven by correlated unobservables), but does not alter our conclusions about the distributional implications of our sorting results.[22] Second, we assume that households choose where to live conditional on moving in a given year (the year of sale). In other words, we do not model the decision of whether and when to move, but just where to move conditional on moving. Third, all households are assumed to face the same choice-set in the CSA. Differences in consideration sets can impact preference estimates (Kuminoff, 2009). In addition, recent work has also shown that subtle forms of housing discrimination, e.g., the number of houses or sample of neighborhoods shown by a realtor, can drive wedges in choice sets that are correlated with race and ethnicity (Christensen and Timmins, 2018). However, the U.S. Department of Housing found that on most discrimination measures, Hispanic homebuyers in Miami faced similar levels of discrimination as the overall incidence of random discrimination (irrespective of race) in the Miami sample (Turner et al., 2002).[23] While we do not argue that housing discrimination is not an issue in this context, we believe our results are still relevant given that there has been little work to assess whether heterogeneous sorting with respect to flood risk even exists, whether it be driven by discrimination, preferences, or other factors such as, e.g., differential beliefs (Bakkensen and Barrage, 2017) or access to information (Hausman and Stolper, 2019). We note that the underlying sorting mechanisms are an important area of future research. In addition, to the extent that systemic discrimination or other channels would not be undone by flood insurance program reforms, the parameters recovered from our model could still be used to estimate our policy counterfactuals.

[21] Table B.5 in the appendix shows aggregate statistics from the Census for all movers from our study area between 2009 and 2013. It reveals that 69.8 percent of all moves within our study area were within-county moves, and 76.8 percent of all moves were within the CSA.

[22] These results are available from the authors upon request.

[23] See, for example, Exhibit A4-4 in the supplemental materials.

Household Preferences A household's preference for a residence j at time t depends on the characteristics of the residence. Many of these characteristics are observed by the econometrician and include structural and geographic characteristics such as the distance to various (dis)amenities (e.g. the coast, highways). There are also aspects of residences that factor into a household's decision that are not observed by the econometrician. Let X_{jt} denote attributes of a residence that are observed and ξ_{jt} describe those that are not. A subset of observable attributes $X_{1jt} \in X_{jt}$ include housing structure-related variables. In practice, these are indicators for single family houses, condominiums, pre-FIRM status (also a proxy for age), and distance-to-coast bins, where the omitted category is for residences greater 5 kilometers from the coast. A second set of attributes $X_{2jt} \in X_{jt}$ includes indicators for whether a residence is located in a Special Flood Hazard Area (SFHA) (i.e. A or a V zone), whether BFE is assigned,[24] surface elevation, the distance to the nearest river and park, local environmental quality, and school quality. We also allow households to have preferences over neighborhood sociodemographics by including tract-level per capital income, share of population that is black and share that is Hispanic. As contemporaneous demographics are likely to be endogenous, we include these characteristics as determined in 1990 instead of using 2010 census characteristics.[25] For attributes that are not constant within a choice (e.g. school quality), an average is taken among each observed house in that choice. In order to assign neighborhood choices with the nearest flood zone boundary, we assign the boundary identifier of the choice to be that of the house closest to any boundary of all houses in that choice set. Among the set of characteristics in X_{2jt}, we separately denote the indicator for belonging to a high risk floodplain, our attribute of interest, as $SFHA_j$, which proxies for flood risk.

For flood risk and X_{2jt}, we allow households to have heterogeneous tastes based on its race/ethnicity and income quintile, denoted by $Z^i = (1, z_1^i, \ldots, z_K^i)$. These observable

[24] Recall that 91% of our sample does not have a base flood elevation assigned by the NFIP, and so we include an indicator for BFE assignment in the utility function even though we use the actual level of base flood elevation in computing the insurance premium, when applicable.

[25] We include neighborhood demographic characteristics mainly to serve as controls. As these lagged demographics may still be endogenous, we refrain from interpreting the coefficients on these characteristics.

characteristics include indicators for black and Hispanic (with the omitted group being white or Asian) and for four of the five income quintiles (where the omitted group is the lowest income quintile).[26] To additionally capture heterogeneous tastes for positive water-based amenity value, we include an indicator for a residence being within 100 meters of the coast in X_{2jt}. Lastly, we allow for tastes to vary based on an idiosyncratic component that is household- and residence- specific, ϵ^i_{jt}.

Given the attributes of residences, households trade off between enjoying the services provided by the residences with the flow cost of living in that location, P_{jt}, which enters linearly into household utility.[27] Since we do not observe the prices of houses that individuals do not choose, we calculate all rental prices for residential choices by taking the average house price of houses that sold in that location and then annuitizing the average price with a 5% discount rate. We then account for insurance premium subsidies here by subtracting the discount in annual insurance premium from the annual rent. A household i receives the following indirect utility from choosing to move to residence j at time t:

$$V^i_{jt} = \alpha_{x1} X_{1jt} - \alpha_p P_{jt} + \xi_{jt} + \alpha^i_r SFHA_j + \alpha^i_{x2} X_{2jt} + \epsilon^i_{jt} \tag{1}$$

where

$$\alpha^i_\ell = \alpha_{0,\ell} + \sum_{k=1}^{K} \alpha_{k,\ell} z^i_k \; for \; \ell = \{r, x2\} \tag{2}$$

In anticipation of the need to deal with unobserved factors that are correlated with flood risk and price (elaborated in the next section), we re-write the indirect utility, V^i_{jt}, so that it can be separated into choice- and individual- specific components:

$$V^i_{jt} = \delta_{jt} + \left(\sum_{k=1}^{K} \alpha_{k,r} z^i_k\right) SFHA_j + \left(\sum_{k=1}^{K} \alpha_{k,x2} z^i_k\right) X_{2jt} + \epsilon^i_{jt} \tag{3}$$

[26] In robustness checks, we allow race-income specific preferences as well.

[27] This setup assumes that household budget constraints enter linearly into the utility, ruling out income effects. Income limitations on the choice of residence are more likely to appear through differential choice sets than in choice probabilities.

where

$$\delta_{jt} = \alpha_{0,r} SFHA_j + \alpha_{x1} X_{1jt} + \alpha_{0,x2} X_{2jt} - \alpha_P P_{jt} + \xi_{jt} \tag{4}$$

The choice-specific component, δ_{jt}, represents the mean utility of the base (or omitted) group, which consists of whites and Asians in the lowest income quintile. The parameter $\alpha_{k,r}$, the coefficient on the interaction between the individual's type and the neighborhood's floodplain, represents the additional utility from living in $SFHA_j$ that a household of type k receives *relative to* the base group. The parameter $\alpha_{k,x2}$ is similarly interpreted with respect to the set of attributes in X_{2jt}. These heterogeneous preference parameters, or the coefficients on individual-specific components of utility ($\alpha_{k,r}$, $\alpha_{k,x2}$), are distinguished from the base group parameters on the choice-specific components of utility ($\alpha_{0,r}, \alpha_{x1}, \alpha_{0,x2}, \alpha_P$) because they will be estimated in stages.[28]

Conditional on moving at time t, household i chooses to live in residence $d_t^i = j$ if it yields the highest utility among all other alternatives:

$$d_t^i = j \text{ if } V_{jt}^i \geq V_{j't}^i \;\; \forall \;\; j' \neq j \tag{5}$$

Further assuming that household idiosyncratic tastes for choices are distributed i.i.d. Type I Extreme Value, the expected probability that a household chooses residence j has the following closed form expression (McFadden, 1978):

$$P_{jt}^i = Pr\left(V_{jt}^i \geq V_{j't}^i \;\; \forall \;\; j' \neq j \mid X, P, Z\right) = \frac{e^{V_{jt}^i}}{\sum_{j'} e^{V_{j't}^i}} \tag{6}$$

With $N_t \in N$ residents moving at time t, the predicted share of each residence that is chosen can be calculated by averaging over the probability that individuals choose each location in that period:

$$s_{jt} = \frac{1}{N_t} \sum_{i}^{N_t} Pr\left(V_{jt}^i \geq V_{j't}^i \;\; \forall \;\; j' \neq j \mid X, P, Z\right) \;\; \forall \; j, \; t \tag{7}$$

[28]For the subset of neighborhood characteristics (rental price and attributes X_{1jt}) where we have assumed homogeneous preferences, the coefficients on these variables apply to all groups.

5 Estimation

Estimation of the problem will proceed in two stages. Stage 1 recovers heterogeneous preference parameters and mean utilities using Maximum Likelihood Estimation. Stage 2 follows with a regression that decomposes the mean utility estimates from stage 1 to recover the remaining base group parameters. We refer to this regression as a "mean utility decomposition." It is in this stage that we include boundary fixed effects and employ an instrumental variables strategy to deal with the endogeneity of price. Standard errors are bootstrapped using 500 draws of the sample with replacement.[29] We detail each step below.

Stage 1 In the first stage, we build the following log-likelihood function based on predicted choice probabilities that are consistent with our locational choice model:

$$\ell\ell(d, X, P, Z) = \sum_{t}^{T} \sum_{i}^{N_t} \sum_{j}^{J_t} 1(d_t^i = j) \cdot log\ P_{jt}^i \tag{8}$$

The indicator, $1(d_t^i = j)$, is equal to 1 if a household i actually chooses to live in neighborhood j. We maximize the log-likelihood to estimate the parameters in the household's utility. Recall that location-specific attributes (such as flood risk) has been characterized by a set of J mean utilities, δ_{jt}. In this stage, we recover these mean utilities first, instead of the coefficients on the various attributes $(\alpha_{0,r}, \alpha_{0,x1}, \alpha_{x2}, \alpha_p)$ that contribute to these mean utilities. This procedure then returns the set of mean utility parameters, δ_{jt}'s, and household-specific taste parameters, $(\alpha_{k,r}, \alpha_{k,x2})$, that best explain the actual choices made in the data according to our model. Practically, we normalize the mean utility of one choice in each period to be 0 and then solve for the mean utilities of the remaining choices using a Berry (1994) contraction mapping routine. The contraction mapping routine to recover the δ_{jt}'s is nested in an outer loop of the likelihood estimation procedure that varies the household-specific taste parameters, $(\alpha_{k,r}, \alpha_{k,x2})$.

[29] This is to account for estimation error for the second stage estimates, which performs estimation using first stage estimates.

A benefit of estimating mean utility parameters first instead of the choice-specific parameters directly is that we postpone dealing with endogeneity concerns associated with the choice- and period-specific unobservable, ξ_{jt}, until the second stage, where mean utility is linear in parameters.[30] Furthermore, using a contraction mapping yields computational savings, which is important given the large number of choice alternatives in our setting.

Stage 2 With the first stage estimates in hand, the second stage regresses the estimates of residence mean utilities on neighborhood attributes to recover the preferences for these attributes:

$$\hat{\delta}_{jt} = \alpha_{0,r} SFHA_j + \alpha_{x1} X_{1jt} + \alpha_{0,x2} X_{2jt} - \alpha_P P_{jt} + \xi_c + \xi_t + \xi_{jt} \tag{9}$$

The coefficients on attributes for which households have heterogeneous preferences ($\alpha_{0,r}, \alpha_{0,x2}$) represent the preferences of the base group, while those on the remaining attributes (α_{x1}, α_P) represent the average preferences of all households. Here, we additionally introduce county (ξ_c) fixed effects to control for unobserved differences between counties, and year (ξ_t) fixed effects to adjust for macroeconomic price trends.

