THE EFFECTS OF FLOOD ZONE DESIGNATIONS ON LAND DEVELOPMENT

by

Samantha Joy Poteet

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Applied Economics

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2024

ACKNOWLEDGEMENTS

I would like to acknowledge my advisor Nick Hagerty for all of his thoughtful mentorship and guidance. I would like to acknowledge the rest of my committee members Diane Charleton, Justin Gallagher, and Melissa LaPalo for their helpful feedback. I would also like to thank my friends and family for their continued encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES – CONTINUED

ABSTRACT

In 2017, Hurricane Harvey hit Harris County, Texas and caused devastating flood damages. These flood damages were exacerbated by a rapid land development. This paper estimates the impact of the Federal Emergency Management Agency (FEMA) high risk flood zone designation, the Special Flood Hazard Area (SFHA), on land development. FEMA's flood maps convey information to homeowners regarding their property's flood risk and requires flood insurance for most properties in the SFHA. Using a spatial regression discontinuity design I find evidence of a 64% decrease in land development just inside of the SFHA boundary line. These results suggest FEMA can significantly impact the allocation of land development with the SFHA designation. Currently, FEMA underestimates flood risk, accurately assessing flood risk can help better prepare homeowners for future flooding events and allocate future land development in a more socially optimal way.

INTRODUCTION

Flooding in the US is the most frequent and costly natural disaster (Cigler, 2017). The costs of flood damages are expected to keep rising due to development in high-risk flood zones and climate change (Winsemius et al., 2016). Given the increasing frequency and financial toll of flooding within the US, it is important to learn how we can mitigate the impacts and better prepare for future flooding events. One potential flood adaptation response is to not over develop land in high-risk flood zones. Over the last two decades, much of the new land development has occurred in inland areas due to the rising housing costs along the coast and in urban centers (Levitt and Eng, 2021). However, rapid development is problematic in areas that have high flood risk since it increases the cost and severity of future flood events.

Due to increasing flood risk throughout the US, it is crucial to accurately assess flood risk to mitigate the adverse effects of flood events. The Federal Emergency Management Agency (FEMA) estimates flood risk at the property level by defining different flood map designations for different levels of risk. Properties in FEMA's highest risk flood zones are in the 100-year flood plain or defined by FEMA as the Special Flood Hazard Area (SFHA). While the flood maps do not perfectly reflect flood risk, the flood map designations are an important tool for household climate adaptation and inform FEMA's flood policies. Additionally, FEMA uses the flood map designation to determine insurance premiums. For properties in the SFHA designation, FEMA requires properties with a federally backed mortgage or those that have previously received disaster aid to have flood insurance, and in some cases requires additional construction regulations.

The additional requirements associated with the SFHA, as well as the information conveyed about flood risk, are intended to disincentivize new development in high-risk areas.

However, there are some constraints that undermine the SFHA designation's impact. First, FEMA systematically underestimates flood risk in inland areas (Wing et al., 2022). Second, FEMA flood maps are not regularly updated and do not account for future flood risk. Third, local government officials often have incentives to manipulate the flood maps by keeping the SFHA small (Pralle, 2019). Given the discrepancies between FEMA estimation of flood risk and the true flood risk of a property, FEMA flood maps might be incentivizing land development in the areas that are actually vulnerable to future floods.

Inaccurate flood maps can have devastating consequences for communities that experience extreme flood events. For example, in the summer of 2017, Hurricane Harvey caused devastating floods in Harris County, Texas. Analysis of flood damages found that nearly three quarters of damaged residential buildings were outside the SFHA (Hunn et al., 2018). FEMA's underestimation of flood risk in Houston left thousands of households uninsured and unprepared for Hurricane Harvey. In the decade prior to Hurricane Harvey, Houston had the fifth largest population growth and land development in the US (Bounoua et al., 2018). The probability of an extreme flood event like Hurricane Harvey increased 21 times because of the surge in Houston's land development (Zhang et al., 2018).

In this paper I estimate the impact of FEMA's SFHA designation on land development in Harris County, Texas. I use property-level data for SFHA boundaries from the National Flood Hazard Layer (NFHL) and 30 by 30 meter land development data from the United States Geological Survey (USGS). Given the granularity of this data, I am able to use a spatial regression discontinuity design. This design compares properties next to each other on either side of the SFHA boundary and relies on the assumption that the SFHA boundaries do not reflect any true discontinuities in flood risk. The errors in FEMA's flood zone designation help validate this assumption.

This paper broadly contributes to the literature on the effects of FEMA's flood maps on the housing market. More specifically, this paper contributes to the literature by estimating the effect of the SFHA on land development in an inland setting, which expands on previous work. This paper shows evidence of a 64% decrease in land development for the area just inside the SFHA boundary. These results emphasize the need for FEMA to consider how flood map revisions will impact land development and the subsequent effects on homeowners and future flood events.

LITERATURE REVIEW

The current literature on the effects of FEMA's flood maps on the US housing market is mixed. There are several papers finding small decreases in property values from being in the higher risk flood zone (Harrison et al., 2001; Bin et al., 2008; Daniel et al., 2009; Bin and Landry, 2013; Beltrán et al., 2018; Ortega and Taşpınar, 2018; Eichholtz et al., 2021; Hino and Burke, 2021). Bin and Kruse (2006) does not find statistically significant negative effects of flood risk on property values in coastal areas. Other studies find an increase in property values from being in the flood zone in areas in Texas and Florida (Atreya and Czajkowski, 2019; Ostriker and Russo, 2022). Gourevitch et al. (2023) takes a different approach in accessing the impact of FEMA's flood maps on the property values by comparing the SFHA to more scientifically accurate flood maps and finds that underestimating flood risk has led to $200 billion in property over evaluation within the US housing market. These inconsistencies in the literature show a gap in our understanding of how being in the SFHA affects the desirability of properties.

Other studies estimate the extent to which homeowners are responsive to new flood risk information. (Gallagher, 2014) finds that following a local flood event there is a spike in flood insurance up-take. More recent work finds that white communities respond to new flood risk information by increasing insurance up-take and predominately minority communities decrease insurance up-take (Weill, 2023). Fairweather et al. uses more scientifically accurate flood risk information instead of FEMA's flood maps in a randomized control trial to show that home buyers are responsive to flood risk information and make more informed purchasing decisions.

There is one other paper that estimates the impact of FEMA's flood maps on land development. Ostriker and Russo (2022) add to the literature on FEMA's flood map's impact on the housing market in Miami-Dale County, Florida by using a spatial regression discontinuity design to estimate the flood map's impact on property values and land development and uses these estimates in a structural model. They find a 9% decrease in land development within the SFHA using a one-mile bandwidth on either side of the SFHA boundary. I use a spatial regression discontinuity approach to estimate the impact of SFHA boundaries on land development in Harris County, Texas that uses an optimal bandwidth selection outlined in Calonico et al. (2020). By using this method of bandwidth selection and estimating impacts in Harris County, Texas, I find results with larger magnitude than in previous work.

The differences in these results might be due to two reasons. First, Harris County has SFHA designations in areas along its coastline and inland areas where as Miami-Dale County is largely a coastal area. Since the literature on the impacts of FEMA's flood maps on property values varies by location, particularly in areas where land development is correlated with flood risk, then it is reasonable that the impacts on land development would also vary by location. Second, culturally Harris County, Texas and Miami-Dale County, Florida are very different. There might be an unobserved characteristic of the culture in southern Florida that makes residence less responsive to flood risk information. Alternatively, there might be an unobserved characteristic in Harris County that makes the residence more responsive to flood risk.

BACKGROUND

In the US, few private insurance companies offer flood insurance since a single flood event can be difficult to predict and can affect thousands of properties simultaneously. To fill this gap in coverage the US government began subsidizing flood insurance and created the National Flood Insurance Program (NFIP) in 1968. In order to determine insurance premiums, FEMA is responsible for creating Flood Insurance Rate Maps (FIRMs) or flood maps that are used by the NFIP. These maps evaluate a property's probability of being flooded within a given year. The 100-year flood plain, also known as the SFHA or high-risk zone, has an annual probability of being flooded by 1% or more. The 500-year flood plain, also known as moderate risk flood zones, have a 0.2% to 1% probability of being flooded in a given year. A property's risk is determined by proximity to water, slope, and elevation.

Properties with a federally backed mortgage in a SFHA are required to purchase flood insurance. This requirement has historically been poorly enforced and flood insurance uptake remains low due to incomplete information on risk (Chivers and Flores, 2002). The NFIP heavily depends on subsidized and grandfathered insurance premiums within the SFHA, that are lower than the actuarially accurate risk premiums (Tier). Actuarially fair insurance premiums would make homeowners internalize the expected costs of living in a high risk flood zone. Overall, the insurance mandate should discourage land development in the SFHA. However, since national flood insurance is poorly enforced and premiums are lower than the actuarially fair premium, the insurance mandate falls short of optimally discouraging land development in the SFHA.

Currently FEMA flood maps are incomplete, outdated, and underestimate flood risk to the extent that FEMA estimates that 13 million Americans live in the SFHA and more scientifically accurate flood maps estimate that it is closer to 41 million Americans (Wing et al., 2022). FEMA is supposed to update the flood maps for any given area every 5 years.

However, nearly two thirds of flood maps have not been updated in the last five years (Scata, 2017). FEMA has created updated flood maps for some coastal flood hazard areas in the US. However, due to the catchment size FEMA uses in estimating flood risk, flood mapping of risk from rivers and rainfall is incomplete (Wing et al., 2022). The underestimation of flood risk becomes clear when considering one-third of all NFIP claims are for properties located outside of the SFHA boundaries. Underestimating flood risk to the extent that properties with true high flood risk are incorrectly classified as being outside of the SFHA could lead to a decrease in development in the SFHA. However, underestimating flood risk for properties within the SFHA may lead to further under pricing of insurance premiums, inflating property values, and increase demand for land development within the SFHA.

There is a bundle of treatments that the SFHA designation represents. The SFHA designation requires properties to purchase flood insurance, gives information about a properties flood risk, gives eligibility for streamline BCA process for FEMA's Hazard Mitigation Grant program, and in some cases requires additional construction regulation. Since there are multiple mechanisms through which the SFHA designation impact land development, results are interpreted as how all of these treatments combined affect land development. Each of these mechanisms are intended to disincentivize land development in the SFHA.