Equation (9) can be estimated by OLS; however, we are concerned with two important endogeneity issues. First, cost of living in a neighborhood, P_{jt}, will likely be correlated with unobserved neighborhood quality, ξ_{jt}. In this respect, we follow the approach taken in Bayer and Timmins (2007) by constructing instruments based on the exogenous attributes of distant communities that affect the price of neighborhood j. The logic behind this instrument is based on the equilibrium sorting model: the cost of living in a community j depends, in part, on the availability of residences in distant communities that may be considered substitutes. The exclusion restriction is satisfied with this instrument because while the attributes of farther-away communities can affect price in equilibrium, the attributes of these distant communities should not directly enter into the utility of living in residence j. We use the

[30] Berry (1994) shows that we can recover the set of mean utilities in this way by inverting choice shares; in other words, there is a unique vector of δ_{jt}'s that sets the predicted shares of choice alternatives equal to the observed shares.

share of urban, open land in nearby communities as an instrument, which is a measure of undeveloped land.[31]

To implement this instrument, let $\widehat{\ }$'s indicate first-stage estimates. Using a guess of the price coefficient, $\alpha_P^{(0)}$, we adjust the estimated mean utility for a location with the cost of living there by moving price to the left side of the equality in (9), $\hat{\delta}_{jt} + \alpha_P^{(0)} P_{jt}$. Next, we estimate the following modified version of equation (9) with the adjusted mean utilities as the dependent variables and additionally include the share of undeveloped land within a 1-, 3-, and 5- kilometer radius, denoted by $\widetilde{U}_j$,

$$\hat{\delta}_{jt} + \alpha_p^{(0)} p_j = \alpha_{0,r} SFHA_j + \alpha_{x1} X_{1jt} + \alpha_{0,x2} X_{2jt} + \alpha_{\widetilde{U}} \widetilde{U}_j + \xi_c + \xi_t + \widetilde{\xi}_{jt} \tag{10}$$

Since we have allowed characteristics of neighboring residences (within 5 kilometers of choice j) to directly affect mean utility, the error $\widetilde{\xi}_{jt}$ now captures attributes of distant neighborhoods (i.e. farther than 5 kilometers) that affect cost of living in j. With the estimates from the modified mean utility regression (10), which we denote with $*$'s, we can formulate a modified version of the mean utility, where $\widetilde{\xi}_j$ are set to 0,

$$\tilde{\delta}_{jt} = \alpha_{0,r}^* SFHA_j - \alpha_p p_j + \alpha_{x1}^* X_{1jt} + \alpha_{0,x2}^* X_{2jt} + \alpha_{\widetilde{U}}^* \widetilde{U}_j + \xi_c + \xi_t \tag{11}$$

We then solve for the vector of prices, p_j^{IV}, that sets predicted shares based on $\tilde{\delta}_{jt}$ equal to actual shares:

$$\sigma_{jt} = \frac{e^{\tilde{\delta}_j + \left(\sum_{k=1}^K \alpha_{k,r} z_k^i\right) SFHA_j + \left(\sum_{k=1}^K \alpha_{k,x2} z_k^i\right) X_{2jt}}}{\sum_{j'} e^{\tilde{\delta}_{j'} + \left(\sum_{k=1}^K \alpha_{k,r} z_k^i\right) SFHA_{j'} + \left(\sum_{k=1}^K \alpha_{k,x2} z_k^i\right) X_{2j't}}} \tag{12}$$

The difference in predicted and actual shares is driven by variation in developed land that is more than 5 kilometers away (i.e. $\widetilde{\xi}_j$), which we do not expect to directly influence the utility of living in a given location. Thus, the prices, $p_{j,t}^{IV}$, that clear the market will also reflect this variation. In practice, the instrumental variables estimates may be sensitive to the initial

[31] This data is assessed from digitized files provided by the Florida Fish and Wildlife Conservation Commission and Florida Natural Areas Inventory. For details, see https://www.fnai.org/LandCover.cfm.

guess for the price coefficient. We repeat this procedure with updated guesses of the price coefficient until the final price coefficient estimate is stable.

Second, water-related amenities are likely to be correlated with flood risk, i.e., $E\left[\xi_{jt} SFHA_j\right] \neq 0$. Recall that we categorized residences using distance-to-coast bins and BFE. Inclusion of these distance bins as fixed effects avoids the comparison of houses with vastly different levels of coastal access. Conditional on distance to the coast, inclusion of average elevation also provides some control of the quality of coastal view. Finally, we include boundary fixed effects to compare residences just on either side of a flood zone boundary, which subsume our county-level fixed effects ξ_c. To the extent that these controls are imperfect, however, the unobserved amenity correlates should attenuate our flood risk parameter estimate.

6 Sorting Over Flood Risk

Do individuals sort over flood risk? We first present our base group estimates from the mean utility decomposition and then follow with estimates of the heterogeneous sorting parameters.[32] The first two columns of Table 3 give our main estimates from the mean utility decomposition and standard errors. These estimates are a result of applying the Bayer and Timmins (2007) instrumental variables strategy for price, which exploits variation in the share of undeveloped land in nearby communities. Boundary fixed effects are included but not shown. For comparison, we present OLS estimates in the final two columns.[33]

First, the magnitude of the IV price coefficient (-1.739) is much larger than the OLS estimate (-0.005). The F-statistic from the first stage regression is 26.65, significantly larger than the rule of thumb of 10 suggested by Staiger and Stock (1997). The direction of the bias

[32] We present the stage 2 estimates before stage 1 because the heterogeneous parameter estimates (recovered from stage 1) are interpreted *relative* to the base group estimates as discussed in section 4.

[33] Recall that households have heterogeneous sorting parameters for a subset of the amenities, namely flood zone, surface elevation, BFE assignment, distance to the nearest river and park, 0.1km coastal bin indicator, environmental quality, school quality, and neighborhood sociodemographics (tract-level per capita income, share black, and share Hispanic). Estimates in Table 3 thus represent the utility of the base group of white and Asian households in the lowest income quintile. For all other attributes (price, single, condominium, pre-FIRM status and distance-to-coast bins greater than 0.1km), the coefficients should be interpreted as the average effects on utility for all groups.

is consistent with cost-of-living being positively correlated with the quality of unobserved amenities. Our main coefficient of interest, $SFHA$, finds that living in a high risk floodplain decreases utility for those in the base group (we return to magnitudes shortly).[34] Focusing on the coefficients of other amenities in the decomposition, we also find that their signs are generally reasonable: households like school quality, single family homes, and newer houses (proxied by pre-FIRM status), and dislike environmental nuisances (although it is not statistically significant) and distance from the coast, parks, and rivers. We note that for many of these attributes, the coefficients from OLS estimation are counter-intuitive, highlighting the need for the IV.

Next, we present the heterogeneous parameter estimates in Table 4. We use the price coefficient (α_p) to convert all estimates into a dollar measure for marginal willingness to pay (MWTP) rather than present estimates in utils.[35] Note that we include the MWTP estimates for the "Base Group" in Table 4, which are simply the mean utility estimates from the first stage converted into a dollar value.[36] We do this to aid the interpretation of the heterogeneous parameter estimates. For example, a particular (non-base) group's MWTP to live in a floodplain is the sum of the base group's MWTP and its (heterogeneous) parameter estimate (converted to a dollar value). Panel A focuses on preferences to avoid flood risk. We also list the average and standard deviations of income by group for reference in the last two columns. On average, the base group is willing to pay \$710 per year to *avoid* living in a high risk flood zone. Assuming a 5 percent discount rate, this represents 6.3 percent of average housing prices given an average price in the sample of \$225,434. Relative to this group, the MWTP to *avoid* flood risk is close to 87 percent as high for Hispanic owners, and about 68 percent as high for black owners. The MWTP also increases with income, where those in the two highest income quintiles have willingnesses to pay that are approximately 9 to 28

[34] Less than 1 percent of our sample lives in the V zone, which is representative of the spatial distribution of the population with respect to flood risk in this area. With so few observations in the V zone, we are unable to estimate heterogeneous preferences to live in the V zone separately by household type and therefore combine the V and A zones in estimation.

[35] Raw utility estimates are presented in Table B.6 of the appendix.

[36] Specifically, the base group MWTP for flood risk is computed as $\frac{\alpha_{0,r}}{\alpha_P} \times 1000$ (since annual rent is in \$1,000's of dollars).

percent higher compared to the base group. While all groups dislike flood risk, low income and minority groups are more likely to sort into floodplains. If protection from flood risk is a normal good, then it is intuitive that willingness to pay to avoid flood risk increases with income or wealth. This may also explain the low MWTP estimates for minorities, who have an average income that is $24 thousand (hispanic) and $57 thousand (black) less than whites and Asians (who have an average income of $111 thousand overall).[37]

We present flood risk estimates that ignore NFIP price supports in the second two columns of Panel A. Ignoring premium discounts biases the MWTP estimate downward (consistent with our hedonic estimates), where the MWTP to avoid flood risk is actually positive. As such, the model that ignores these discounts attributes higher flood risk exposure than what is actually internalized for these individuals for the housing costs paid. In other words, the same amount of housing cost reduction looks to be compensating a much larger increase in exposure to flood risk than in reality. Again, these results motivate the importance of accounting for related program discounts in sorting and hedonic analyses.

These results provide a basis for the concern that the correlation between environmental risk and vulnerably groups can be driven at least in part by heterogeneous sorting, and, in particular, the "coming to the nuisance" by more vulnerable groups and, at the same time, the systematic flight from the nuisance by less vulnerable groups (Depro et al., 2015; Banzhaf et al., 2019). While we posit that income is an important underlying driver of heterogeneity, it is unlikely to be the whole story. There could be various factors correlated with income that contribute to heterogeneous sorting, which we cannot disentangle, including differential tastes (Banzhaf and Walsh, 2008), beliefs (Bakkensen and Barrage, 2017), access to information (Hausman and Stolper, 2019) and learning (Ma, 2019), or housing discrimination (Christensen

[37]In Table B.7 of the appendix, we additionally provide preference estimates by race-by-median-income groups. The base group of low-income, white residents is willing to pay $755 per year to avoid living in a high risk flood zone, which is similar to the previous base group estimate. The same general patterns emerge as before: white households are more likely avoid flood risk relative minorities, holding income constant. Conditional on race and ethnicity, flood risk avoidance generally increases with income. See the appendix also for sorting results over other spatial amenities and nuisances. In particular, we find a similar income gradient for elevation, BFE assignment, and proximity to the coast, which confirms our earlier intuition that these variables proxy for access to water-related amenities.

and Timmins, 2018), all of which can moderate adaptation and resiliency to natural disasters. Although we do not have the ability to separate these alternative mechanisms and note this as an important area of future research, such heterogeneity will likely leave low-income and minority groups relatively more exposed to high risk areas. Finally, this evidence of sorting reaffirms the established result that capitalization effects estimated from hedonic price models should be interpreted with care as they may combine preferences to avoid risk with changes in the implicit prices of flood risk (and other amenities) as people move heterogeneously into and out of neighborhoods (Kuminoff et al., 2010; Kuminoff and Pope, 2014).

7 Implications for Policy Reform

Public programs are paramount for disaster preparedness and recovery, yet continued calls for reform highlight concerns over their performance and fiscal costs (Michel-Kerjan, 2010; Deryugina, 2017). The National Flood Insurance Program (NFIP), in particular, has long faced scrutiny for the prevalence of discounted premiums and outdated flood maps (Michel-Kerjan, 2010; Michel-Kerjan and Kunreuther, 2011). Designated as a financially "high risk" program by the Government Accountability Office, historical payouts have exceeded premiums, with NFIP debt at $20.5 billion as of February 2018 and expected costs exceeding revenue by $1.4 billion annually (CBO, 2017; GAO, 2018). In addition, while flood risk maps are the primary source of flood risk information in the United States, almost two-thirds of the flood maps have not been updated in the past five years (Keller et al., 2017). Despite a recommendation from the Department of Homeland Security that the NFIP should improve its management of floodplain mapping, funding for flood map updates has been slow to follow and politically uncertain (OIG, 2017).

Moreover, previous NFIP reform attempts have been unsuccessful despite historical bipartisan support, due, in part, to concerns about insurance affordability if price supports end (DHS, 2018; Kousky, 2018). In July 2012, Congress passed, with bi-partisan support, the Biggert-Waters Flood Insurance Reform Act that slowly phased out some key price discounts, including grandfathering and discounts for pre-FIRM properties. However, this legislation was tempered in the eventual 2014 Homeowner Flood Insurance Affordability Act, which slowed the removal of pre-FIRM discounts and re-instated grandfathering through Congressional mandate.[38] Even with efforts to improve fiscal soundness, the affordability of flood insurance remains a key policy goal of the NFIP (Kousky and Kunreuther, 2014). A small but growing area of literature assesses the potential distributional consequences of the current the NFIP and affiliated programs (Bin et al., 2012, 2017; Kahn and Smith, 2017; Noonan and Sadiq, 2018).

7.1 Empirical Approach

We use the structural parameters recovered from the sorting model to assess how potential reforms to insurance prices and flood-risk information under the National Flood Insurance Program may lead to differential changes in household welfare and hazard exposure. First, we estimate the compensating variation across race and income groups for three simulated types of price changes, as well as the predicted reallocation of household types across flood-risk zones. The three price reforms include discontinuation of: (1) pre-FIRM insurance rates, a preferential rate structure at the property-level for housing stock built before the first flood insurance rate map (FIRM) was released in their community; (2) Community Rating System (CRS) discounts, a price reduction of up to 45% off flood insurance premium prices determined by flood activities at the community level; and (3) preferred rate grandfathering, a rule that allows properties with pre-existing flood insurance contracts to maintain preferential rates after new flood maps are released. [39] We note that there are no current efforts to remove or alter the CRS program. Pre-FIRM rate structures are currently being phased out and grandfathering was eliminated in 2012 and then brought back by act of Congress in 2014. The specific reforms that we consider are therefore highly relevant to the present

[38] Premium changes under the Biggert-Waters reform did not begin until 2013 and therefore do not affect our sample (FEMA, 2013a). For an assessment of the impact of the reforms on housing markets, see recent empirical work by, e.g., Gibson et al. (2017) and Indaco et al. (2018).