DATA

Overview

The data for this paper includes publicly available data from the National Flood Hazard Layer (NFHL) for flood risk assessment, the United States Geological Survey (USGS) National Dataset of Rasterized Building Footprints for a measure of land development, USGS National Map of the 3D Elevation Program (3DEP) for a measure of elevation, and the National Land Cover Database (NLCD) that identifies land type. Table 4.1 displays summary statistics for the treatment variable, running variable, and all outcome variables.

Main Data

For the treatment variable, I use the SFHA flood zone area data from the NFHL. This data contains spatial data delineating the different flood zones within Harris County and identifies the flood zones that are classified as being in the SFHA.

To create an object that identifies the mulipolygon of the SFHA and the multipolygon of the non-SFHA, I aggregate these flood zones up to the SFHA level. First, I take the union of the flood zone polygons grouping by SFHA designation. Then, to create a single geometry for each SFHA designation I group by SFHA designation and combine the unionized polygons.

For the outcome variable, I use the National Dataset of Rasterized Building Footprints from the United States Geological Survey. A subset of this dataset for the state of Texas contains a raster layer of 30 by 30 meter cells covering the state. The value of each cell in this raster layer is the number of building centroids that intersect the cell.

To narrow down the Texas dataset to just the observations within Harris County I crop the raster layer. First, I use a separate shape file of Harris County from the United States Census Bureau to create a spatial extent object that delineates the outer boundaries of the county's geometry. Then, I crop the raster layer down to cells within the extent boundaries. This step restricts the sample to only the area corresponding to Harris County. After cropping the raster layer, I use the original Harris County shape file to mask the remaining raster layer to remove any cells that are not within Harris County. Finally, I transform the masked raster layer into an sf object with point type geometry, resulting in a centroid for each raster cell. This transformation is necessary for further spatial analysis and calculating the running variable.

For the running variable, I calculate the distance of each raster cell centroid from the SFHA boundary line. Then, I bind together the results from the distance calculation with the cleaned buildings raster cell centroid data.

To create a dataset that distinguishes whether the raster cell lies within the SFHA, I spatially join the SFHA polygon data frame with the data frame containing distance at the cell level. Figure 4.1 represents a map of the SFHA boundaries and the distance of each building cell from the boundary line.

To address building observations that intersect with the SFHA boundary line and cannot be completely assigned to the to either side of the SFHA, I exclude observations within 15 meters of the boundary line. Since each observation in the building data frame represents a centroid of a 30 by 30 meter cell and the radius of each cell is 15 meters.[1] Finally, I set distance to a negative value for observations outside the SFHA boundary. Figure 4.2 shows the distribution of distance in this dataset.

[1]The process of identifying intersecting building observations with the SFHA boundary line may not be entirely precise. It is possible that some centroid observations are within 15 meters of the SFHA boundary but the polygon the centroid represents does not actually intersect the SFHA boundary. In future work, with fewer computational constraints, I would refine the exclusion process to only remove observations that actually intersect the SFHA boundary line. First, I would introduce a grid cell identifier and join the polygons of the buildings data raster grid with the polygons of the SFHA boundaries. Then, I would drop observations that are duplicated and therefore only excluding observations that intersect SFHA boundary line.

Additional Data

In additional regression discontinuity plot visualizations, I use restricted samples of the data that only contains census tracts and census blocks with observations on either side of the SFHA. To create census tract and census block identifier variables, I intersect the cleaned data frame with a census tract and block shape file of Harris County. To create the restricted sample data frames that only includes tracts or blocks with observations on either side of the SFHA boundary line, I group by tracts or blocks and create a variable that identifies tracts or blocks with observations on either side of the boundary. Then, I drop tracts or blocks that do not have observations on either side of the boundary.

Covariate Data

I perform covariate continuity checks to verify the assumption that the SFHA boundaries are arbitrary, so development should not vary across the SFHA boundary for any other reason. I use the USGS National Map 3DEP dataset for the elevation covariate. I use the USGS National Building Rasterized Dataset that was also used for the outcome variable to create the latitude and longitude covariates. For the other covariates and to drop observations in water, I use the NLCD to identify cells in open water, agricultural land, and wetlands.

The USGS National Map 3DEP has 1 arc-second, approximately 30 meters, spatial resolution and provides a vertical measure of elevation in meters. The raw elevation data for Harris County is composed of three raster layers subsets. I individually process each of the elevation data subsets by cropping the raster files to the extent of Harris County, Texas. After cropping the raster layer, I use the original Harris County shape file to mask the remaining raster layer to remove any cells that are not within Harris County. Next, I transform the masked raster layer into an sf object with polygon type geometry. I spatially join the elevation data frame with the building data frame at the 30 by 30 meter cell level. To obtain latitude and longitude variables, I extract the coordinates from this data frame's geometry column.

The NLCD is raster data covers the contiguous US for 2021 with a 30 by 30 meter spatial resolution. Following the same steps for processing the building count and elevation data, I crop and mask the NLCD to the extent of Harris County. To merge this data onto the main data frame, I convert the NLCD raster data into a vector data frame. Then, I extract the mean values from the raster data frame at the location of the vector data frame. After creating a vector of values corresponding to the locations in poly, I merge the NLCD vector data frame onto the main data frame. Finally, I filter out observations that are in open water since regardless of the SFHA designation this land would not be developed.

Table 4.1: Summary Statistics

	min	median	max	mean	sd
Buildings	0.00	0.00	8.00	0.23	0.56
Distance in km	-3.43	-0.35	2.73	-0.43	0.64
SFHA	0.00	0.00	1.00	0.21	0.41
Elevation in m	-7.31	21.61	98.92	26.35	17.69
Latitude	29.51	29.87	30.17	29.86	0.15
Longitude	-95.96	-95.40	-94.91	-95.40	0.25
Wetlands	0.00	0.00	1.00	0.07	0.26
Agricultural land	0.00	0.00	1.00	0.15	0.36

Note: Each column represents the minimum, median, maximum, mean, and standard deviation value for each different variable. Buildings is the main outcome variable representing the number of buildings for each 30 by 30 meter cell. Distance in km represents the running variable in distance from the SFHA boundary line for each cell. SFHA is a binary variable representing the treatment. Elevation, latitude, longitude, wetlands, and agricultural land all are outcome variables used for continuity checks.

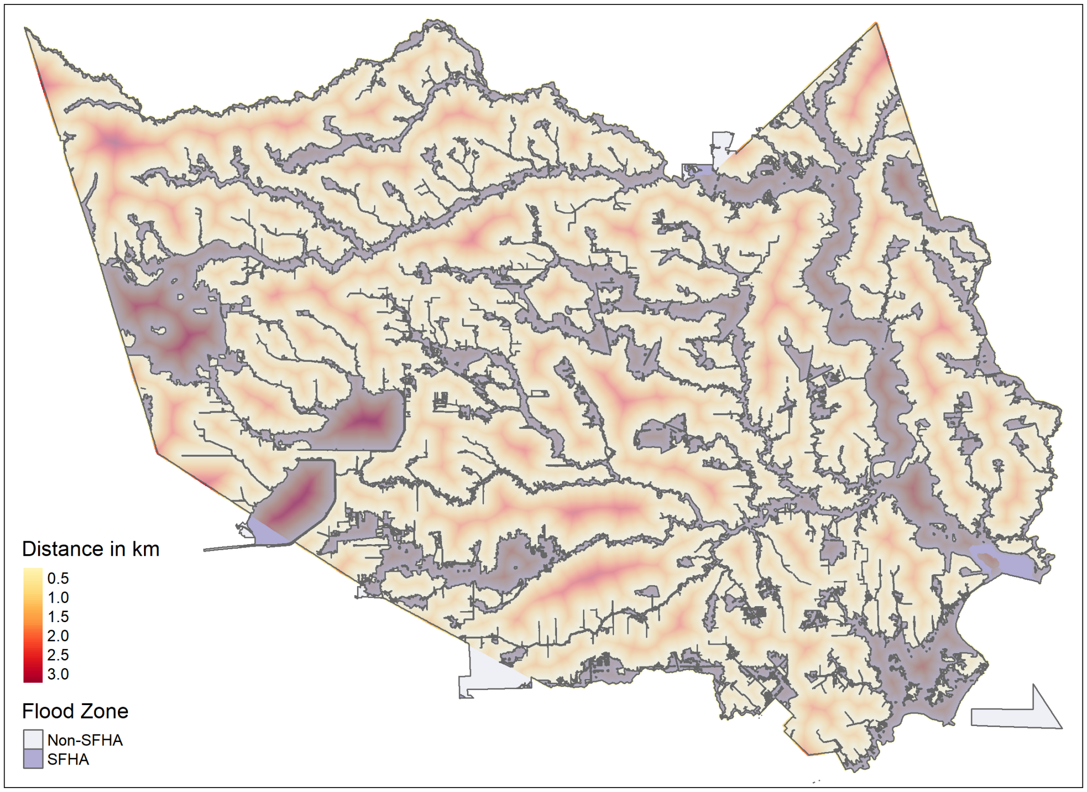

Figure 4.1: Map of Harris County, TX SFHA boundaries and building distances.