[39] Grandfathering benefits can be passed on to new property owners if flood insurance policies do not lapse.

policy discussion. Second, we utilize recent risk-map changes in Florida to assess the value of new risk-map information by comparing household welfare under up-to-date information (from new risk maps) versus welfare based on location choices made using previous risk maps. We then compare these benefits to the costs of map revisions.

We apply our sorting estimates to counterfactual exercises with the necessary assumption that the drivers of sorting and, importantly, heterogeneity remain fixed. The specific source and the time-invariance of heterogeneity are both important caveats in interpreting our predicted NFIP reform impacts. What drives heterogeneity is crucial for regulators to select the appropriate policy to counteract undesirable sorting outcomes.[40] The persistence of heterogeneity over time will also affect the distribution of welfare impacts. While our data do not allow us to identify the mechanism leading to heterogeneous sorting and the drivers may also change in the long-run, the results from the next set of empirical exercises demonstrate the scope for both overall and distributional impacts of reform, which is novel to this literature.[41] However, to the extent that the underlying sorting mechanisms are not immediately impacted by changes to the flood insurance program, our results would represent short- and medium-term policy impacts.

7.2 CRS and Pre-FIRM Price Supports

We first examine a counterfactual scenario in which NFIP premiums are set to risk-based rates after the removal of pre-FIRM and CRS price supports. Specifically, we calculate the compensating variation associated with these price changes following McFadden (1999):

$$CV_{PE} = \frac{1}{-\alpha_P}\Big(E \max_{j \in 1,\dots,J} \big\{V(p^1, risk, X, \epsilon)\big\} - E\ max \big\{V(p^0, risk, X, \epsilon)\big\}\Big) \tag{13}$$

[40] For example, should the policy lever be in an information campaign to adjust heterogeneous beliefs or to enforce realtor reporting of shown properties to prevent subtle forms of housing discrimination?

[41] Due to these limitations, we abstract from general equilibrium price and sorting changes in our counterfactual exercises, which would require additional strong assumptions on top of those already imposed.

Figure 3: Losses as a Percentage Income, Remove NFIP Price Supports

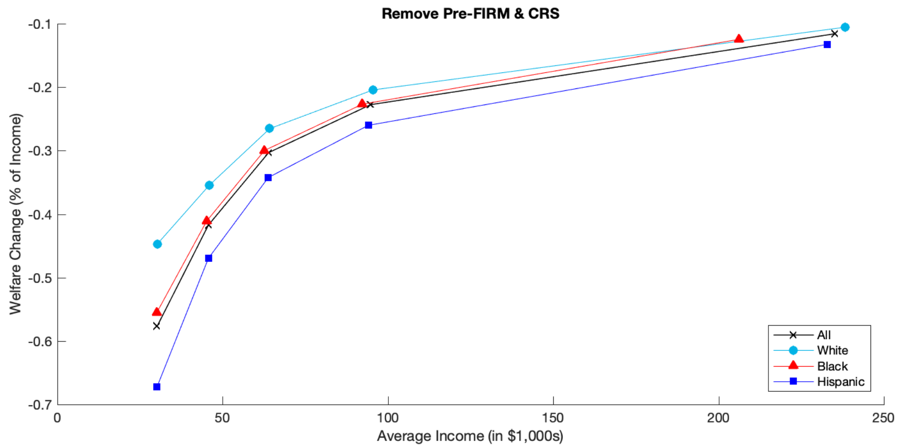


Table 5 presents the dollar value of impacts overall and by race and income groups upon removing all price supports other than grandfathering (top panel), pre-FIRM discounts only (middle panel), or CRS discounts (bottom panel). On average, removing both pre-FIRM and CRS discounts causes individuals to lose $209 per year (shown in the first cell of of the top panel).[42] The magnitude of losses increases with income. Not only do higher income households dislike flood risk, they are most likely to live in high risk areas due to their demand for coastal amenities. By race, the per household impacts on white households are generally lowest.

We view these impacts relative to income in Figure 3, which plots the losses as a percentage of the average income in each race and income-bin cell. While the highest income quintile groups lose the most in levels, removal of all price supports presents the largest burden for the lowest income groups (conditional on race).[43] This suggests that a policy that removes these discounts would be regressive. Compared to the impacts on whites, Hispanics experience somewhat larger impacts as a share of income.

Table 6 presents the change in distribution of race and income across the low risk zone X, the high inland flood risk zone A, and the high coastal flood risk zone V after removing

[42]Note that we currently assume that there are no costs to move as we do not observe homebuyers previous locations, so these welfare impacts can be seen as a lower bound on the true welfare costs (i.e., the true cost of the policy reform may be more negative).

[43]Corresponding figures that separately remove subsidies are similar (presented in Appendix Figure B.3).

both the CRS and pre-FIRM discounts. We see two important trends. First, price reforms would lead to fewer people in harm's way, with a migration towards low risk areas and away from higher risk A zone areas of between 2 and 14% and from coastal zone V areas, which (currently) receive large discounts, by between 2 and 54%. However, we also note that the higher risk zones will tend to become increasingly minority and low income as we predict a stronger out-migration of white and high-income groups from these areas. Thus, our results highlight that policy change could have large distributional impacts with potentially long lasting implications for disaster vulnerability, recovery, and fiscal policy (Arrow et al., 1996; Robinson et al., 2016; Banzhaf et al., 2019).

7.3 Grandfathering

An important feature of the NFIP program is grandfathering, which allows a household who would have faced a higher risk zone after a flood map update, the option to maintain its original flood insurance premium from a pre-existing policy. We can assess the losses from removing grandfathering from the following thought experiment: using a snapshot of updated flood maps as of 2016, we compare the change in household welfare from a map update with grandfathered rates to the change in welfare from the same map update except without grandfathered rates. In the case where a zone is mapped into a lower risk area, we retain the premium based on the lower risk level in the grandfathering scenario, assuming that households are given this option. We operationalize this by first mapping all houses according to flood insurance maps as of 2016. We then calculate the 2016 zone premiums according to current NFIP rates, CRS discounts, and community boundaries.

Table B.9 in the appendix presents a transition matrix for current zones to future zones. In general, most zone X houses remain as zone X after the map changes. However, there was a large share of zone A houses that would eventually be "downgraded" to zone X, and similarly, a small portion of V zone houses during our sample would become X or A zone houses.[44] We note that the direction of the change in flood zone can impact the size of

[44] A large share comes from Broward County (FIPS 12011). The V zone "downgrades" come primarily

losses from removing grandfathering. In addition, since some areas within a residential choice (partially defined by the original flood zone) could have experienced an update whereas other areas (within the same choice) do not, we replace the flood zone dummy from our model with the share of houses that are A or V zones (according to the updated flood map) in our counterfactual analysis.

Column (1) of Table 7 presents the impacts of removing grandfathering. Within each group, we then stratify by race, income, and the current zone (actually chosen). Overall, we find that all groups lose without the grandfathering option, where the average loss is \$305 per year. The average loss for white households (\$241) is lower than that for Hispanic households (\$391), and similar to that for black households (\$238). We again assess the regressivity of a policy that removes grandfathering in Figure 4(a). We divide the average welfare loss of each race and income group by the group's average income. As before, we see that removing grandfathering causes disproportionate burden on the lowest income groups. For example, the annual losses represent almost 0.75 percent of income for the bottom income quintile compared to less than 0.2 percent for the top quintile. These losses also disproportionately impact Hispanic households. At the bottom quintile, the difference in income share between white and Hispanic households is about 0.3 percentage points.

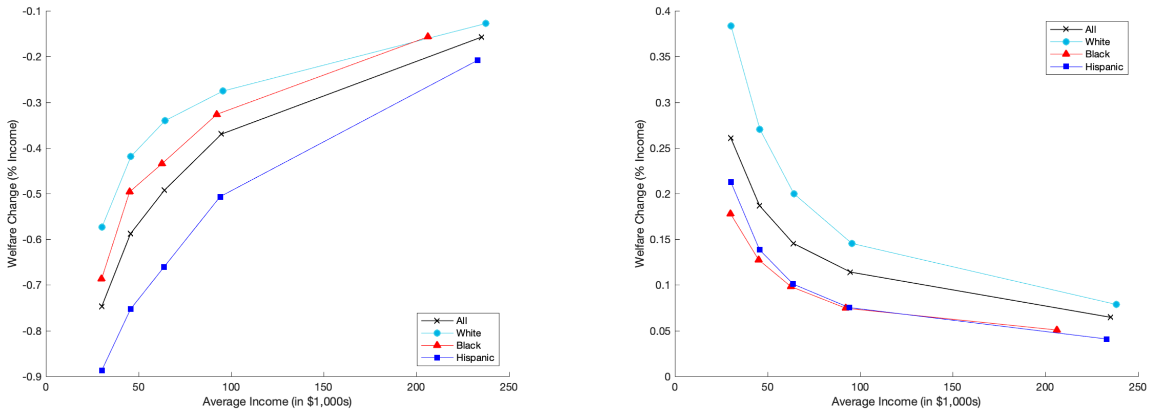
7.4 Flood Map Updates

Improvement of flood maps is critical for optimal household decision-making, yet funding to FEMA's flood mapping program is politically uncertain. To quantify the benefits of flood map revisions, we use the updated flood insurance maps to learn about the value of the information provided by these maps. We calculate the value of information following Leggett (2002) as

$$cv = \frac{1}{\alpha_P}\left[ln\left(\sum_{j,t} e^{V^1_{jt}}\right) - ln\left(\sum_{j,t} e^{V^0_{jt}}\right) - \sum_{j} \pi^0_{jt}\left(V^1_{jt} - V^0_{jt}\right)\right] \tag{14}$$

from Miami Dade county (FIPS 12086).

Figure 4: Welfare Changes from Map Updates as a Percentage of Income



(a) Remove Grandfathering

(b) Value of Information

The term, V_{jt}, again represents the conditional value of choosing choice j at time t. The superscripts, 0 and 1, on V_{jt} respectively index before and after flood map release, and π^0_{jt} is the probability of selecting neighborhood j given the pre-update (or old) flood maps. As before, α_P, refers to the marginal utility of income. The "log-sum" terms (or inclusive values) in equation (14) is the expected value from choosing optimally (less an Euler's constant, which is eventually differenced out); this is a result from taking the expectation of the maximum of the utility, i.e. $E\left[max_j\left\{e^{V_{jt}}\right\}\right]$, where the ϵ_{jt}'s have been integrated out based on their assumed extreme value distribution, similar to the measure of compensating variation used previously. The difference between the first two inclusive values in the brackets gives the change in welfare before and after the flood map update. The last term, derived by Leggett (2002), adjusts for any potential loss an individual might incur. It does so by giving more weight to alternatives that are only more attractive under the old information set, where the weight is the choice probability association with the old information set. Intuitively, cv is loss from making a sub-optimal decision that would have seemed optimal with the old map.

Column (2) of Table 7 presents the value of information as calculated based on equation (14). On average, there are positive gains from release of current flood map information, where the average value of the update is \$103 per year. In terms of a dollar value, maps provide the least value to black households ($60 compared to $144 and $70 for white and Hispanic households, respectively), and the highest value to high-income households ($152). However, plotting the benefits as a share of income as before in Figure 4(b), we find that information provision is progressive. On average, the benefits are 0.25 percent of income for households in the bottom quintile compared to 0.1% for the top quintile. Not only do flood maps have positive value, it potentially has a progressive impact across those who use the information. These results once again highlight how policy change can have important distributional consequences in the presence of heterogeneous sorting.[45]

8 Policy Discussion

Are these impacts large? Appendix Table B.10 presents a simple aggregation exercise with impacts scaled up to all households in the Miami-Dade, Port St. Lucie, Ft. Lauderdale Combined Statistical Area (CSA).[46] Recall that these distributional costs are not simply the sum of the increases in insurance premiums, but also account for the re-sorting that will occur in response to premium changes. We find that price reforms can have large distributional costs en masse. Removal of all three price supports would lead to a $353 million annual loss for affected homeowners in the CSA. The removal of pre-FIRM price supports alone, currently being phased out under the 2012 and 2014 reforms, costs approximately $131 million per year.

While we do not estimate a full benefit-cost analysis of removing NFIP price supports, there are several important societal benefits from bringing the program into fiscal balance based on how the program is financed.[47] One potentially large benefit is the resulting

[45]Our results assume equal access to and understanding of this information. As flood maps are easily available online and flood risk much be disclosed in the buying process, we assume this to be true but note that unequal access to or understanding of information across groups is an important area of future work.

[46]We also scale the cost estimates (down) by the fraction of flood insurance uptake across the region (29.1%). We do not scale the value of new flood map information down because the flood map information is freely and publicly available regardless of flood insurance status and therefore would benefit all residents.