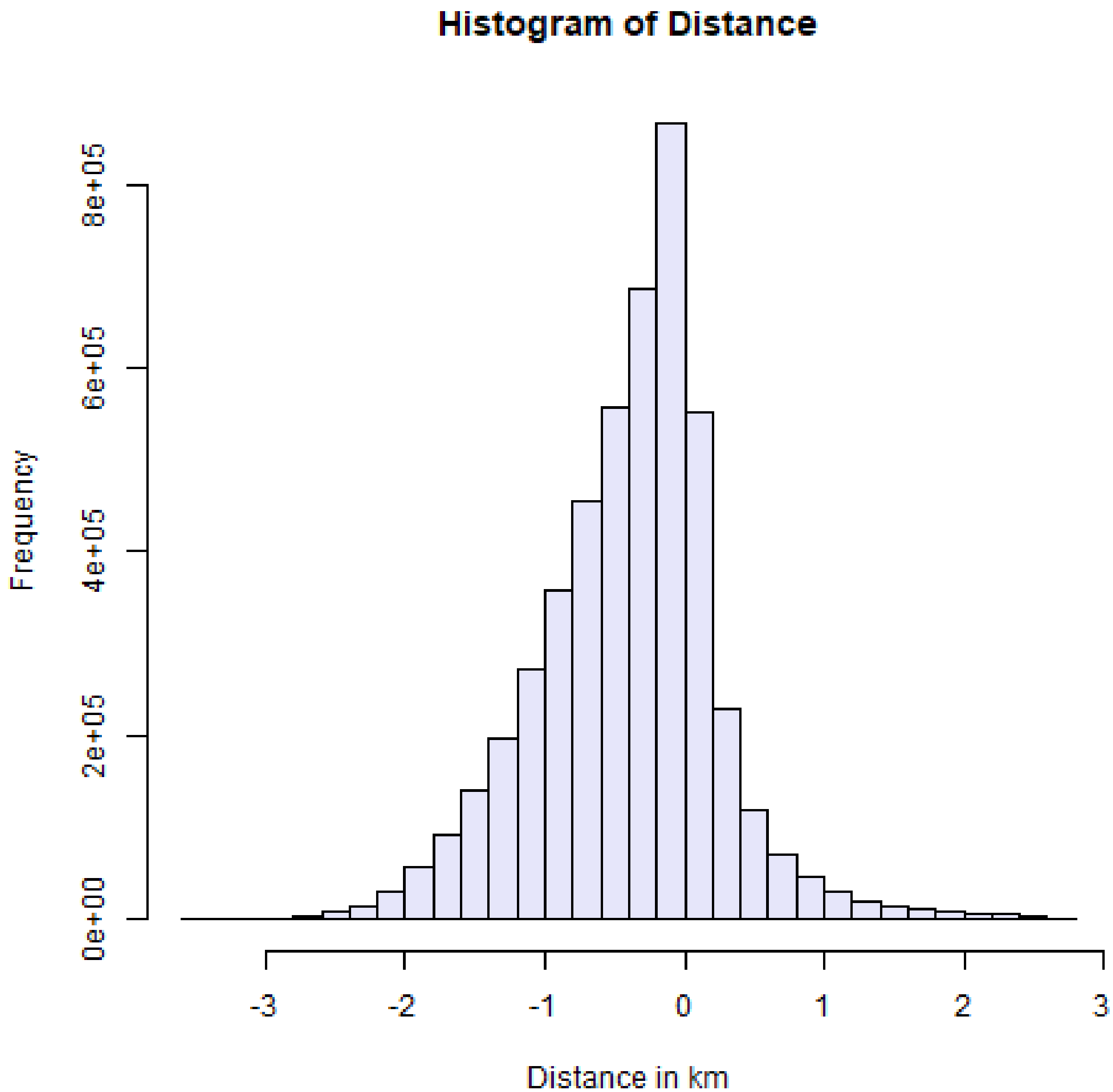

Figure 4.2: Histogram of distance from the SFHA boundary line. Positive distance represents buildings inside the SFHA.

EMPIRICAL STRATEGY

Main Specification

In an ideal experimental setting to assess the impact of FEMA's SFHA designation on land development. The SFHA designation would be randomized. This means that properties in both the SFHA and non-SFHA would on average have the same flood risk levels and share the same characteristics that affect land development, including economic activity, population, and property amenities. These two groups would be identical without the SFHA designations and therefore would have the same potential for land development.

In reality, the SFHA designation is not random and is correlated with flood risk. Properties within the SFHA may have different population densities, levels of economic activity, and property characteristics that influence land development. To account for these differences, I use a spatial regression discontinuity design. This strategy compares areas that are just on either side of the SFHA boundaries. These areas are likely more similar to each other compared to all of the areas on either side of the boundary since the SFHA boundary line does not represent any true discontinuities in flood risk. This design address omitted variables bias by assuming that all unobserved characteristics vary smoothly across the boundary line.

In this design, the treatment variable is whether or not a cell is within the SFHA, and the running variable is the distance from the SFHA boundary line, where positive distances indicate that a cell is in the SFHA. The cutoff in this design is when the distance of a cell to the SFHA boundary line is zero. The variation in this design comes from difference in whether or not a property is within the SFHA boundaries.

To implement the spatial regression discontinuity design I run regressions of the following form.

$$Buildings_i = \alpha + \tau SFHA_i + \beta Distance_i + \gamma SFHA_i * Distance_i + \epsilon_i \tag{5.1}$$

Where i represents individual 30 by 30 meter cells, SFHA is a binary variable for whether or not a property is in the SFHA, and Distance is the number of kilometers a property is away from the SFHA boundary line. In this equation represents the estimate of the discontinuity in land development at the SFHA boundary line. I estimate this equation using a local linear regression, triangular kernel weights, and I select bandwidths using rdbwselect to construct data-driven bandwidth choices as outlined in Calonico et al. (2020).

Using triangular kernel weights in this design put greater weight on the observations closer to the SFHA boundary line. I re-construct the weights by subtracting the absolute value of the distance divided by the optimal bandwidth from one. This process ensures that the weights are symmetric around the cutoff and will linearly decrease as they approach the bandwidth. When the absolute value of the distance of an observation exceeds the optimal bandwidth, the weight is set to zero.

This design relies on two assumptions: Firstly, that developers do not have control over what land is within the SFHA boundaries, and secondly, characteristics of the land around the SFHA boundary are continuous, except for variations in land development and characteristics directly determined by SFHA designations. These identification assumptions are likely held in general but situations where property owners petition to be removed from the SFHA violates the first assumption. Following previous work to address this issue, future work could use the initial FEMA flood maps that have not been altered by map revisions Ostriker and Russo (2022).

Additional Specification

To improve the accuracy of the estimates and standard errors, I run specifications with additional adjustments. To allow for geographic correlation in the error term, I cluster standard errors at census tract level. This adjustment provides a more accurate estimate of the uncertainty around the regression coefficients. This is particularly impactful in situations where land development is influenced by local zoning regulations, proximity to amenities, shared socio-economic status within tracts, and when the area within tracts are subject to the same environmental factors. To reduce omitted variable bias in some specifications, I use geographic fixed effects at the census tract and block level. These adjustments control variables that influence land development and may be correlated within small geographic units but vary across units. Small geographic fixed effects also address concerns of the sample becoming skewed when aggregating the data across all boundary segments, ending up with more building data representing the area outside the SFHA. Census tract and block level fixed effects only compare the number of buildings across the SFHA boundary line in the same census tract or block.

The main specification uses rdrobust and optimal bandwidth selection but is too computationally intensive to include fixed effects. The specifications run using fixed effects are run using bandwidths extracted from the optimal bandwidth of the specification with only clustering. The ideal specification would use both small geographic fixed effects and the optimal bandwidth for the fixed effects model.

Robustness Checks

To verify the assumption that, in absence of the SFHA designation, land development would be continuous across the boundary and that the discontinuity is driven by the SFHA designation, I check for continuity of various covariates. Using the main specification with clustered standard errors at the census tract level, I estimate the effects of the SFHA on elevation, latitude, longitude, agricultural land, and wetlands. To test if my estimated coefficients are robust, I rerun the main specification using rdrobust with 50% , 200%, and 300% of the optimal bandwidth and I check for continuity at varying cutoffs near the SFHA boundary. In future work, there is room for additional specifications. One possible adjustment is to estimate the main specification on a restricted sample that does not include SFHA boundaries that border levees, parks, and bayous. This sample restriction would exclude areas that were not going to be developed regardless of the SFHA high risk designation and likely reducing the magnitude of estimates found in this paper.

RESULTS

Main Results

Table 6.1 reports the estimates for the main regression discontinuity specifications, with buildings per cell as the outcome variable and the cutoff set at the SFHA boundary line. These results estimate the local average treatment effect that ranges from a 0.129 to a 0.156 decrease in buildings per cell just inside the SFHA boundary. The magnitude of the point estimates is consistently more than half of the average number of buildings in the full data set. The first column displays the point estimate from equation (1) using the rdrobust function developed by Calonico et al. (2020). This specification results in a 64% decrease in land development within 0.187 km, or 613 feet, of the SFHA boundary line. These results are statistically significant at the 1% level, ruling out any possibility of a null or positive effect on land development. The RD plot for building density relative to distance from the SFHA boundary line is illustrated in Figure 6.1 and Figure 6.2, where negative values of distance represent the distance from the SFHA boundary line for cells in the non-SFHA and positive values of distance represent the distance from the boundary line for observations in the SFHA.

To account for potential spatial correlation of land development within census tracts, Column 2 estimates the main specification with clustered standard errors at the census tract levels. The standard errors in Column 2 slightly increase suggesting that there is some spatial correlation and spatial heterogeneity in the treatment effect. This correlation in land development may be driven by zoning regulations, property values, and proximity to amenities or urban areas in Harris County.

Other factors that are correlated with both land development and the SFHA boundaries, like topographical characteristics and socioeconomic factors, can vary between small geographic areas. To control for these localized effects, Columns 4 and 5 introduce tract and block fixed effects. With this adjustment, the results show a decrease in land development just inside the SFHA boundary with a slightly smaller effect compared to the main specification.

Figure 6.3 plots the regression discontinuity with the sample restricted to tracts that have observations on either side of the boundary. Similarly, Figure 6.4 uses a sample restricted to only the census blocks that have observations on either side of the boundary. This sample restriction controls for geographic specific factors and shows the causal effect at the discontinuity point. Each of these figures show a smaller discontinuity than the main specification, which is consistent with the change in estimates seen in Table 2 when adding fixed effects.

These figures address concerns about spillovers by providing evidence that there is no sorting around the SFHA boundary. Prior to the threshold the shape of the conditional expectation function is decreasing on each side of the SFHA boundary. This suggests that the decrease in development is driven by a total reduction in development rather than development sorting out the SFHA. The restricted samples only include census tracts and census blocks that intersect the SFHA boundary. These sample restrictions are useful when using the shape of the global conditional expectation function to understand the extent of spillovers since properties in these areas will have a higher propensity to sort into the non-SFHA.

Robustness Checks

Continuity Checks

Table 6.2 shows the point estimates of the covariate regression discontinuities. Each column shows one of the point estimates for elevation, latitude, longitude, agricultural land, and wetlands respectively. Each of these coefficients are estimated from the main specification with clustered standard errors at the census tract level. The results for all of these estimates are not statistically different from zero, showing that each covariate is continuous across the SFHA boundary line. This supports the assumption that nothing else is impacting land development discontinuously at the SFHA boundary line.

Figure 6.5 represents the RD plot for elevation across the SFHA boundary line where elevation is in meters above sea level. The global shape of the conditional expectation function shows that that elevation is continuously trending downwards approaching the SFHA boundary from outside the SFHA. This shows that the SFHA has areas with lower elevation than the non-SFHA areas. Figures 6.6 and 6.7 represent the RD plots for latitude and longitude. Each of these plots are continuous across the SFHA boundary and show that there is no discontinuity spatially in observations across the SFHA boundary. This shows that the discontinuity in land development is not due to abrupt changes in spatial location. Verifying that elevation, latitude, and longitude are continuous across the SFHA is important from a policy perspective. By showing that geographic features outside the control of the government are not driving the discontinuity in land development it informs how much FEMA's SFHA designation can influence land development in high flood risk areas.

Additional plots of continuity checks reinforce that external factors are not driving the discontinuity in land development. The RD plot in Figure 6.8 shows the distribution for agricultural land across the SFHA boundary line. The conditional expectation function shows that there is more agricultural land within the SFHA, and the trend is continuous across the SFHA boundary. While agricultural land typically has lower building density than urban areas, its continuity across the SFHA boundary indicates that it does not play a role in the observed discontinuity in land development. Similarly, shown in Figure 6.9, the distribution of wetlands is continuous across the SFHA. There are physical and regulatory constraints to development on wetlands resulting in lower building density. However, since wetlands do not vary discontinuously across the SFHA boundary line it is not influencing the decrease in land development.

Various Bandwidths

Table 6.3 estimates the main specification with varying bandwidths, showing the optimal bandwidth in column 1 and 50 percent, 200 percent, and 300 percent of the optimal bandwidth in columns 2, 3, and 4, respectively. Each of these variations' estimates show a statistically significant discontinuity, indicating the results are robust to changes in bandwidth choice.

Various Cutoffs

Table 6.4 shows the estimates of the base specification at various cut offs points. In these models, land development is continuous across the placebo cutoffs 0.1, -0.1, 0.2, and -0.2. This supports the assumption that in absence of the treatment, the trend of land development would be continuous.

Table 6.1: Main Results

	(1)	(2)	(3)	(4)	(5)
SFHA	-0.151	-0.156	-0.156	-0.134	-0.129
	(0.002)	(0.012)	(0.009)	(0.009)	(0.009)
(Intercept)			0.237		
			(0.012)		
distance_km			-0.213	-0.283	-0.255
			(0.063)	(0.049)	(0.047)
distance_km × SFHA			0.600	0.762	0.775
			(0.091)	(0.077)	(0.078)
FE: Tracts				X	
FE: Blocks					X
Num.Obs.	1316992	1140290	1140290	1140290	1140290
Bandwidth	0.187	0.157	0.157	0.157	0.157
R2			0.021	0.156	0.194
Clustered Std.Errors		by Tract	by Tract	by Tract	by Tract

Note: Column 1 contains the point estimate from equation (1) using the rdrobust function created by Calonico, Cattaneo, and Titiunik 2015. Column 2 contains the point estimate from rdrobust with clustered standard errors by census tracts. Column 3 contains the estimate from manually constructing the triangular kernel weights and estimating equation 1 to replicate the estimate found in column 2. Columns 4 and 5 estimate equation 1 with tract and block fixed effects.

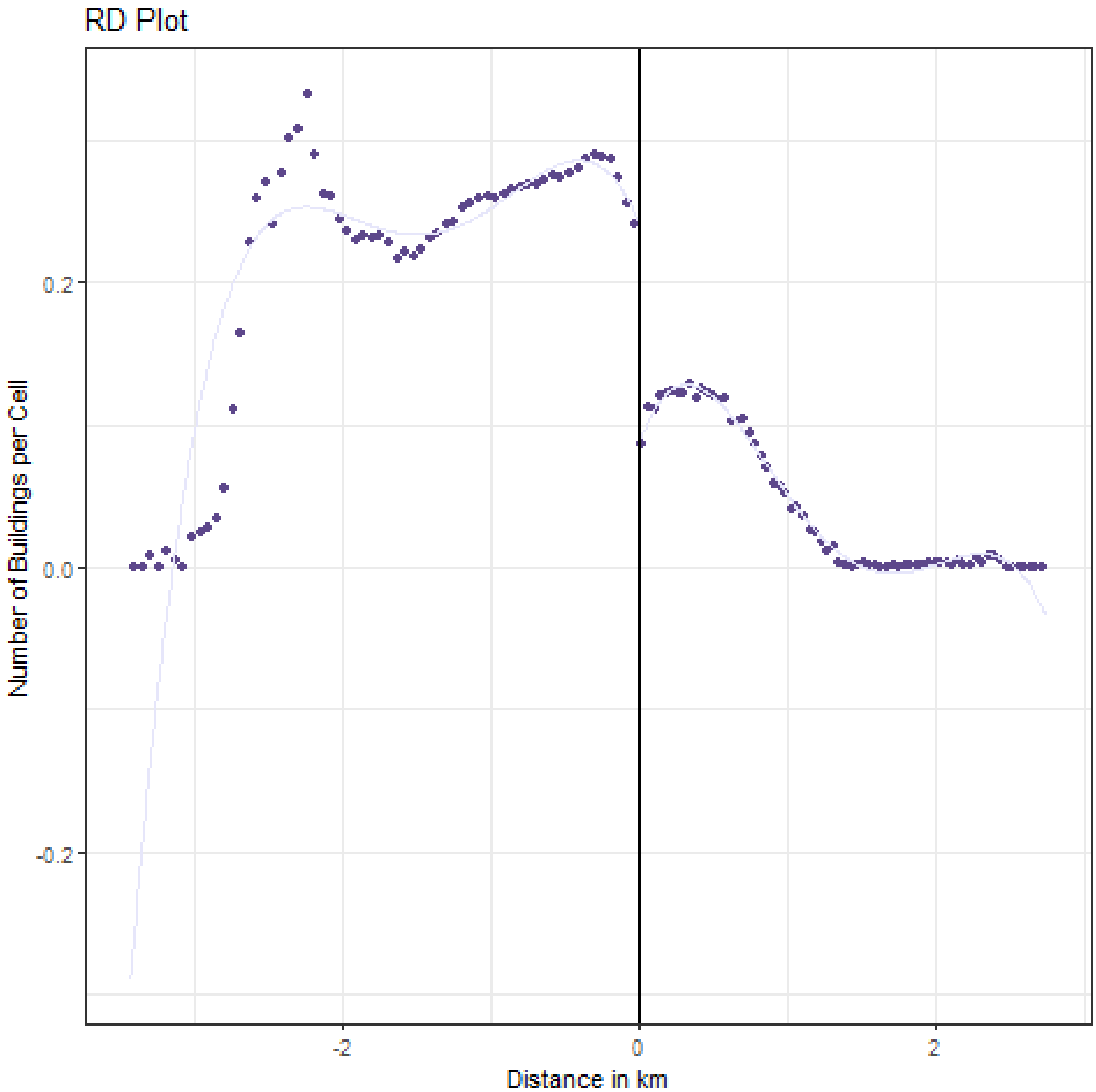

Figure 6.1: RD plot of building density relative to distance from SFHA boundary. Positive distances represents buildings in the SFHA. Each point represents a binned mean of observations. The number of bins is selected using IMSE-optimal evenly-spaced method using polynomial regression.

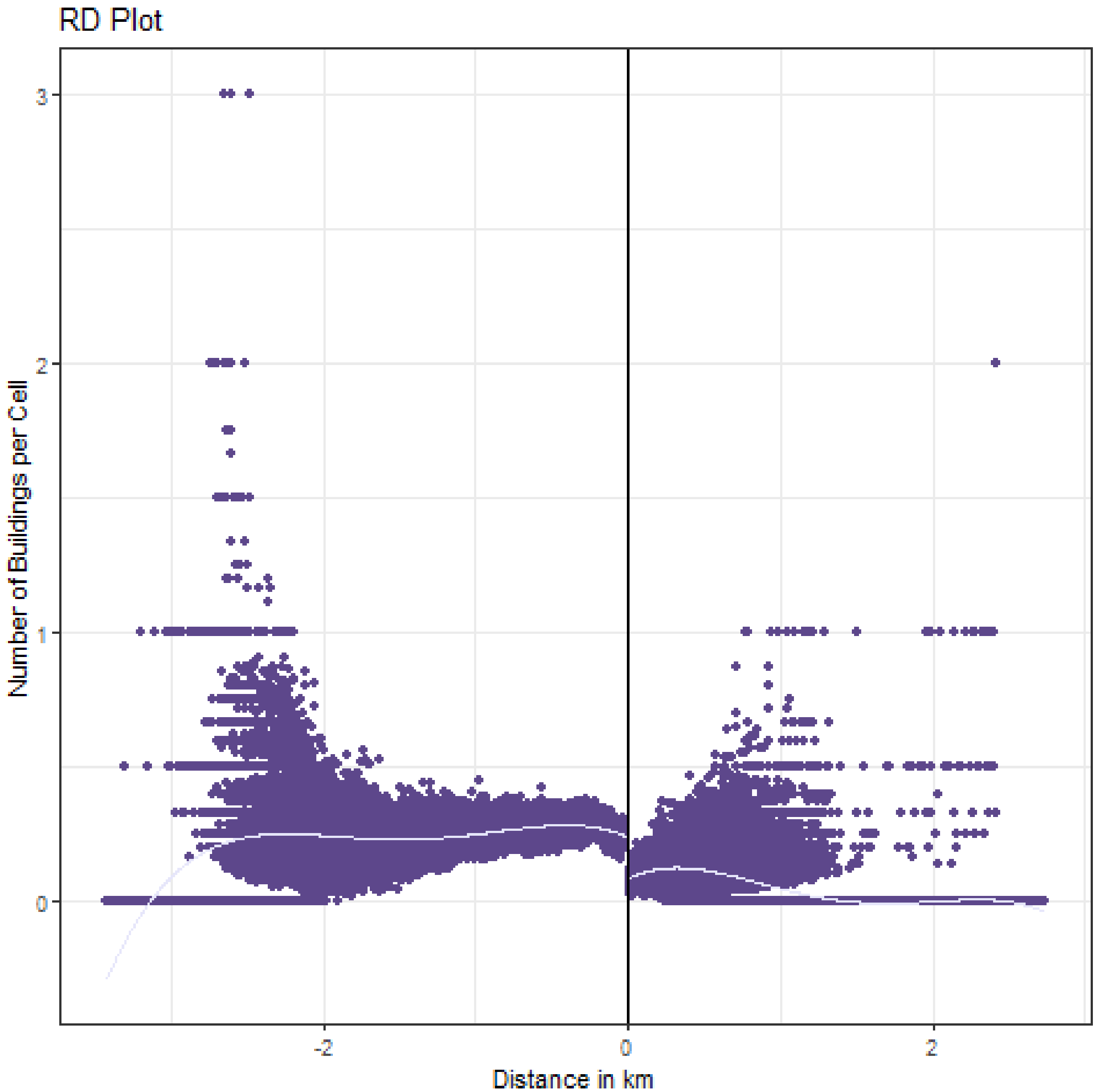

Figure 6.2: RD plot of building density relative to distance from SFHA boundary. Positive distances represents buildings in the SFHA. Each point represents a binned mean of observations. The number of bins is selected by using the mimicking variance evenly-spaced method using spacings estimators.