[47]The NFIP funds its debt from three main sources: (1) cross-subsidization (i.e., higher rates) from other policies (CBO, 2017), (2) funds borrowed (plus affiliated interest payments – currently about $300 million per year) from the U.S. Treasury, and (3) (infrequent) NFIP debt cancellation through Congress. Thus, welfare benefits of fiscal soundness would depend on the allocation of savings to each group, and other factors

migration from higher- to lower-risk areas through the re-sorting process. From Table 6, we find a significant, albeit heterogeneous, shift from high risk flood zones, and especially the coastal V zone, to the low risk X zone. As historical damages from flood-related events have topped $1.18 billion since 2000 in our study area, fewer properties in harm's way would be expected to reduce future damages.[48] It also hints that migration could likely be an important (albeit costly) channel to mitigate climate risks. This resorting also highlights the importance of accounting for behavioral responses to policy change in estimating costs and benefits. For example, we find that household welfare costs from insurance price reforms are significantly lower relative to costs estimated assuming no behavioral (resorting) response. In particular, expected welfare loss experienced by these households are, on average, only 24 percent of the price discounts removed. Lastly, our results hint that the current policy is potentially incentivizing individuals to undertake more risk and is suggestive evidence of a moral hazard response to the price supports currently in place. While policy reform may well be a desirable goal, policy change could have large distributional and efficiency impacts that should be considered in designing reform solutions (Zeckhauser, 1981; Arrow et al., 1996; Robinson et al., 2016; Banzhaf et al., 2019).

Turning to the value of new flood map information, we find that flood risk map updates are valuable sources of information and are appealing from both a distributional and efficiency perspective. Aggregate benefits of new maps to residents of South Florida are an estimated $244 million per year. While no public record of the costs to revise the maps across the CSA exists, we estimate the average cost to update maps since 2000 has been approximately $4.8 million per county, implying a one-year benefit-to-cost ratio of 7.3 from new maps.[49] Given that large coastal counties are more expensive to map, even if costs were triple this estimate,

including the marginal cost of public funds (Browning, 1976).

[48] Figure calculated by the authors from the NCEI's Storm Events Database.

[49] According to the CBO (2017), the NFIP allocates approximately $200 million per year for flood map updates. Given the current effective map date for the more than 20,000 communities in the NFIP (available from FEMA's Community Status Book Report), we estimated the average annual cost per new county map from 2000 to 2018 to be $4.8 million. As there are seven counties within the CSA, we calculate 7.26 = $244 million/(4.8*7). In our conversations with map developers at the NFIP, they note that map costs can range greatly across the size and complexity of counties, with larger coastal counties being more expensive.

the benefit-cost ratio for a single year of map use would still exceed 2.4. Assuming the maps remain valid for five years and a 5% interest rate, the discounted present net benefits of these new map revisions is approximately $1.3 billion. More generally, the value of new flood maps is determined by many factors including the distribution of properties as well as the magnitude (and frequency) of map changes: all else equal, older or more outdated maps are more likely to have larger benefits from a revision since they would have greater inaccuracies relative to more recently developed maps. As flood risk remains a critical concern for Florida (Hallegatte et al., 2013), these results highlight the importance of high quality flood risk information for decision making.

9 Conclusion

This paper examines sorting over flood risk and the implications for policy reform. Using 2009 to 2012 housing sales data from Florida's Miami-Dade-Ft. Lauderdale-Port St. Lucie Combined Statistical Area, we build and estimate a residential sorting model with a boundary discontinuity design to recover heterogeneous sorting parameters that account for flood insurance price discounts. We then use our estimates to assess the distributional impacts of removing NFIP price supports, as well as to calculate the value of flood map updates.

We find clear evidence of heterogeneous sorting, specifically that low income and minority groups are more likely to sort into high flood risk areas. We also highlight how the presence of heterogeneous sorting has important policy implications. In particular, we note the need to account for behavioral responses in estimating the consequences of policy change, as policy reform can potentially trigger unwanted consequences that can fall more heavily on traditionally disadvantaged communities. While reform may well be a desirable policy goal, distributional impacts can have long lasting implications for disaster vulnerability, recovery, and fiscal policy (Arrow et al., 1996; Robinson et al., 2016; Banzhaf et al., 2019). In addition, we show how behavioral responses can mitigate the costs of policy reform through the resorting process. Lastly, our results reaffirm the importance of high quality risk information for households to make decisions. In addition to providing new valuation estimates for flood risk, our results shed light on the distributional impacts of natural disaster policy reform and its potential role in shaping disproportionate flood risk exposure in the United States.

References

Albouy, D., W. Graf, R. Kellogg, and H. Wolff (2016). Climate amenities, climate change, and american quality of life. *Journal of the Association of Environmental and Resource Economists 3*(1), 205–246.

Arrow, K. J., M. L. Cropper, G. C. Eads, R. W. Hahn, L. B. Lave, R. G. Noll, P. R. Portney, M. Russell, R. Schmalensee, V. K. Smith, et al. (1996). Is there a role for benefit-cost analysis in environmental, health, and safety regulation? *Science 272*(5259), 221–222.

Bakkensen, L. A. and L. Barrage (2017). Flood risk belief heterogeneity and coastal home price dynamics: Going under water? Technical report, National Bureau of Economic Research.

Banzhaf, H. S. and R. P. Walsh (2008). Do people vote with their feet? An empirical test of Tiebout's mechanism. *American Economic Review 98*(3), 843–863.

Banzhaf, S., L. Ma, and C. Timmins (2019). Environmental justice: The economics of race, place, and pollution. *Journal of Economic Perspectives 33*(1), 185–208.

Bayer, P., F. Ferreira, and R. McMillan (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy 115*(4), 588–638.

Bayer, P., R. McMillan, A. Murphy, and C. Timmins (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica 84*(3), 893–942.

Bayer, P. and C. Timmins (2007). Estimating equilibrium models of sorting across locations. *The Economic Journal 117*(518), 353–374.

Beltrán, A., D. Maddison, and R. J. Elliott (2018). Is flood risk capitalised into property values? *Ecological Economics 146*, 668–685.

Bernstein, A., M. T. Gustafson, and R. Lewis (2019). Disaster on the horizon: The price effect of sea level rise. *Journal of Financial Economics*.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics 25*(2), 242–262.

Bin, O., J. Bishop, and C. Kousky (2017). Does the national flood insurance program have redistributional effects? *The BE Journal of Economic Analysis & Policy 17*(4).

Bin, O., J. A. Bishop, and C. Kousky (2012). Redistributional effects of the national flood insurance program. *Public Finance Review 40*(3), 360–380.

Bin, O., J. B. Kruse, and C. E. Landry (2008). Flood hazards, insurance rates, and amenities: Evidence from the coastal housing market. *Journal of Risk and Insurance 75*(1), 63–82.

Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *Quarterly Journal of Economics 114*(2), 577–599.

BLS (2012). Consumer price index - South urban, Housing. Available at http://www.bls.gov/cpi/data.htm, (assessed April 2016).

Browning, E. K. (1976). The marginal cost of public funds. *Journal of Political Economy 84*(2), 283–298.

Cavallo, E., S. Galiani, I. Noy, and J. Pantano (2013). Catastrophic natural disasters and economic growth. *Review of Economics and Statistics 95*(5), 1549–1561.

CBO (2017). The national flood insurance program: Financial soundness and affordability. *Available online at https : //www.cbo.gov/publication/*53028.

Christensen, P. and C. Timmins (2018). Sorting or steering: Experimental evidence on the economic effects of housing discrimination. Technical report, National Bureau of Economic Research.

Depro, B., C. Timmins, and M. ONeil (2015). White flight and coming to the nuisance: can residential mobility explain environmental injustice? *Journal of the Association of Environmental and Resource Economists 2*(3), 439–468.

Deryugina, T. (2017). The fiscal cost of hurricanes: Disaster aid versus social insurance. *American Economic Journal: Economic Policy 9*(3), 168–98.

Deryugina, T., L. Kawano, and S. Levitt (2018). The economic impact of hurricane katrina on its victims: evidence from individual tax returns. *American Economic Journal: Applied Economics 10*(2), 202–33.

DHS (2018). An affordability framework for the national flood insurance program. *Available online at https : //www.fema.gov/media − library/assets/documents/*163171.

Dixon, L., N. Clancy, S. A. Seabury, and A. Overton (2006). The national flood insurance programs market penetration rate.

Epple, D. and H. Sieg (1999). Estimating equilibrium models of local jurisdictions. *Journal of Political Economy 107*(4), 645–681.

Fan, Q. and M. Davlasheridze (2016). Flood risk, flood mitigation, and location choice: evaluating the national flood insurance programs community rating system. *Risk Analysis 36*(6), 1125–1147.

Fan, Q., K. Fisher-Vanden, and H. A. Klaiber (2018). Climate change, migration, and regional economic impacts in the united states. *Journal of the Association of Environmental and Resource Economists 5*(3), 643–671.

FEMA (2013a). Biggert-waters flood insurance reform act of 2012 (bw12) timeline. *Available online at* $https://www.fema.gov/media-library-data/20130726-1912-25045-8239/bw_timeline_table_04172013.pdf$.

FEMA (2013b). National flood insurance condominium coverage. *Available online at* $https://www.fema.gov/media-library-data/20130726-1450-20490-0505/f029_condocov0410.pdf$.

FEMA (2017). National flood insurance program community rating system coordinators manual. *Available online at* $https://www.fema.gov/media-library-data/1493905477815-d794671adeed5beab6a6304d8ba0b207/633300_2017_CRS_Coordinators_Manual_508.pdf$.

Flood Smart (2016a). National flood insurance program. Available online at: $https://www.floodsmart.gov/floodsmart/$.

Flood Smart (2016b). Why does my mortgage lender require me to buy flood insurance? Available online at: $https://www.floodsmart.gov/floodsmart/pages/faqs/why-does-my-mortgage-lender-require-me-to-buy-flood-insurance.jsp$.

Gallagher, J. and D. Hartley (2017). Household finance after a natural disaster: The case of hurricane katrina. *American Economic Journal: Economic Policy 9*(3), 199–228.

GAO (1983). Siting of hazardous waste landfills and their correlation with racial and economic status of surrounding communities. Available online at: $http://archive.gao.gov/d48t13/121648.pdf$.

GAO (2018). National flood insurance program. *Available online at:* $https://www.gao.gov/key_issues/disaster_assistance/national-flood-insurance-program$.

Gibson, M., J. T. Mullins, and A. Hill (2017). Climate change, flood risk, and property values: Evidence from New York City. Technical report, Working Paper.

Hallegatte, S., C. Green, R. J. Nicholls, and J. Corfee-Morlot (2013). Future flood losses in major coastal cities. *Nature Climate Change 3*(9), 802–806.

Hallstrom, D. and V. Smith (2005). Market responses to hurricanes. *Journal of Environmental Economics and Management 50*(3), 541–561.

Harrison, D., G. T. Smersh, and A. Schwartz (2001). Environmental determinants of housing prices: The impact of flood zone status. *Journal of Real Estate Research 21*(1-2), 3–20.

Hausman, C. and S. Stolper (2019). Inequality, information failures, and air pollution. Technical report, Working Paper.

Hornbeck, R. (2012). The enduring impact of the american dust bowl: Short-and long-run adjustments to environmental catastrophe. *American Economic Review 102*(4), 1477–1507.

Indaco, A., F. Ortega, and S. Taspinar (2018). The effects of flood insurance on housing markets. *IZA Discussion Paper*.

Kahn, M. E. and V. K. Smith (2017). The affordability goal and prices in the national flood insurance program. Technical report, National Bureau of Economic Research.

Keller, M., M. Rojanasakul, D. Ingold, C. Flavelle, and B. Harris (2017). Outdated and unreliable: Femas faulty flood maps put homeowners at risk. *Available online at: https : //www.bloomberg.com/graphics/*2017 − *fema* − *faulty* − *flood* − *maps/*.

Klaiber, H. A. and N. V. Kuminoff (2013). *Equilibrium Sorting Models of Land Use and Residential Choice*. The Oxford Handbook of Land Economics.

Klaiber, H. A. and D. J. Phaneuf (2010). Valuing open space in a residential sorting model of the twin cities. *Journal of Environmental Economics and Management 60*(2), 57–77.

Kousky, C. (2018). Financing flood losses: A discussion of the national flood insurance program. *Risk Management and Insurance Review 21*(1), 11–32.

Kousky, C. and H. Kunreuther (2014). Addressing affordability in the national flood insurance program. *Journal of Extreme Events 1*(01), 1450001.

Kousky, C., B. Lingle, and L. Shabman (2016). Nfip premiums for single-family residential properties: Today and tomorrow. *Resources for the Future Policy Brief 16-10*.

Kousky, C. and L. Shabman (2014). Pricing flood insurance: how and why the nfip differs from a private insurance company. *Resources for the Future Discussion Paper*, 14–37.

Kuminoff, N. V. (2009). Decomposing the structural identification of non-market values. *Journal of Environmental Economics and Management 57*(2), 123–139.

Kuminoff, N. V., C. F. Parmeter, and J. C. Pope (2010). Which hedonic models can we trust to recover the marginal willingness to pay for environmental amenities? *Journal of Environmental Economics and Management 60*(3), 145–160.

Kuminoff, N. V. and J. C. Pope (2014). Do "capitalization effects" for public goods reveal the publics willingness to pay? *International Economic Review 55*(4), 1227–1250.

Kuminoff, N. V., V. K. Smith, and C. Timmins (2013). The new economics of equilibrium sorting and policy evaluation using housing markets. *Journal of Economic Literature 51*(4), 1007–62.

Leggett, C. G. (2002). Environmental valuation with imperfect information the case of the random utility model. *Environmental and Resource Economics 23*(3), 343–355.

Logan, J. R., Z. Xu, and B. J. Stults (2014). Interpolating us decennial census tract data from as early as 1970 to 2010: A longitudinal tract database. *The Professional Geographer 66*(3), 412–420.

Ma, L. (2019). Learning in a hedonic framework: Valuing brownfield remediation. *International Economic Review 60*(3), 1355–1387.

McFadden, D. (1978). *Modelling the choice of residential location*. Institute of Transportation Studies, University of California.

McFadden, D. (1999). To-pay in random utility models. *Trade, Theory, and Econometrics*.

Michel-Kerjan, E. and H. Kunreuther (2011). Redesigning flood insurance. *Science 333*(6041), 408–409.

Michel-Kerjan, E. O. (2010). Catastrophe economics: The national flood insurance program. *Journal of Economic Perspectives 24*(4), 165–86.

Mohai, P., D. Pellow, and J. T. Roberts (2009). Environmental justice. *Annual Review of Environment and Resources 34*, 405–430.

NFIP (2016). National flood insurance program flood insurance manual. Available online at: $https://www.fema.gov/media-library/assets/documents/127010$.

Noonan, D. S. and A.-A. A. Sadiq (2018). Flood risk management: exploring the impacts of the community rating system program on poverty and income inequality. *Risk Analysis 38*(3), 489–503.

OIG (2017). Fema needs to improve management of its flood mapping programs. *Available online at:* $https://www.documentcloud.org/documents/4066233-OIG-17-110-Sep17.html$.

Rabines, A. (2018). In miami beach, buyers and builders turn to elevated homes amid rising concerns of floods. *Available online at* $https://therealdeal.com/miami/2018/12/21/in-miami-beach-buyers-and-builders-turn-to-elevated-homes-amid-rising-concerns-of-floods/$.

Robinson, L. A., J. K. Hammitt, and R. J Zeckhauser (2016). Attention to distribution in us regulatory analyses. *Review of Environmental Economics and Policy 10*(2), 308–328.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *The Journal of Political Economy 82*(1), 34–55.

Shilling, J. D., C. Sirmans, and J. D. Benjamin (1989). Flood insurance, wealth redistribution, and urban property values. *Journal of Urban Economics 26*(1), 43–53.

Sieg, H., V. K. Smith, H. S. Banzhaf, and R. Walsh (2004). Estimating the general equilibrium benefits of large changes in spatially delineated public goods. *International Economic Review 45*(4), 1047–1077.

Sinha, P., M. L. Caulkins, and M. L. Cropper (2018). Household location decisions and the value of climate amenities. *Journal of Environmental Economics and Management 92*, 608–637.

Smith, A. B. and R. W. Katz (2013). US billion-dollar weather and climate disasters: Data sources, trends, accuracy and biases. *Natural Hazards 67*(2), 387–410.

Smith, V. K., J. C. Carbone, J. C. Pope, D. G. Hallstrom, and M. E. Darden (2006). Adjusting to natural disasters. *Journal of Risk and Uncertainty 33*(1-2), 37–54.

Staiger, D. and J. H. Stock (1997). Instrumental variables regression with weak instruments. *Econometrica 65*(3), 557–586.

Strobl, E. (2011). The economic growth impact of hurricanes: Evidence from US coastal counties. *Review of Economics and Statistics 93*(2), 575–589.

Taylor, D. E. (2000). The rise of the environmental justice paradigm: Injustice framing and the social construction of environmental discourses. *American Behavioral Scientist 43*(4), 508–580.

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of political economy 64*(5), 416–424.

Timmins, C. (2007). If you cannot take the heat, get out of the cerrado... recovering the equilibrium amenity cost of nonmarginal climate change in brazil. *Journal of Regional Science 47*(1), 1–25.

Tra, C. I. (2010). A discrete choice equilibrium approach to valuing large environmental changes. *Journal of Public Economics 94*(1), 183–196.

Turner, M. A., S. L. Ross, G. C. Galster, J. Yinger, et al. (2002). Discrimination in metropolitan housing markets: National results from phase i hds 2000. *Washington, DC: The Urban Institute*.

Walsh, R. (2007). Endogenous open space amenities in a locational equilibrium. *Journal of Urban Economics 61*(2), 319–344.

Zeckhauser, R. (1981). Preferred policies when there is a concern for probability of adoption. *Journal of Environmental Economics and Management 8*(3), 215–237.

Zhang, L. (2016). Flood hazards impact on neighborhood house prices: A spatial quantile regression analysis. *Regional Science and Urban Economics 60*, 12–19.

Table 1: Differences in Mean Attributes by Zone (A vs. X)

Distance from Boundary	Full Sample		5km		3km		1km		0.5km		0.3km	
	Mean Δ	T-stat.	Mean Δ	T-stat.	Mean Δ	T-stat.	Mean Δ	T-stat.	Mean Δ	T-stat.	Mean Δ	T-stat.
Price	20450.21	1.89	20199.20	1.51	19752.20	1.46	10097.14	0.81	2951.64	0.25	-5281.57	-0.57
Single Family	-0.18	-8.31	-0.18	-6.92	-0.18	-6.88	-0.16	-6.41	-0.13	-5.84	-0.11	-4.50
Condominium	0.18	8.51	0.18	7.15	0.18	7.08	0.16	6.57	0.13	6.00	0.11	4.53
Age	-0.35	-0.30	-0.44	-0.20	-0.23	-0.10	-0.48	-0.30	0.16	0.13	0.63	0.60
Pre-FIRM	0.11	3.75	0.11	1.97	0.12	2.19	0.10	2.61	0.12	3.65	0.13	4.50
ICR within 3km	-0.14	-2.07	-0.15	-1.68	-0.14	-1.61	-0.14	-1.65	-0.14	-1.94	-0.17	-2.59
School Quality	-4.11	-3.29	-4.09	-1.19	-4.57	-1.37	-3.60	-1.66	-3.12	-1.90	-3.20	-2.25
Dist. to Park	-3.81	-4.54	-3.72	-3.33	-3.60	-3.27	-2.96	-3.11	-2.65	-3.03	-2.17	-2.82
Dist. to River	34.72	8.01	34.09	3.79	33.57	3.80	26.40	4.23	23.73	4.77	22.27	5.00
Share Hispanic ('90)	0.06	3.09	0.06	1.97	0.06	2.26	0.05	2.45	0.05	2.85	0.06	3.24
Share Black ('90)	-0.02	-1.89	-0.02	-1.62	-0.01	-1.47	-0.01	-1.42	0.00	-0.47	0.01	1.13
Per Capita Income ('90)	821.17	0.92	771.04	0.95	524.47	0.63	278.21	0.32	-109.98	-0.12	-399.15	-0.45

Note: The table assess differences in mean attributes between A and X flood zones (i.e. A-X) for the full sample (columns 1 and 2) and sub-samples of houses at various distances from the nearest boundary (columns 3-10). T-statistics test the null hypothesis that the difference in mean attributes for a specified sample is 0. All t-tests are clustered at the boundary ID level.

Table 2: Hedonic Regressions

Panel A. Progression of Controls

Dep. Var.:		Add Flood Controls		BFE (<1km)
Annual Rent	(1)	(2)	(3)	(4)
SFHA	-2,203***	-1,642***	-1,120***	-658.6***
	(80.15)	(91.70)	(85.84)	(100.1)
Elevation		-477.8***	-168.0***	-263.8***
		(33.89)	(31.81)	(47.82)
Relative BFE		460.6***	904.5***	1,081***
		(16.47)	(16.30)	(22.21)
Distance to Coast:				
<0.1km			14,392***	11,022***
			(268.4)	(400.5)
<0.5km			11,854***	7,948***
			(177.5)	(295.2)
<1km			9,908***	6,663***
			(192.9)	(263.0)
<2km			6,000***	5,022***
			(148.5)	(199.5)
<3km			3,521***	2,539***
			(141.9)	(177.7)
<4km			2,269***	808.2***
			(152.6)	(196.2)
<5km			2,161***	255.3
			(150.6)	(186.3)
Observations	48,174	48,174	48,174	31,601

Panel B. Alternative Specifications

	Other BFE Distance Buffers			Ignore Price Supports
Sample Restriction:	<800m	<500m	<300m	None
SFHA	-657.6***	-542.7***	-681.6***	-18.84
	(103.5)	(113.9)	(126.0)	(83.21)
Observations	29,044	23,194	17,594	48,174

Note. All regressions include county and year fixed effects. The (omitted) base group for flood zone is zone X. A set of controls that are consistent with the sorting model are included but not shown, including house characteristics (single, condo, age, Pre-FIRM) and neighborhood characteristics (ICR's within 3km, school quality, distance to the nearest park and river, 1990 Census share Hispanic, share black, and median income.). Column (4) of Panel A limits the sample to houses within 1 kilometer of a flood boundary and includes boundary fixed effects. Panel B re-estimates the boundary fixed effects specifications with different buffer distances (columns 1-3), and re-estimates the specification in column (3) of panel A but ignores price supports (column 4).

Table 3: Mean Utility Decomposition (J=8,618)

	Price IV		OLS	
	est.	s.e.	est.	s.e.
Rent (α_p)	-1.739	0.302	-0.005	0.001
SFHA	-1.235	0.379	-0.085	0.039
BFE Assigned	-14.446	2.649	-0.294	0.076
Elevation	-0.275	0.159	0.144	0.014
Coast <0.1km	15.284	3.294	-1.842	0.159
Income	0.273	0.059	-0.062	0.002
Black	-14.035	2.177	-2.264	0.118
Hispanic	-15.611	2.856	-2.247	0.158
School Quality	11.318	5.514	-1.482	0.413
Dist. to River	-1.570	0.530	-0.131	0.039
ICR within 3km	0.078	0.077	0.069	0.010
Dist. to Park	-0.415	0.081	0.002	0.005
Single	4.220	0.779	0.755	0.039
Condo	-11.800	2.299	0.644	0.047
Pre-FIRM	-8.149	1.404	0.148	0.019
Distance to Coast:				
<0.5km	12.111	2.340	-0.535	0.054
<1km	10.525	2.006	-0.567	0.049
<2km	6.268	1.296	-0.578	0.039
<3km	2.526	0.709	-0.546	0.036
<4km	0.143	0.466	-0.536	0.039
<5km	-1.170	0.416	-0.404	0.037

Note. The estimates from the mean utility decomposition are in "utils" and will be converted to a dollar value using the coefficient α_p on Rent (in 2010 \$1,000's) in Table 4. The specification includes county and year fixed effects, which are not shown. All distances are in kilometers, and elevation is in meters. Standard errors are bootstrapped using 500 sample draws with replacement.

Table 4: Heterogeneous Parameter Estimates for Flood Zone

	Include Discount		Ignore Discount		Income (in $1,000's)	
Base Group	est.	s.e.	est.	s.e.	mean	s.d.
White, Quintile 1	-710.49	218.11	818.34	169.15	30.28	5.89
Relative to Base Group	est.	s.e.	est.	s.e.	mean	s.d.
Black	229.25	29.40	257.18	32.99	53.85	57.17
Hispanic	91.74	20.96	102.92	23.51	86.50	134.08
Quintile 2	-15.94	24.12	-17.89	27.06	45.69	4.30
Quintile 3	-31.00	24.60	-34.78	27.60	63.90	6.39
Quintile 4	-62.65	25.43	-70.29	28.53	94.74	12.59
Quintile 5	-198.03	27.12	-222.16	30.42	235.17	245.23

Note. Base group estimates (for white/Asian households in the first income quintile) are recovered from stage 2, and all standard errors are bootstrapped. The heterogeneous parameter estimates (by race/ethnicity and income group) are recovered from stage 1. All estimates have been converted to a (real 2010) dollar value using the estimated coefficient on rent and represent an annual (flow) MWTP per household. Estimates for non-base group categories should be added to the base group estimate to recover the preference for that group. The unconditional average income for the white/Asian group is $111 thousand. Average house price is $225,434.