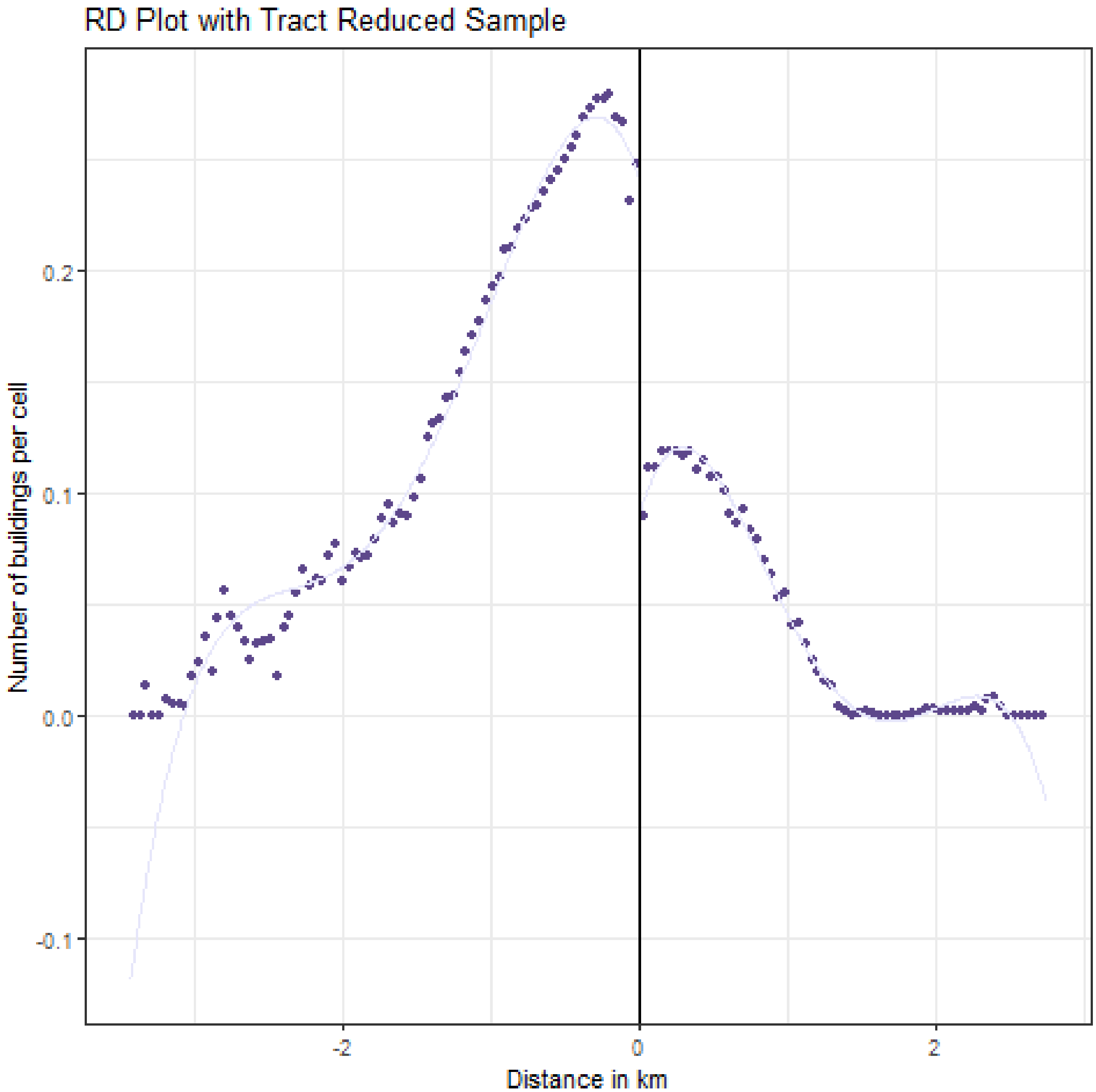

Figure 6.3: RD plot of building density relative to distance from SFHA boundary with sample restricted to tracts that have observations on both sides of the SFHA boundary. Positive distances represents buildings in the SFHA. Each point represents a binned mean of observations. The number of bins is selected using IMSE-optimal evenly-spaced method using polynomial regression

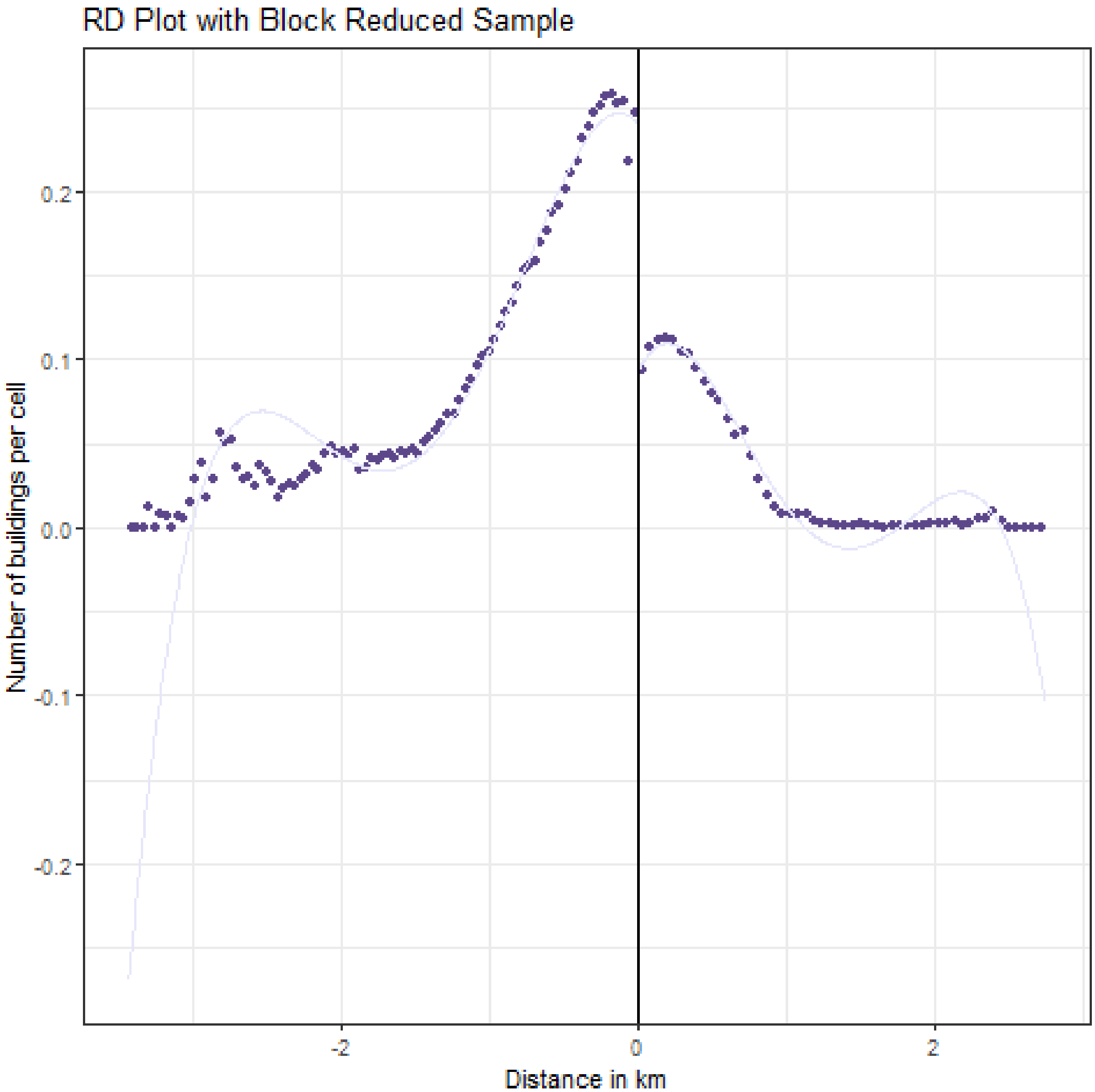

Figure 6.4: RD plot of building density relative to distance from SFHA boundary with sample restricted to blocks that have observations on both sides of the SFHA boundary. Positive distances represents buildings in the SFHA. Each point represents a binned mean of observations. The number of bins is selected using IMSE-optimal evenly-spaced method using polynomial regression

Table 6.2: Continuity Checks

	(1)	(2)	(3)	(4)	(5)
	Elevation	Latitude	Longitude	Ag Land	Wetlands
SFHA	-1.456	0.005	0.007	0.015	0.018
	(2.067)	(0.014)	(0.029)	(0.028)	(0.009)
	[-5.203, 2.526]	[-0.021, 0.033]	[-0.050, 0.061]	[-0.042, 0.065]	[-0.005, 0.030]
Num.Obs.	1333429	1211080	1672872	1304800	429539
Bandwidth	0.190	0.169	0.257	0.185	0.056
Clustered.SE	by Tract	by Tract	by Tract	by Tract	by Tract

Note: Columns 1 through 5 represent the main specification run with rdrobust on elevation, latitude, longitude, agricultural land, and wetlands as the outcome variables. The conventional standard errors are in parentheses and the robust confidence intervals are in the square brackets below each coefficient.