Table 5: Impact of Removing Price Supports by Race and Income (2010 $USD)

All	Overall	Q1	Q2	Q3	Q4	Q5
Overall	-209.31	-174.07	-190.68	-193.73	-215.65	-272.44
White	-191.67	-135.33	-162.86	-170.45	-195.12	-251.36
Black	-184.18	-167.02	-185.65	-187.83	-209.29	-257.07
Hispanic	-234.19	-202.73	-214.60	-218.48	-244.70	-308.69
Pre-FIRM	Overall	Q1	Q2	Q3	Q4	Q5
Overall	-191.29	-157.01	-172.40	-176.45	-198.09	-252.56
White	-175.39	-122.50	-147.12	-155.23	-178.85	-232.15
Black	-166.58	-149.73	-167.63	-170.63	-191.66	-237.64
Hispanic	-214.23	-182.93	-194.21	-199.13	-225.38	-287.64
CRS	Overall	Q1	Q2	Q3	Q4	Q5
Overall	-96.04	-78.84	-85.69	-83.74	-97.46	-134.47
White	-89.88	-58.36	-74.37	-76.22	-90.56	-125.31
Black	-94.89	-85.77	-96.73	-94.76	-107.11	-138.60
Hispanic	-102.89	-89.60	-91.20	-88.85	-105.69	-149.60

Note. Table calculates the compensating variation required after removing all or one of the price supports under the NFIP by race and/or income quintile and represent an annual (flow) welfare change per household in real 2010 $USD.

Table 6: Percent Changes in Race/Income Distribution by Zone

All	Zone X	Zone A	Zone V
White	11.55	-11.74	-54.36
Black	2.24	-2.32	-2.01
Hispanic	13.86	-14.27	-29.86
All	Zone X	Zone A	Zone V
Q1	4.70	-4.87	-4.03
Q2	4.89	-5.06	-4.90
Q3	4.84	-4.98	-10.11
Q4	5.52	-5.64	-19.45
Q5	7.71	-7.77	-47.75

Note. Table aggregates the changes in predicted shares after all NFIP price supports are removed.

Table 7: Values from Map Updates (2010 $USD)

Group	(1) Remove Grandfathering	(2) Value of Better Risk Information
Overall	-305.87	103.38
Zone X	-304.12	102.34
SFHA	-307.69	104.46
White	-241.13	144.09
Black	-238.29	59.75
Hispanic	-391.28	70.47
Q1	-225.37	78.72
Q2	-268.53	85.35
Q3	-315.03	92.98
Q4	-350.40	108.07
Q5	-370.16	151.83

Note. Column 1 of this table provides the compensating variation required from removing the grandfathering option given the current flood map updates from the 2016 FIRMs and represent an annual (flow) welfare change per household in real 2010 $USD. Column 2 then calculates the value of this information.

Online Appendix

Appendix A: Data Sources and Construction

We begin with all arms-length sales for owner-occupied residential properties from the Miami-Dade, Port St. Lucie, Fort Lauderdale Combined Statistical Area (CSA) from 2009 to 2012. The data are provided by Dataquick, Inc., and include information on selling price, date of sale, numbers of bedrooms and bathrooms, and mortgage information. Structural characteristics are available as recorded from county assessor offices; those that we use for flood premium calculation are: exact latitude and longitude location, year built, house type (i.e. single family, mobile, condominium), and mortgage loan value (in thousands). To characterize the neighborhoods in which houses are located, we use GIS shapefiles obtained from the Yale University Map Department to map each house to nearby spatial amenities, including the distances to the nearest park, river, and coast. To control for local environmental quality, we join each house with levels of industrial contaminants, as operationalized by the count of sites within 3 kilometers of the property, in the year of property sale, that are listed on Florida's Institutional Controls Registry (ICR). ICR data are available from the Florida Department of Environmental Protection, and include brownfields, superfunds, solid waste sites, and those from various environmental programs such as Resource Conservation and Recovery Act, Comprehensive Environmental Response Compensation and Liability Act and statewide programs.[50] Public school quality has also been documented as an important amenity over which people sort (Black, 1999; Epple and Sieg, 1999; Bayer et al., 2007), so we include Florida Comprehensive Assessment Test (FCAT) scores for each school district, in each year, from the Florida Department of Education to proxy for school quality. The FCAT performance score evaluates achievement in the categories of reading, mathematics, science, and writing, out of a combined maximum of 400 possible points. Finally, as prospective buyers may have preferences over their future neighbors, we collect tract-level per-capita income and race/ethnicity population shares. As these neighborhood demographics

[50] We do not count drycleaning and sites that could not be categorized.

are endogenously determined based on residents' locational choices, we use predetermined neighborhood demographics from the 1990 census, which are mapped to 2010 census tracts using the Longitudinal Tract Data Base (Logan et al., 2014).

Digitized Flood Insurance Rate Maps (FIRMs), accessed from FEMA's Map Services Center and the Florida Geographic Data Library, are used to assign the underlying flood risk level to each house (in all zones) based on the FIRM that was in effect as of the property's sale date. We first map all houses to their flood zone (including a Base Flood Elevation (BFE), if determined in the FIRM), and then combine this information with available structural characteristics from the sales data (e.g. year of sale and single family or 2-4 family house) and the year a house was built. Base Flood Elevation is the computed elevation to which floodwater is anticipated to rise (NFIP, 2016), and the flood zones are A, V, or X.[51] This information is sufficient to determine the effective premium rate that each house should face, based on the NFIP Technical Manual (NFIP, 2016). Figure B.1 presents an example of an NFIP rate assignment. Owners in Special Flood Hazard Areas with such lender flood-insurance requirements must purchase an amount of coverage that is the lesser of (1) the outstanding principal balance of the loan, (2) the maximum NFIP coverage limit ($250,000 for residential buildings), or (3) the total insurable value of the property (Flood Smart, 2016a).[52] As an example, suppose a single-family dwelling in any of the A zones that was built before the NFIP has a loan of $150k. Assuming the property has no basement, the "basic" premium rate is $1.21 per $100 of building coverage for the first $60k of coverage, and then $1.11 per $100 of coverage using the "additional" rate for building coverage beyond $60k, up to $250k. The annual premium would be calculated as $\$1725 = \$1.21 \times \$600 + \$1.11 \times (\$1500 - \$600)$.

[51]Included in the A zone is A, AO, AH, zones A1-A30, AE, A99, AR, AR/AE, AR/AO, and AR/A1-A30. Included in the V/VE zones are zones VE, and V1-V30. Zone B and X shaded designates a flood risk of more than 0.2 percent but less than 1 percent per year. Zone C and X unshaded designates a risk of flood averaging less than 1 in 500 years. We group zones B, C, X shaded, and X unshaded together, termed in this paper as "Zone X", to designate flood risk of less than 1 percent per year.

[52]These provisions were introduced in the Flood Disaster Protection Act of 1973 and later strengthened in 1994 with the Flood Insurance Reform Act. In conversations with real estate agents and insurance agents, there is high compliance with insurance mandates when a home is purchased through a mortgage. Properties may later fall out of compliance if insurance is dropped at a later date. Non-mandatory insurance is available to homes outside of the SFHA, where flood damage could still occur albeit with low probability.

If the loan value was \$300k, then the coverage required would be capped at \$250k, making the calculated annual premium equal to $\$2836 = \$1.21 \times \$600 + \$1.11 \times (\$2500 - \$600)$.

Three structural characteristics, required for assigning the building premium rate, are missing from our housing data. These are (1) for condominiums, the size of the building in which a condominium is located (e.g. high rise versus low rise); (2) whether the property contains a basement or enclosure, and (3) the elevation of the lowest floor. We make a few necessary assumptions in the face of these data limitations, namely that (1) unit owners of condominiums purchase individual policies, rather than being covered through their homeowner's association (HOA), (2) houses have no basements or enclosures, (3) the lowest floor elevation of a building is 0, equal to the ground in elevation (i.e., not raised above or lowered below), and (4) homes have an elevation certificate if required by county ordinance. We provide a few remarks on these necessary assumptions and their implications.

First, by assuming that condominium owners purchase individual flood insurance policies, we consider all condominium units as single family residences, whereas the policy rates for the alternative option, obtaining coverage through the HOA, depends on the number of units in the building (i.e., data we do not have). Policies through an HOA generally provide greater coverage for a lower premium (FEMA, 2013b), so our assumption will tend to overstate the flood risk internalized by these households and attenuate our estimated preferences to avoid risk.

Second, unobserved basements could alter the coverage premium greatly, which would also cause us to miscalculate the price discount. In practice, however, basements are not very common in the South Atlantic states. Data from the Census Survey of Construction Microdata Files, which include annual single-family housing starts, completion, and/or sales in the U.S., at the Census Division level, show that approximately 82 percent of new construction in the New England and Mid-Atlantic states between 1999 and 2012 includes basements, compared to only 20 percent for South Atlantic states (as far north as Delaware).[53] The proportion for Florida is even lower, given that this coastal state has predominantly a limestone substrate

[53] Statistics are based on the authors' calculations.

and a shallow water table, making basement construction costly and difficult. A search of Realtor.com, a real-estate listings website, reveals that only 28 out of 25,186 homes listed for sale in Miami-Dade County in December of 2018 are listed as having a basement. Of the subset of houses in our data with basement information, only 0.04 percent have a basement.

Third, for a subset of homes in the SFHA,[54] the (post-FIRM) premium rate depends on the elevation difference, which is the difference between the elevation of the lowest floor of the building and the base flood elevation that is assigned to the building's FIRM. In our sample, 4.7 percent of houses belong to these zones and are post-FIRM. If houses have actually been raised, then our assumption that the lowest-floor elevation is the same as the surface elevation would cause us to over-estimate the relevant flood-insurance premium. In practice, however, homes elevated high on pylons are not common in South Florida, unlike other areas along the Gulf Coast. In addition, some localities have zoning ordinances specifically limiting people's ability to build above ground elevation. For example, until July 2018, a local height restriction ordinance in Miami Beach forced homes to have their first floor no more than seven feet above ground level (Rabines, 2018).

A final assumption we make is that all houses have elevation certificates except for a subset of houses built before 1995 in Miami-Dade County and before 1992 in Martin County, when these counties did not require elevation certificates. The post-FIRM premium rate for properties in AO and AH zones, which constitute 19.8% of our sample, additionally depends on whether a property has an elevation certificate. We test our model by alternatively assuming that no property had an elevation certificate, and results are robust to this assumption.

With the appropriate NFIP rate, we next recover the relevant insurance premium by applying this rate to the "building coverage." The building coverage amount is set as either the recorded loan amount or $250,000 (whichever is lower), given the NFIP coverage limits described above.[55] We assume no coverage for dwelling contents is purchased, because

[54] Specifically, the areas are un-numbered A, AE and VE zones.

[55] This assumes that homeowners purchase the minimum required building coverage. In our case, because our sales data are matched to mortgage applications in HMDA, and HMDA regulates reporting for federally regulated and backed lenders, SFHA houses in our data are required to cover the remaining principal on their mortgage, which is the initial loan amount as recorded in the application at the time of property sale.

coverage for contents is not required. In addition, the value of the home's contents would not be capitalized into the home's price. Table B.3 presents summary statistics for our calculated coverage and insurance premium. Based on our assumptions, the annual covered amount is \$159,664 on average, with a median of \$154,982. The full premium calculated prior to any discounts is, on average, \$2,113 per year, with a median of \$808. The pre-FIRM discounts, which are built into the NFIP rates, then provide an average discount of almost \$1,000 relative to the full premium.

The next step in calculating flood insurance premiums for properties is to incorporate the CRS program discounts. NFIP lists the cities and counties that participate in the CRS, as well as their discount rates.[56] Participating "communities," as defined by county or city limits, can receive premium discounts for its properties by engaging in community-level flood mitigation or education projects. The Program grants discounts to communities based on 19 different activities that can reduce insurance premiums (NFIP, 2016), where activities that do more to mitigate flood risk earn more "credits" toward receiving discounts. These discounts can range from 5 percent to as much as 45 percent depending the type of CRS activity. In theory, this can alter the flood risk faced by the homeowner, a point to which we return later. We match each house to the appropriate CRS community using GIS. We find that houses in our sample receive CRS discount rates of between 0 and 25 percent, with an average of 12 percent.

To implement our flood risk boundary discontinuity design, we also map each property to the closest flood zone boundary using Geographic Information Systems: we first split FIRM flood map polygon boundaries into segments, which are assigned a unique identifier, and then we find the closest (in terms of distance) segment to each house. Neighborhood choices, used in the sorting model, are represented by a certain type of house within a census tract (to be described in the model section). In order to assign neighborhood choices with the nearest flood zone boundary, we assign the boundary identifier of the choice to be that of the house closest to any boundary of all houses in that choice set.

[56]In Florida, county CRS rates cover rural areas within the county.

After characterizing housing properties, we lastly recover the race and income of buyers in our sales data so that we can categorize households into different "types." We merge the sales data with mortgage applications data, collected under the Home Mortgage Disclosure Act (HMDA),[57] which include (self-reported) race and income for the primary mortgage applicant. The merge is based on loan information present in both of the data sources, and we follow the procedure outlined in Bayer et al. (2016). Key matching variables include the census tract of the property of interest, the application date, the loan amount, and the name of the lending institution. Household "types" are defined by race and income, where income is categorized into bins based on quintiles of the observed income distribution.