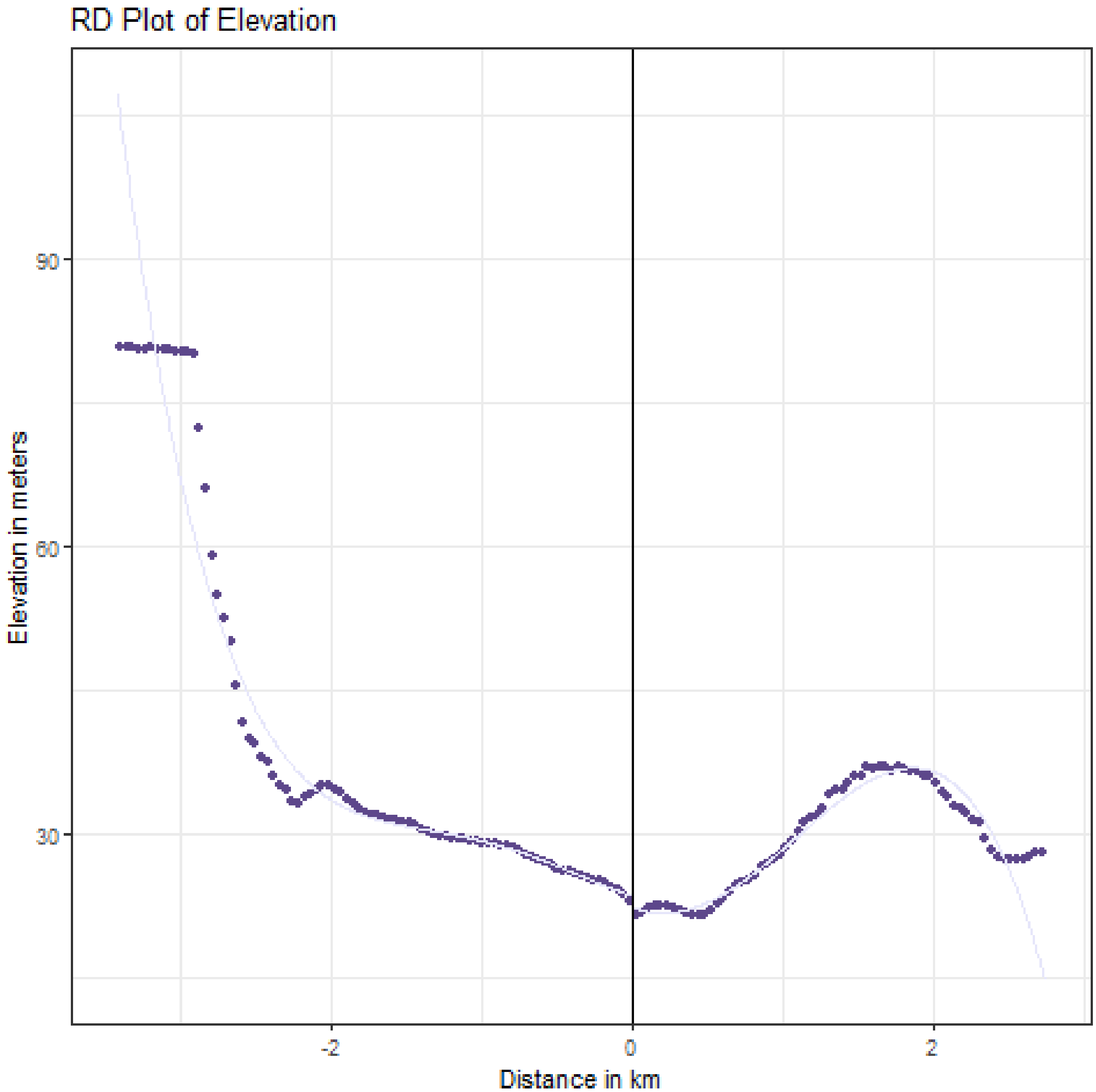

Figure 6.5: RD plot of elevation relative to distance from SFHA boundary. Positive distances represents buildings in the SFHA. Each point represents a binned mean of observations. The number of bins is selected using IMSE-optimal evenly-spaced method using polynomial regression.

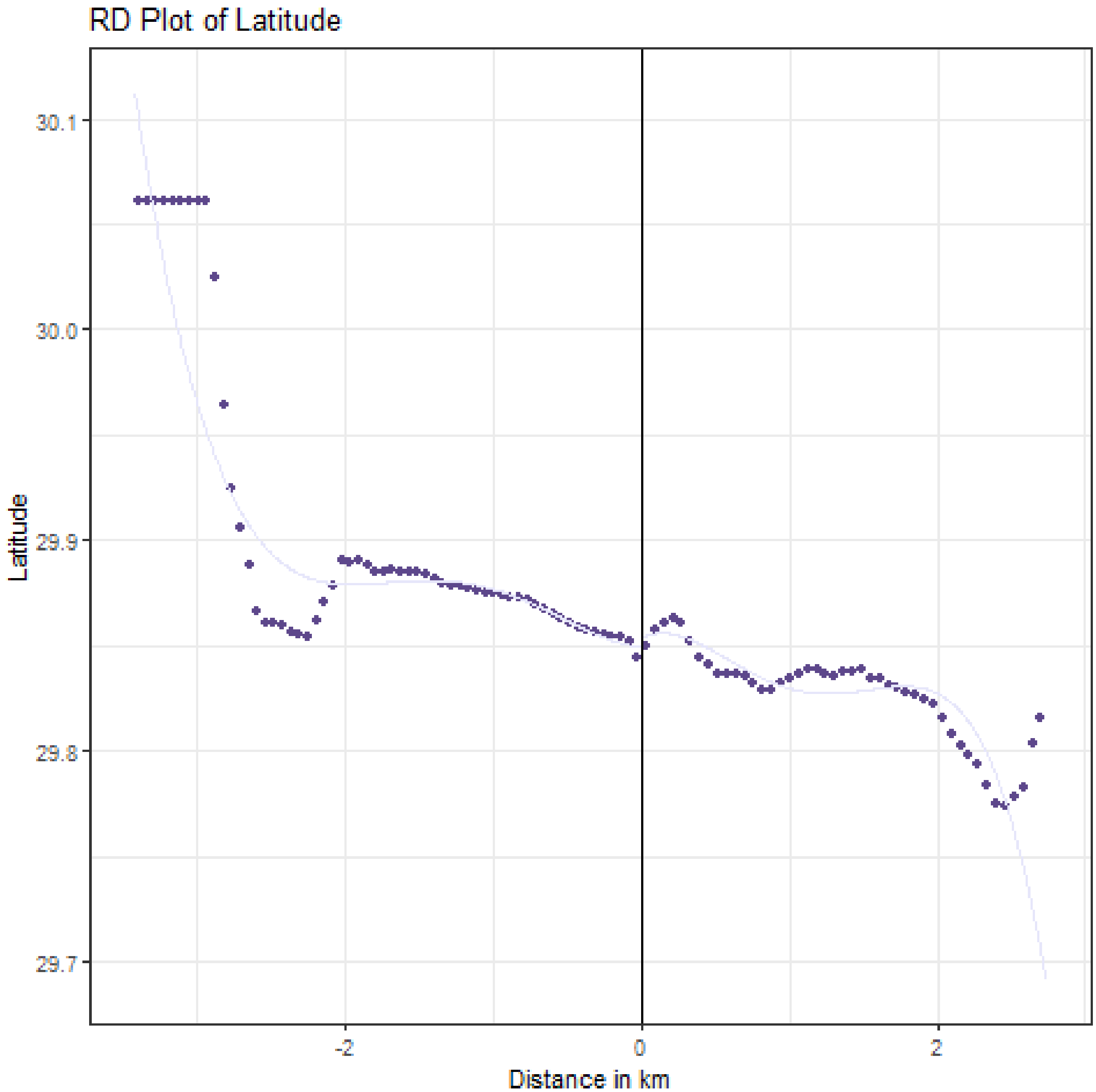

Figure 6.6: RD plot of latitude to distance from SFHA boundary. Positive distances represents buildings in the SFHA. Each point represents a binned mean of observations. The number of bins is selected using IMSE-optimal evenly-spaced method using polynomial regression.

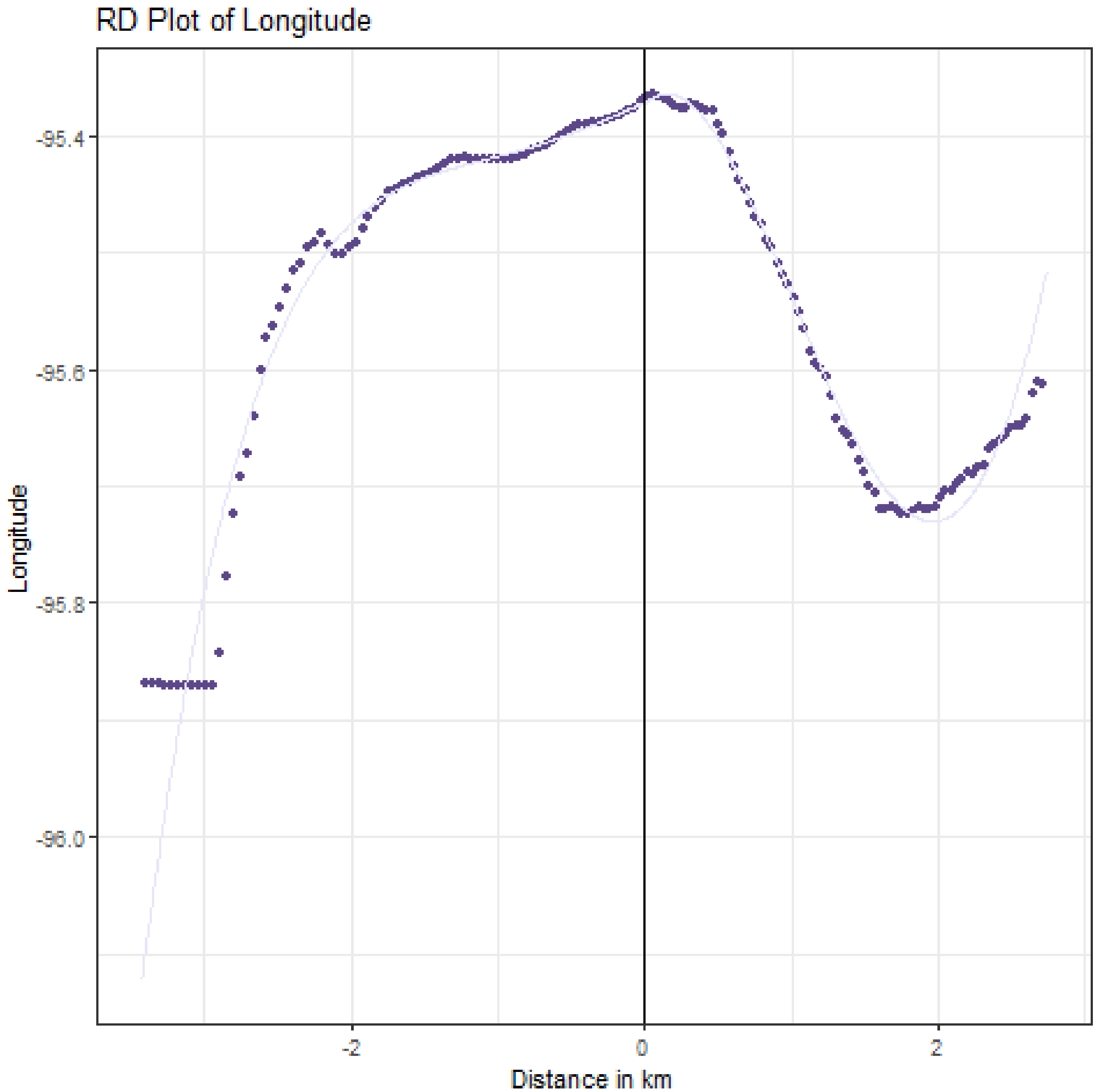

Figure 6.7: RD plot of longitude relative to distance from SFHA boundary. Positive distances represents buildings in the SFHA. Each point represents a binned mean of observations. The number of bins is selected using IMSE-optimal evenly-spaced method using polynomial regression.

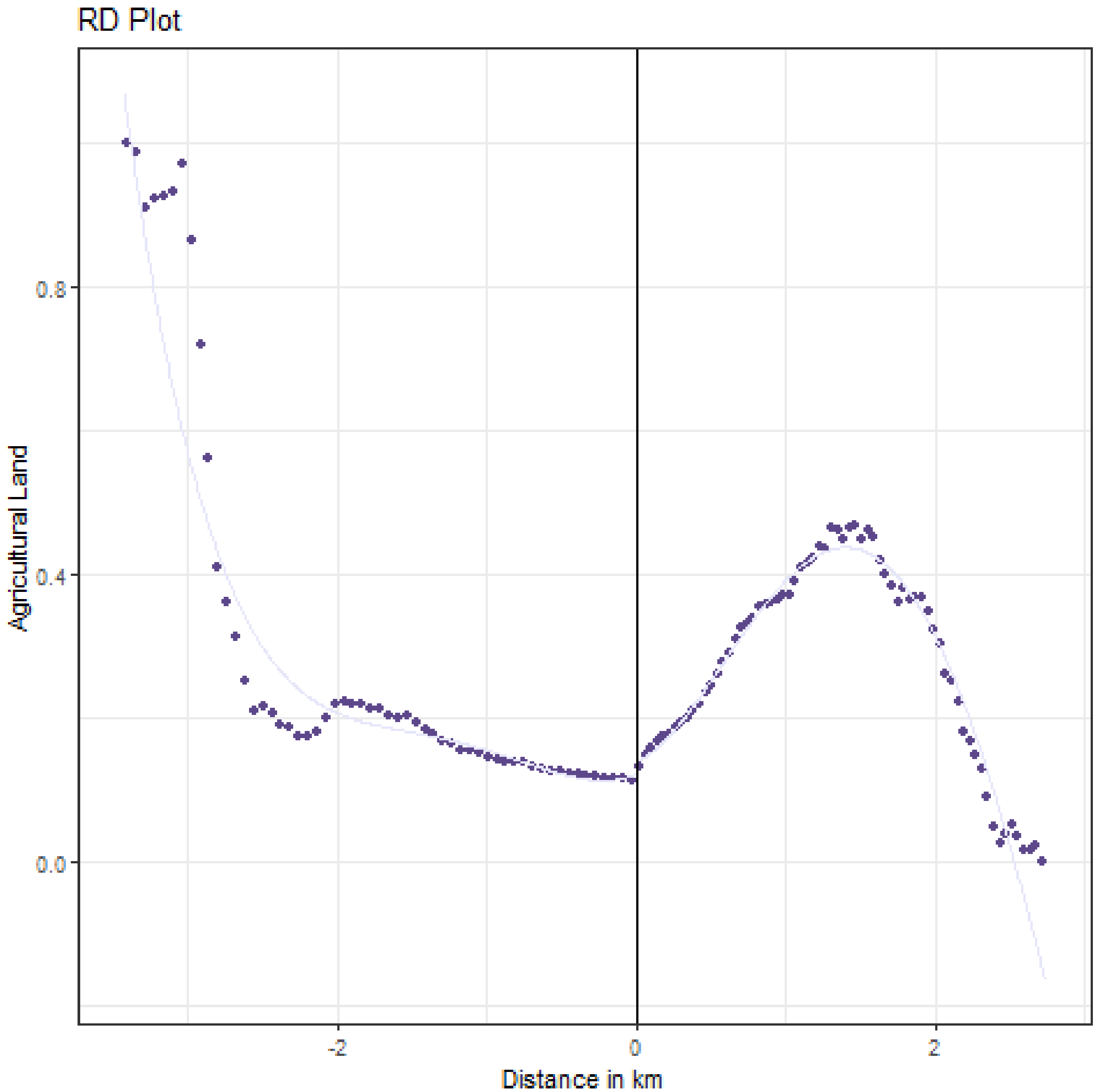

Figure 6.8: RD plot of density of agricultural land relative to distance from SFHA boundary. Positive distances represents buildings in the SFHA. Each point represents a binned mean of observations. The number of bins is selected using IMSE-optimal evenly-spaced method using polynomial regression.

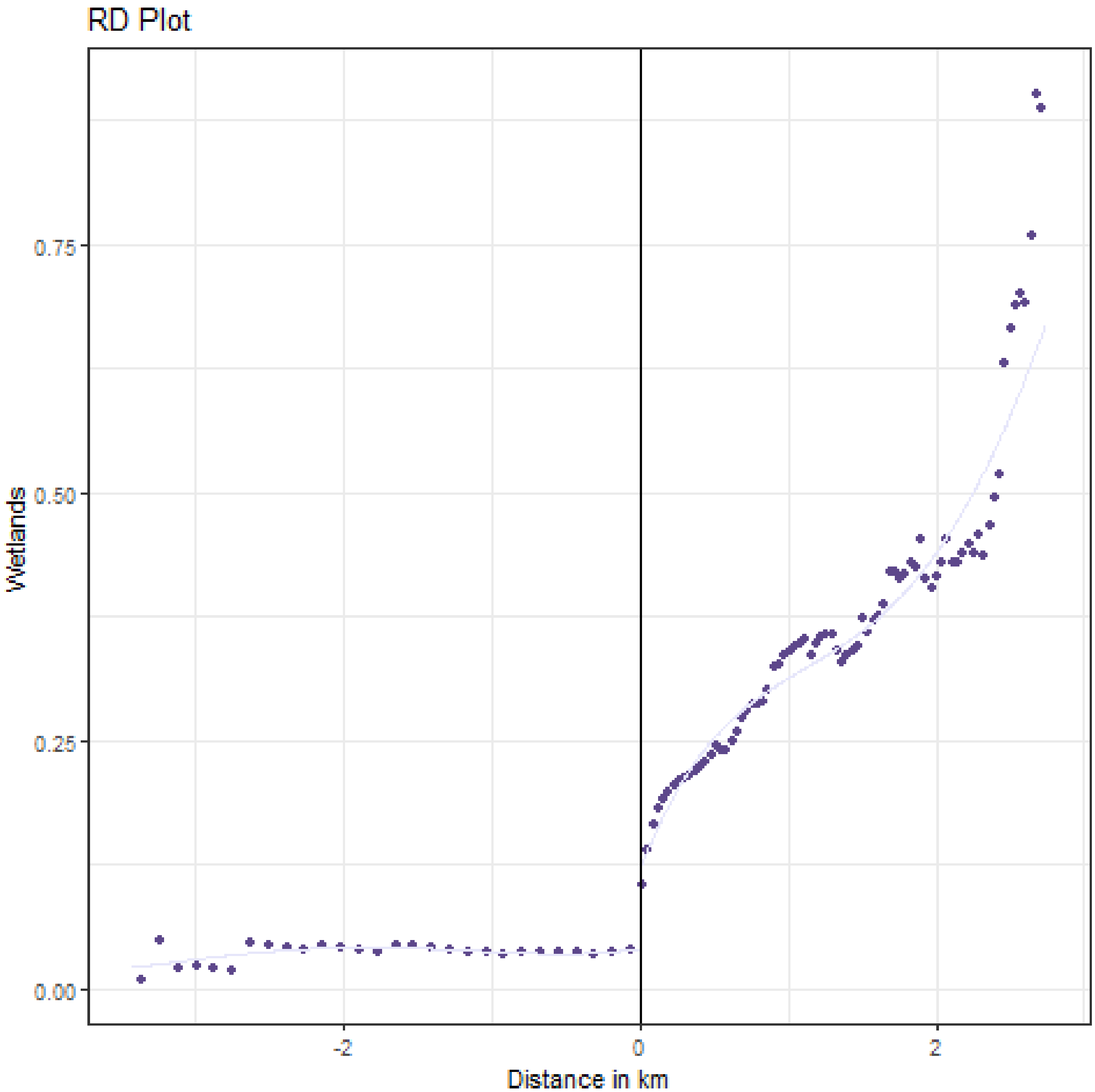

Figure 6.9: RD plot of density of wetlands relative to distance from SFHA boundary. Positive distances represents buildings in the SFHA. Each point represents a binned mean of observations. The number of bins is selected using IMSE-optimal evenly-spaced method using polynomial regression.

Table 6.3: Various Bandwidths

	(1)	(2)	(3)	(4)
SFHA	-0.156	-0.226	-0.144	-0.148
	(0.002)	(0.003)	(0.001)	(0.001)
Num.Obs.	1140290	611707	1926056	2520034
Bandwidth	0.157	0.078	0.313	0.470

Note: Column 1 is the main specification using the optimal bandwidth. Column 2, 3, and 4 are 50, 200, and 300 percent of the optimal bandwidth respectively.

Table 6.4: Various Cutoffs

	(1)	(2)	(3)	(4)
SFHA	-0.001	-0.002	-0.001	0.005
	(0.012)	(0.017)	(0.015)	(0.018)
Num.Obs.	416463	759186	652693	1537494
Bandwidth	0.074	0.080	0.177	0.193

Note: Each column represents the main specification with a placebo cutoff.

CONCLUSION

This paper provides evidence of a 64% decrease in land development for areas just inside the SFHA boundaries using a spatial regression discontinuity design. The conditional expectation function illustrated in the RD plots, provides evidence suggesting there is no sorting around the SFHA boundary line. This implies that if FEMA were to accurately estimate flood risk and expand the SFHA land development would decrease inside the high flood risk areas in Harris County, Texas, rather than reallocating development to non-SFHA.

The direct implications of expanding the SFHA will subject more properties to the flood insurance requirement and inform households that are living in high flood risk areas. The indirect implications of expanding the SFHA will vary for different communities. Expanding the SFHA to accurately estimate flood risk would benefit future home buyers allowing them to make informed purchasing decisions. This would also incentivize developers to act in the interest of future home buyers by reducing land development in high flood risk areas. However, there is $200 billion dollars in unpriced flood risk in the US housing market and accounting for flood risk in the housing market would result in more home equity losses in low-income communities (Gourevitch and Weill, 2023; Gourevitch et al., 2023).