Some features of data availability and processing affect the final estimating sample. First, we lose Palm Beach County because no digitized flood map was available at the time of our analysis.[58] Second, we cannot match every property sale to a mortgage application and therefore cannot determine the race and income of homebuyers for a subset of transactions. The HMDA data allow us to recover matches for 47% of all housing transactions. We can, however, compare the race and income distributions for our merged data to the analogous distributions for owner-occupied housing from the 2013 American Housing Survey of Miami-Ft. Lauderdale-Hollywood, Florida, as shown in Table B.4. Compared to owner-occupied properties surveyed in the AHS, our HMDA-matched sample is quite similar, especially for the white, Asian, and Pacific Islander groups, which differ by less than 0.1%. Our HMDA sample has slightly fewer black households (11.9% versus 13.9%) and slightly more Hispanic households (38.9% versus 36.7%) relative to the AHS, but overall, the two samples are closely comparable. Median income for our sample of homeowners, at $64,000, is somewhat higher than median income for the general owner-occupied population, at $56,000. Our final sample includes 48,174 individual house sales between 2009 and 2012 across six counties and 953

[57]Enacted by Congress in 1975, HMDA requires financial institutions, including depository institutions (e.g. banks and credit unions) and for-profit mortgage lending institutions, to make loan data publicly available. For reporting requirements in 2010, see https://www.ffiec.gov/hmda/reportde2010.htm.

[58]In a phone interview, the NFIP representative described the digitization process as being prioritized by population density. However, these counties are more densely populated than other counties in Florida. Thus, the rationale for the omission is unclear.

census tracts in Florida. The six counties are Miami-Dade, Broward, St. Lucie, Martin, Indian River, and Okeechobee.

Appendix B: Additional Figures and Tables

Figure B.1: NFIP Premium Rate Example

ANNUAL RATES PER $100 OF COVERAGE (Basic/Additional)

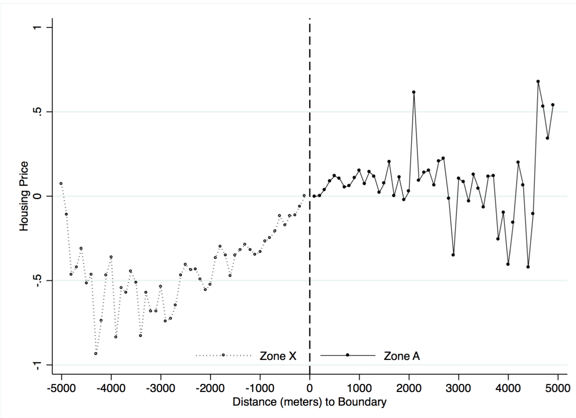
FIRM ZONES A, AE, A1–A30, AO, AH, D[3]

	OCCUPANCY	SINGLE FAMILY		2-4 FAMILY		OTHER RESIDENTIAL		NON-RESIDENTIAL BUSINESS[4]		OTHER NON-RESIDENTIAL[4]	
		Building	Contents	Building	Contents	Building	Contents	Building	Contents	Building	Contents
BUILDING TYPE	No Basement/Enclosure	1.21 / 1.11	1.52 / 1.99	1.21 / 1.11		1.21 / 2.34		1.32 / 2.46		1.32 / 2.46	
	With Basement	1.29 / 1.64	1.52 / 1.67	1.29 / 1.64		1.21 / 1.95		1.39 / 2.40		1.39 / 2.40	
	With Enclosure[5]	1.29 / 1.96	1.52 / 1.99	1.29 / 1.96		1.29 / 2.44		1.39 / 3.04		1.39 / 3.04	
	Elevated on Crawlspace	1.21 / 1.11	1.52 / 1.99	1.21 / 1.11		1.21 / 2.34		1.32 / 2.46		1.32 / 2.46	
	Non-Elevated with Subgrade Crawlspace	1.21 / 1.11	1.52 / 1.67	1.21 / 1.11		1.21 / 2.34		1.32 / 2.46		1.32 / 2.46	
	Manufactured (Mobile) Home[6]	1.21 / 1.11	1.52 / 1.99					1.32 / 2.46		1.32 / 2.46	
CONTENTS LOCATION	Basement & Above[7]				1.52 / 1.67		1.52 / 1.67		2.59 / 4.12		2.59 / 4.12
	Enclosure & Above[8]				1.52 / 1.99		1.52 / 1.99		2.59 / 4.93		2.59 / 4.93
	Lowest Floor Only – Above Ground Level				1.52 / 1.99		1.52 / 1.99		2.59 / 2.16		2.59 / 2.16
	Lowest Floor Above Ground Level and Higher Floors				1.52 / 1.39		1.52 / 1.39		2.59 / 1.85		2.59 / 1.85
	Above Ground Level – More Than 1 Full Floor				.35 / .12		.35 / .12		.24 / .12		.24 / .12
	Manufactured (Mobile) Home[6]								2.59 / 2.16		2.59 / 2.16

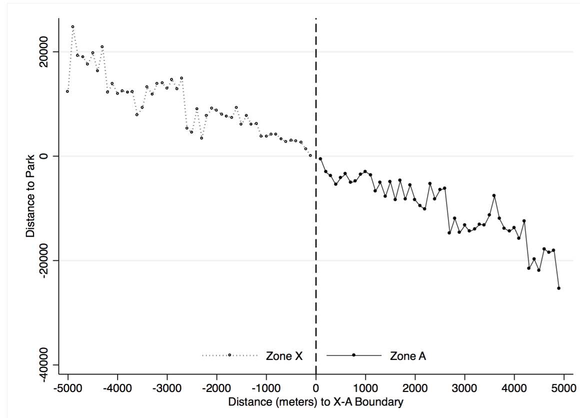
FIRM ZONES V, VE, V1–V30

	OCCUPANCY	SINGLE FAMILY		2-4 FAMILY		OTHER RESIDENTIAL		NON-RESIDENTIAL BUSINESS[4]		OTHER NON-RESIDENTIAL[4]	
		Building	Contents	Building	Contents	Building	Contents	Building	Contents	Building	Contents
BUILDING TYPE	No Basement/Enclosure	1.57 / 2.79	1.96 / 4.78	1.57 / 2.79		1.57 / 5.16		1.75 / 5.97		1.75 / 5.97	
	With Basement	1.69 / 4.15	1.96 / 4.05	1.69 / 4.15		1.69 / 7.69		1.85 / 8.85		1.85 / 8.85	
	With Enclosure[5]	1.69 / 4.89	1.96 / 4.77	1.69 / 4.89		1.69 / 8.57		1.85 / 9.88		1.85 / 9.88	
	Elevated on Crawlspace	1.57 / 2.79	1.96 / 4.78	1.57 / 2.79		1.57 / 5.16		1.75 / 5.97		1.75 / 5.97	
	Non-Elevated with Subgrade Crawlspace	1.57 / 2.79	1.96 / 4.05	1.57 / 2.79		1.57 / 5.16		1.75 / 5.97		1.75 / 5.97	
	Manufactured (Mobile) Home[6]	1.57 / 8.75	1.96 / 4.77					1.75 / 16.83		1.75 / 16.83	
CONTENTS LOCATION	Basement & Above[7]				1.96 / 4.05		1.96 / 4.05		3.41 / 10.45		3.41 / 10.45
	Enclosure & Above[8]				1.96 / 4.77		1.96 / 4.77		3.41 / 11.29		3.41 / 11.29
	Lowest Floor Only – Above Ground Level				1.96 / 4.77		1.96 / 4.77		3.41 / 9.46		3.41 / 9.46
	Lowest Floor Above Ground Level and Higher Floors				1.96 / 4.19		1.96 / 4.19		3.41 / 8.17		3.41 / 8.17
	Above Ground Level – More Than 1 Full Floor				.59 / .51		.59 / .51		.57 / .73		.57 / .73
	Manufactured (Mobile) Home[6]								3.41 / 15.73		3.41 / 15.73

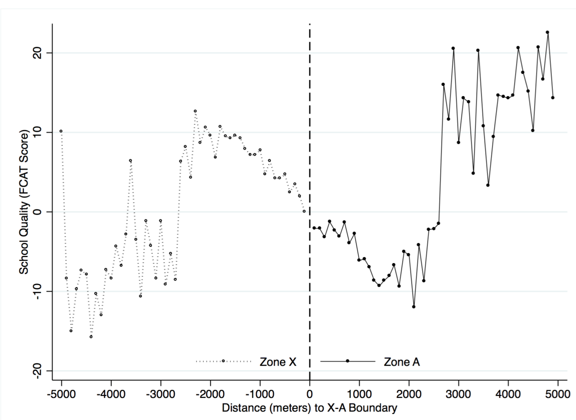
Figure B.2: Price and Spatial Amenities by Distance to Flood Boundary



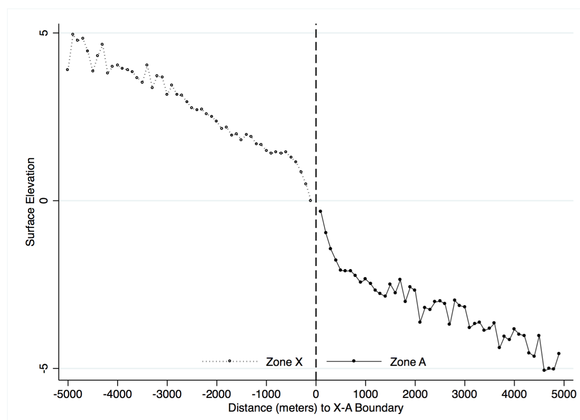
(a) Housing Price



(b) Distance to Nearest Park



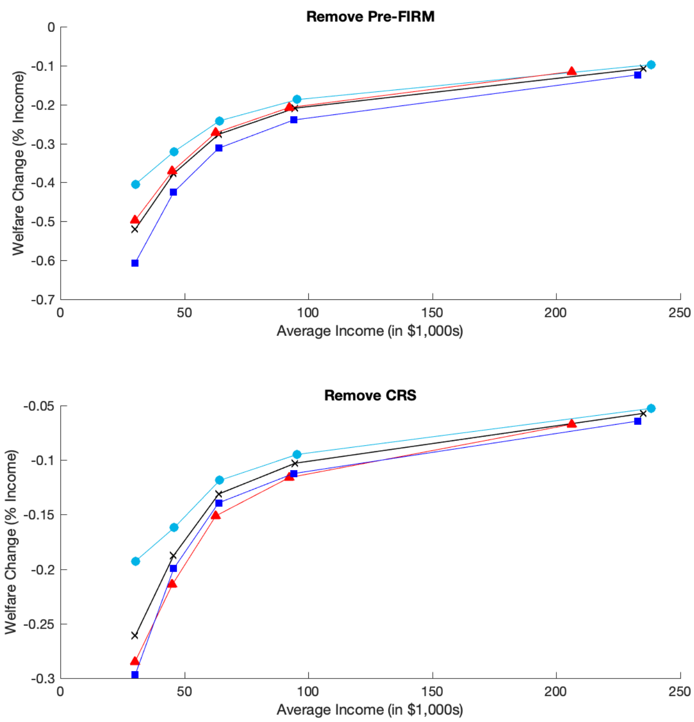
(c) School Quality



(d) Surface Elevation

Note. Each figure plots the coefficients from a regression of some attribute against distance-to-flood boundary dummy variables at 100-meter increments from the X zone (on the left) to the A zone (on the right). All points are normalized to the 100-meter distance on the X side of the boundary. The figure supports previous researchers' observations that houses in higher risk areas seem to sell at a premium compared to their lower risk counterparts. Compared to the overall average of characteristics on high and low risk sides of the flood boundary, A zone properties have higher levels of attributes that would generally be considered attractive: houses in those regions are, among other traits, closer to parks and have higher school quality as proxied by FCAT scores. Moreover, the lower surface elevation in the higher risk A areas is indicative of access to water-based amenities that would also be an important contributor to price. That these correlations are apparent for observable attributes implies that there are likely to be differences in unobserved attributes across flood zones that are also capitalized into housing values. While the overall average of these characteristics across flood zones are very different, the differences in attributes are much smaller if we focus on the immediate window near the flood boundary, especially within 1 kilometer.

Figure B.3: Losses as a Percentage Income, Remove CRS or PreFIRM Price Supports



Note. Figure plots the losses of removing either Pre-FIRM (top) or CRS (bottom) subsidies as a percentage of income.

Table B.1: Summary Statistics for Housing (Full Sample)

A. Structural and Neighborhood Characteristics

Variable	Mean	Median	St. Dev.	Min.	Max.
Price (in 2010 $'s)	219,841	171,985	168,702	9,625	1,399,301
# of Bathrooms	1.83	2.00	0.84	0.00	12.00
Year Built	1975	1978	16	1900	2010
Any Basement	0.0004	0	0.02	0	1
Enviro. Nuisances	0.60	0.00	1.43	0.00	15.00
School Quality	270	265	16	202	313

B. Flood-Related Characteristics

Variable	Mean	Median	St. Dev.	Min.	Max.
Dist. to River	218.5	227.9	50.5	33.6	294.8
Dist. to Park	14.4	12.4	11.5	0.0	90.9
Dist. to Coast	10.4	10.2	7.4	0.0	68.2
Surface Elevation	2.4	2.1	1.4	-1.3	20.8
Zone X (low risk)	0.405	0.000	0.491	0.00	1.00
Zone A (high risk)	0.593	1.000	0.491	0.00	1.00
Zone V (high risk)	0.002	0.000	0.041	0.00	1.00
Pre-FIRM	0.62	1.00	0.49	0.00	1.00
BFE Assigned	0.09	0.00	0.28	0.00	1.00
Relative BFE	-7.88	-8.00	1.69	-15.00	0.00

C. Homebuyer Characteristics

Variable	Mean	Median	St. Dev.	Min.	Max.
White	0.47	0.00	0.50	0.00	1.00
Asian	0.02	0.00	0.15	0.00	1.00
Black	0.12	0.00	0.32	0.00	1.00
Hispanic	0.39	0.00	0.49	0.00	1.00
Income (in 2010 $1,000's)	90.42	63.53	122.23	4.81	9745.98

Note. "BFE" refers to base flood elevation and "Relative BFE" is the surface elevation minus the BFE. "Enviro. Nuisances" refers to the number of sites listed on Florida's Institutional Controls Registry and "School Quality" evaluates achievement in the categories of reading, mathematics, science and writing, with a maximum score of 400 points (see Appendix 9 for details). All distances to spatial amenities are in kilometers. Surface elevation and BFE are measured in meters. The number of observations for all variables is 48,174, with the exception of the "Relative BFE," which only has 4,212 observations since not all areas are assigned a base flood elevation.

Table B.2: Summary Statistics for Housing (Boundary Discontinuity Sample)

A. Structural and Neighborhood Characteristics

Variable	Mean	Median	St. Dev.	Min.	Max.
Price (in 2010 $'s)	225,434	169,664	180,957	13,413	1,399,301
# of Bathrooms	1.85	2.00	0.79	0.00	8.00
Year Built	1973	1974	16	1900	2010
Any Basement	0.0006	0	0.02	0	1
Enviro. Nuisances	0.72	0.00	1.59	0.00	15.00
School Quality	267	265	16	202	313

B. Flood-Related Characteristics

Variable	Mean	Median	St. Dev.	Min.	Max.
Dist. to River	224.4	237.4	53.1	33.6	290.9
Dist. to Park	13.8	10.9	12.6	0.0	90.9
Dist. to Coast	8.0	7.0	6.4	0.0	68.2
Surface Elevation	2.3	2.0	1.3	-1.3	20.8
Zone X	0.51	1.00	0.50	0.00	1.00
Zone A	0.49	0.00	0.50	0.00	1.00
Zone V	0.00	0.00	0.05	0.00	1.00
Pre-FIRM	0.70	1.00	0.46	0.00	1.00
BFE Assigned	0.11	0.00	0.31	0.00	1.00
Relative BFE	-7.92	-8.00	1.89	-15.00	0.00

C. Homebuyer Characteristics

Variable	Mean	Median	St. Dev.	Min.	Max.
White	0.44	0.00	0.50	0.00	1.00
Asian	0.02	0.00	0.14	0.00	1.00
Black	0.10	0.00	0.31	0.00	1.00
Hispanic	0.43	0.00	0.50	0.00	1.00
Income (in 2010 $1,000's)	93.92	63.47	132.37	4.81	9745.98

Note. Table presents corresponding summary statistics for the boundary discontinuity design sample (i.e. sales for properties within 1 kilometer of flood boundary).

Table B.3: Flood Risk and Insurance Summary Statistics

A. Coverage and Annual Insurance Premium (in 2010 $'s)

Variable	Mean	Median	St. Dev.	Min.	Max.
Total Coverage (in $'s)	159,664	154,982	67,910	5,000	250,000
Full Premium (IP)	2,113	808	3,808	0	28,668
Discounted IP (pre-FIRM)	1,138	779	2,053	0	23,491
Discounted IP (pre-FIRM + CRS)	984	714	1,728	0	18,793

B. Insurance Premium Discounts

Variable	Mean	Median	St. Dev.	Min.	Max.
Total Subsidy (in $'s)	1,129	50	3,082	0	26,115
Total Subsidy (as %)	19.55	10.00	23.89	0.00	95.32
CRS Discount Rate (%)	12.02	10.00	6.18	0.00	25.00

Note. Table provides summary statistics for insurance coverage and premiums (panel A) and discounts through various NFIP price support channels (panel B). All dollars in January 2010 $USD.

Table B.4: HMDA Merged Sample Comparison with 2013 American Housing Survey

Race/Ethnicity	AHS	HMDA	Income	AHS	HMDA
White	46.9%	46.8%	Median	56,000	64,000
Black	13.9%	11.9%			
Hispanic	36.7%	38.9%			
Asian	2.5%	2.5%			

Note. This table presents race and income distributions in our sample, following the HMDA merge, and compares them with owner-occupied household characteristics data from the 2013 American Housing Survey for the metro area of Miami-Ft. Lauderdale-Hollywood, Florida.

Table B.5: ACS Migration Flows (2009-2013)

Nature of Move:	Count of People	Share of Movers
Within County	524,864	69.8%
Within CSA (study area)	577,212	76.8%
Outside CSA	75,511	10.0%
Outside FL	98,699	13.1%
Total Movers:	751,422	

Table B.6: First Stage Estimates (N=32,027)

	Black	Hispanic	Quintile 2	Quintile 3	Quintile 4	Quintile 5
Flood Zone	0.399	0.160	-0.028	-0.054	-0.109	-0.344
	0.050	0.035	0.041	0.041	0.044	0.048
BFE Assigned	-0.350	-0.358	0.145	0.195	0.475	0.885
	0.082	0.053	0.071	0.078	0.076	0.079
USGS Elevation	0.129	-0.041	-0.064	-0.090	-0.131	-0.224
	0.017	0.016	0.016	0.018	0.021	0.025
Coast<0.1km	-1.055	-0.293	0.017	0.800	1.363	2.273
	0.282	0.102	0.219	0.179	0.164	0.157
Income	0.004	0.010	0.023	0.044	0.064	0.087
	0.002	0.002	0.002	0.002	0.002	0.002
Black	3.647	1.306	-0.665	-1.368	-2.188	-3.711
	0.129	0.125	0.117	0.140	0.166	0.283
Hispanic	-1.234	3.212	0.159	0.180	-0.360	-1.168
	0.198	0.095	0.099	0.102	0.109	0.123
School Quality	3.245	-2.208	1.351	1.233	1.015	1.127
	0.525	0.255	0.201	0.214	0.191	0.194
Dist. to River	0.270	0.165	0.076	0.074	0.085	0.109
	0.017	0.007	0.008	0.008	0.008	0.008
ICR within 3km	-0.085	-0.092	-0.043	-0.086	-0.097	-0.116
	0.015	0.010	0.011	0.012	0.012	0.013
Dist. to Park	0.057	0.018	0.010	0.005	-0.001	0.000
	0.004	0.002	0.002	0.002	0.003	0.003

Note. This table presents first-stage estimates for heterogeneous sorting parameters (in utils) before conversion to a dollar value.

Table B.7: First Stage Estimates - Race by Income (Est/S.E.)

Race/Ethnicity:	White		Black		Hispanic	
Income Quartile:	Q1†	Q2	Q1	Q2	Q1	Q2
Flood Zone	-754.97	-68.25	235.83	295.47	148.02	10.26
	9.09	24.40	35.78	56.84	27.81	29.70
BFE Assigned	-8209.76	264.14	-162.37	-226.49	-300.96	104.82
	59.19	46.24	64.61	98.57	49.65	49.49
USGS Elevation	-157.06	-106.62	52.83	58.85	-66.54	-56.85
	3.83	9.48	14.00	23.37	12.49	13.87
Coast <0.1km	9099.74	837.90	-922.43	300.77	-302.43	582.79
	73.79	95.15	289.05	227.08	128.31	103.39
Income	173.25	31.31	3.83	20.94	3.99	31.91
	1.41	1.19	1.94	2.56	1.41	1.33
Black	-8338.80	-681.75	2463.93	1661.96	1258.67	-552.38
	54.49	109.08	91.64	140.55	88.37	122.07
Hispanic	-8836.85	-370.94	-772.15	-1030.27	1985.24	1520.86
	70.49	81.74	137.79	223.85	75.96	79.18
School Quality	7166.29	318.47	2079.13	944.98	-1326.87	-1273.68
	146.26	96.24	251.23	401.11	170.03	207.89
Dist. to River	-869.11	14.28	156.73	135.69	79.55	117.05
	14.27	3.76	8.18	13.16	5.17	6.26
ICR within 3km	16.10	-16.99	-26.39	-92.81	-27.21	-102.02
	2.56	6.28	9.40	18.93	7.36	9.77
Dist. to Park	-236.99	-3.08	35.27	22.76	9.38	9.84
	2.26	1.22	2.05	3.49	1.75	2.11

Note. This table presents heterogeneous MWTP estimates, where households are categorized into race-by-median-income groups. † represents the base group. Estimates for non-base group categories should be added to the base group estimate to recover the preference for that group.

Table B.8: Heterogeneous Parameter Estimates for Other Attributes

	Base Group	Black	Hispanic	Q2	Q3	Q4	Q5
BFE Assigned	-8308.83	-201.16	-205.96	83.64	112.10	273.39	509.06
	1523.87	47.29	30.33	40.99	44.92	43.51	45.54
USGS Elevation	-158.21	74.06	-23.81	-37.09	-51.93	-75.47	-128.86
	91.43	9.69	9.12	8.97	10.33	11.88	14.26
Coast$<$0.1km	8790.74	-606.67	-168.51	10.04	460.02	784.20	1307.56
	1894.82	161.95	58.39	125.70	102.73	94.20	90.33
Income	157.11	2.11	5.49	13.08	25.10	36.78	50.32
	33.96	1.19	0.88	1.24	1.39	1.39	1.35
Black	-8072.64	2097.83	751.15	-382.21	-786.75	-1258.61	-2134.63
	1252.22	74.31	71.99	67.27	80.42	95.20	162.97
Hispanic	-8979.09	-709.56	1847.48	91.69	103.78	-207.15	-671.82
	1642.84	113.61	54.70	57.05	58.52	62.68	70.48
School Quality	6509.70	1866.49	-1270.19	777.09	708.96	583.82	648.21
	3171.33	301.93	146.76	115.78	123.03	109.88	111.58
Dist. to River	-902.85	155.30	94.97	43.79	42.80	48.73	62.98
	304.95	9.61	3.96	4.65	4.78	4.46	4.57
Enviro. Nuisances	45.03	-49.02	-53.20	-24.57	-49.24	-55.90	-66.67
	44.31	8.74	5.93	6.45	6.78	6.99	7.32
Dist. to Park	-238.52	32.50	10.23	5.86	2.64	-0.59	0.07
	46.50	2.11	1.28	1.27	1.43	1.51	1.49

Note. Base group estimates (for white/Asian households in the first income quintile) are recovered from stage 2, and all standard errors are bootstrapped. The heterogeneous parameter estimates (by race/ethnicity and income group) are recovered from stage 1. All estimates have been converted to a (real 2010) dollar value using the estimated coefficient on rent and represent an annual (flow) MWTP per household. Estimates for non-base group categories should be added to the base group estimate to recover the preference for that group. Variables relating to distance are in kilometers, and elevation is in meters. The unconditional average income for the white/Asian group is $111 thousand.

Table B.9: Zone Transition Matrix

(Count / Percentage) Current FEMA Zone	Future FEMA Zone X	A	V	Total
X	15,395	877	20	16,292
	94.49	5.38	0.12	100
A	6,430	9,204	18	15,652
	41.08	58.8	0.12	100
V	7	4	72	83
	8.43	4.82	86.75	100
Total	10,085	110	21,832	32,027
	31.49	0.34	68.17	100

Note. Table tabulates a house's assigned flood zone according to current FIRMs and its zone according to future FIRMs (as of 2016).

Table B.10: Aggregate Impacts from Policy Reform

Policy	Aggregate Impacts ($ millions)
Remove CRS and Pre-FIRM	-$143.5
Remove Pre-FIRM Only	-$131.1
Remove Grandfathering	-$209.7
Value of Map Revisions	$243.5

Note. Table presents aggregate impacts of policy reform for the Miami-Dade, Port St. Lucie, Ft. Lauderdale CSA. Note that these figures do not represent changes in overall social welfare as not all costs and benefits are included.