The findings in this study also suggest that if FEMA were to continue to remove properties with high flood risk from the SFHA in map revisions, land development would increase in the excluded high flood risk areas. Excluding properties with high flood risk from the SFHA would incentivize developers to act against the some of the interests of home buyers by developing more in high flood risk areas. Continuing to overdevelop in high flood risk areas will exacerbate the severity of future flood damages and costs.

The main limitation of this study is the inclusion of SFHA boundaries that border levees, bayous, and parks. These areas would likely remain undeveloped without the SFHA designation and including them in this analysis is likely inflating the coefficients. This study also does not address the flood map revisions that invalidates the assumption that the boundaries are not manipulated. Previous work addressed this issue by using the initial flood maps (Ostriker and Russo, 2022). Future work should address these limitations along with expanding the geographic scope.

REFERENCES CITED

Ajita Atreya and Jeffrey Czajkowski. Graduated Flood Risks and Property Prices in Galveston County. *Real Estate Economics*, 47(3):807–844, 2019. ISSN 1540-6229. doi: 10.1111/1540-6229.12163. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1540-6229.12163. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1540-6229.12163.

Allan Beltrán, David Maddison, and Robert J R Elliott. Is Flood Risk Capitalised Into Property Values? *Ecological Economics*, 146:668–685, April 2018. ISSN 0921-8009. doi: 10.1016/j.ecolecon.2017.12.015. URL https://www.sciencedirect.com/science/article/pii/S0921800916314732.

Okmyung Bin and Jamie Brown Kruse. Real Estate Market Response to Coastal Flood Hazards. *Natural Hazards Review*, 7(4):137–144, November 2006. ISSN 1527-6988. doi: 10.1061/(ASCE)1527-6988(2006)7:4(137). URL https://ascelibrary.org/doi/10.1061/%28ASCE%291527-6988%282006%297%3A4%28137%29. Publisher: American Society of Civil Engineers.

Okmyung Bin and Craig E. Landry. Changes in implicit flood risk premiums: Empirical evidence from the housing market. *Journal of Environmental Economics and Management*, 65(3):361–376, May 2013. ISSN 0095-0696. doi: 10.1016/j.jeem.2012.12.002. URL https://www.sciencedirect.com/science/article/pii/S0095069612001179.

Okmyung Bin, Thomas W. Crawford, Jamie B. Kruse, and Craig E. Landry. Viewscapes and Flood Hazard: Coastal Housing Market Response to Amenities and Risk. *Land Economics*, 84(3):434–448, August 2008. ISSN 0023-7639, 1543-8325. doi: 10.3368/le.84.3.434. URL https://le.uwpress.org/content/84/3/434. Publisher: University of Wisconsin Press Section: Articles.

Lahouari Bounoua, Joseph Nigro, Ping Zhang, Kurtis Thome, and Asia Lachir. Mapping urbanization in the United States from 2001 to 2011. *Applied Geography*, 90:123–133, January 2018. ISSN 0143-6228. doi: 10.1016/j.apgeog.2017.12.002. URL https://www.sciencedirect.com/science/article/pii/S0143622817302643.

Sebastian Calonico, Matias D Cattaneo, and Max H Farrell. Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, 23(2):192–210, May 2020. ISSN 1368-4221, 1368-423X. doi: 10.1093/ectj/utz022. URL https://academic.oup.com/ectj/article/23/2/192/5625071.

James Chivers and Nicholas E. Flores. Market Failure in Information: The National Flood Insurance Program. *Land Economics*, 78(4):515–521, 2002. ISSN 0023-7639. doi: 10.2307/3146850. URL https://www.jstor.org/stable/3146850. Publisher: [Board of Regents of the University of Wisconsin System, University of Wisconsin Press].

Beverly A. Cigler. U.S. Floods: The Necessity of Mitigation. *State and Local Government Review*, 49(2):127–139, June 2017. ISSN 0160-323X. doi: 10.1177/0160323X17731890. URL https://doi.org/10.1177/0160323X17731890. Publisher: SAGE Publications Inc.

Vanessa E. Daniel, Raymond J. G. M. Florax, and Piet Rietveld. Flooding risk and housing values: An economic assessment of environmental hazard. *Ecological Economics*, 69(2): 355–365, December 2009. ISSN 0921-8009. doi: 10.1016/j.ecolecon.2009.08.018. URL https://www.sciencedirect.com/science/article/pii/S092180090900322X.

Piet M. A. Eichholtz, Eva Steiner, Erkan Yönder, and Jawad M. Addoum. Climate Change and Commercial Real Estate: Evidence from Hurricane Sandy, March 2021. URL https://papers.ssrn.com/abstract=3206257.

Daryl Fairweather, Matthew E Kahn, Robert D Metcalfe, and Sebastian Sandoval-Olascoaga. PRELIMINARY: The Impact of Climate Risk Disclosure on Housing Search and Buying Dynamics: Evidence from a Nationwide Field Experiment with Redfin.

Justin Gallagher. Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States. *American Economic Journal: Applied Economics*, 6(3): 206–233, July 2014. ISSN 1945-7782, 1945-7790. doi: 10.1257/app.6.3.206. URL https://pubs.aeaweb.org/doi/10.1257/app.6.3.206.

Jesse D. Gourevitch and Joakim A. Weill. Flood risks are insufficiently priced into housing markets but better pricing would leave some worse off. *Nature Climate Change*, 13(3): 216–217, March 2023. ISSN 1758-6798. doi: 10.1038/s41558-023-01598-4. URL https://www.nature.com/articles/s41558-023-01598-4. Publisher: Nature Publishing Group.

Jesse D. Gourevitch, Carolyn Kousky, Yanjun (Penny) Liao, Christoph Nolte, Adam B. Pollack, Jeremy R. Porter, and Joakim A. Weill. Unpriced climate risk and the potential consequences of overvaluation in US housing markets. *Nature Climate Change*, 13(3):250–257, March 2023. ISSN 1758-6798. doi: 10.1038/s41558-023-01594-8. URL https://www.nature.com/articles/s41558-023-01594-8. Publisher: Nature Publishing Group.

David Harrison, Greg T. Smersh, and Arthur Schwartz. Environmental Determinants of Housing Prices: The Impact of Flood Zone Status. *Journal of Real Estate Research*, 21(1-2):3–20, January 2001. ISSN 0896-5803. doi: 10.1080/10835547.2001.12091045. URL https://doi.org/10.1080/10835547.2001.12091045. Publisher: Routledge _eprint: https://doi.org/10.1080/10835547.2001.12091045.

Miyuki Hino and Marshall Burke. The effect of information about climate risk on property values. *Proceedings of the National Academy of Sciences*, 118(17):e2003374118, April 2021. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.2003374118. URL https://pnas.org/doi/full/10.1073/pnas.2003374118.

David Hunn, Matt Dempsey, and Mihir Zavier. In Harvey's deluge, most damaged homes were outside the flood plain, new data show, March 2018. URL https://www.houstonchronicle.com/news/article/In-Harvey-s-deluge-most-damaged-homes-were-12794820.php.

Zach Levitt and Jess Eng. Where America's developed areas are growing: 'Way off into the horizon', 2021. URL https://www.washingtonpost.com/nation/interactive/2021/land-development-urban-growth-maps/.

Francesc Ortega and Süleyman Taṣpınar. Rising sea levels and sinking property values: Hurricane Sandy and New York's housing market. *Journal of Urban Economics*, 106: 81–100, July 2018. ISSN 0094-1190. doi: 10.1016/j.jue.2018.06.005. URL https://www.sciencedirect.com/science/article/pii/S0094119018300354.

A. Ostriker and A. Russo. The Effects of Floodplain Regulation on Housing Markets. 2022. URL https://www.semanticscholar.org/paper/The-Effects-of-Floodplain-Regulation-on-Housing-Ostriker-Russo/d570d2fb4e00317c310f7456c829947242002564.

Sarah Pralle. Drawing lines: FEMA and the politics of mapping flood zones. *Climatic Change*, 152(2):227–237, January 2019. ISSN 1573-1480. doi: 10.1007/s10584-018-2287-y. URL https://doi.org/10.1007/s10584-018-2287-y.

Joel Scata. FEMA's Outdated and Backward-Looking Flood Maps, October 2017. URL https://www.nrdc.org/bio/joel-scata/femas-outdated-and-backward-looking-flood-maps.

Melissa Tier. Overcoming Contemporary Reform Failure of the National Flood Insurance Program to Accelerate Just Climate Transitions.

Joakim A. Weill. Flood Risk Mapping and the Distributional Impacts of Climate Information. *Finance and Economics Discussion Series*, (2023-066):1–95, October 2023. ISSN 1936-2854, 2767-3898. doi: 10.17016/feds.2023.066. URL https://www.federalreserve.gov/econres/feds/flood-risk-mapping-and-the-distributional-impacts-of-climate-information.htm.

Oliver E. J. Wing, William Lehman, Paul D. Bates, Christopher C. Sampson, Niall Quinn, Andrew M. Smith, Jeffrey C. Neal, Jeremy R. Porter, and Carolyn Kousky. Inequitable patterns of US flood risk in the Anthropocene. *Nature Climate Change*, 12(2):156–162, February 2022. ISSN 1758-6798. doi: 10.1038/s41558-021-01265-6. URL https://www.nature.com/articles/s41558-021-01265-6. Publisher: Nature Publishing Group.

Hessel C. Winsemius, Jeroen C. J. H. Aerts, Ludovicus P. H. van Beek, Marc F. P. Bierkens, Arno Bouwman, Brenden Jongman, Jaap C. J. Kwadijk, Willem Ligtvoet, Paul L. Lucas, Detlef P. van Vuuren, and Philip J. Ward. Global drivers of future river flood risk. *Nature Climate Change*, 6(4):381–385, April 2016. ISSN 1758-6798. doi: 10.1038/nclimate2893.

URL https://www.nature.com/articles/nclimate2893. Publisher: Nature Publishing Group.

Wei Zhang, Gabriele Villarini, Gabriel A. Vecchi, and James A. Smith. Urbanization exacerbated the rainfall and flooding caused by hurricane Harvey in Houston. *Nature*, 563(7731):384–388, November 2018. ISSN 1476-4687. doi: 10.1038/s41586-018-0676-z. URL https://www.nature.com/articles/s41586-018-0676-z. Publisher: Nature Publishing Group.