Measuring House Price Externalities

Evidence from Hurricane Harvey in Houston

Author: Jan Dudek

Advisor: Diego Puga

June 20th, 2018

Abstract

Hurricane Harvey flooded almost 10 percent of all buildings in Houston. I examine how the valuation of non-flooded homes changed due to a change in their neighborhood – damage to residential and commercial buildings around them. I find the housing externalities from both residential and commercial buildings to be significant and large, with magnitudes around 7 percent. While the externalities due to commercial buildings dissipate at 500 meters, the effect remains present until 800 meters for residential buildings. Finally, I find large "shortage" premia for non-flooded homes located in appealing neighborhoods.

I am very grateful to my advisor Dr. Diego Puga for his guidance in this analysis. I also want to thank professors who attended the Master thesis workshops and gave me valuable comments and suggestions. I want to thank Dr. Robert Stein for aiding me in finding data and giving helpful comments. Finally, I am grateful to Micole De Vera for his help and unrelenting patience.

1 Introduction

A building's valuation is determined by internal factors, such as the number of bedrooms and bathrooms, square footage, or the quality of the building materials used – all of which can be directly affected by the owner by means of investment. It is also determined by the house's location, which gives it access to non-market interactions, such as distance to city center, presence of amenities and commercial stores, or the appeal of surrounding homes. Homeowners cannot directly and unilaterally affect these non-market interactions, which is why they are external to them. For example, if a neighbor fixes her house, she only takes into account the price effect on her own home, but not on neighboring houses, which likely also benefit. Consequently, she will invest in home repairs less than would be socially optimal. Similarly to Rossi-Hansberg et al. (2010), I will define these non-market interactions as housing externalities.

Evidence of the importance of these externalities is the prevalence of neighborhood-level covenants, which define a set of rules about the appearance and maintenance of a house – for example that the lawn must be mowed at least once per two weeks. Given the large urban density in modern cities, it is virtually impossible to avoid housing externalities caused by nearby buildings. As with any externality, housing externalities create inefficiencies, which can only be studied if their magnitude is measured. Without their proper understanding, an analysis of the value of an urban policy, such as building a park or subsidizing home repairs, cannot be conducted. The question also has large economics significance: the average American holds 69 percent of savings in housing property, whose value is directly influenced by these externalities (US Census Bureau, 2017).

Ordinarily, one can observe and estimate a building's hedonic price Rosen (1974), which is determined in a spatial equilibrium and the entire set of implicit prices guides buyers' and sellers' decision on where to locate and what internal and external house characteristics to choose. This means an observed change in the market value of a house may be due to a change in external factors, internal factors, or their combination. In order to identify the magnitude of external effects, a researcher must control for each house's varying internal characteristics. Another issue for identifying the housing externalities is reverse causality: neighborhoods become more attractive and well-equipped once many people wish to move in and are willing to pay a premium, thus increasing their market values; at the same time, people's willingness to pay for a house grows with the neighborhood's attractiveness and the quality of amenities.

Therefore, the most common method of studying housing externalities is through a natural experiment, which introduces exogenous variation in factors producing externalities, but (in the ideal case) keeps all internal home characteristics constant. In this paper, I use the exogenous variation brought by hurricane Harvey in the city of Houston in Texas (USA) in August 2017. While the storm brought only moderate winds, it introduced almost a year's worth of rain within several days. Due to the variation in the coverage and the amount of rain in the London-sized metropolis, some neighborhoods received a meter of rain, while others stayed completely dry. I argue that focusing on individual non-flooded houses, whose internal characteristics remained constant, but whose neighbors received various levels of flooding, allows for the identification of housing externalities.

In this paper, I aim to identify the housing externality effect caused by flooded residential or commercial buildings. However, two other effects are present and influence changes in house prices: The "risk-updating effect" and the "shortage effect." The risk-updating effect describes the market's update of beliefs about future damages, such as a higher risk of flooding. Coulomb and Zylberberg (2016) find that after the Fukushima nuclear accident in Japan, house values in Wales located close to a nuclear plant decreased by 3.5 percent due to the heightened risk perception of nuclear hazard. The "shortage effect," occurs through market fundamentals and may cause house prices to increase in case demand for dry housing in a given area surpasses supply. For example, after the devastating hurricane Katrina in New Orleans, the price of a median house increased by 9 percent even though a large part of the population left and never returned (Vigdor, 2008).

In this analysis, I find housing externalities due to residential and commercial buildings to be statistically significant and large, with magnitudes around negative 7 percent.

Contrary to the literature, I find the externalities to be non-monotonic over space, with negative externalities for areas close to a house, and large positive externalities for flooded areas further away. I present evidence that the positive externality is caused by the shortage effect, and is driven by neighborhoods with high appeal. Finally, I show negative housing externalities dissipate faster for commercial buildings at 500 meters, whereas they stay present for residential buildings up to 800 meters.

2 Literature Review

Various authors have used the concept of a hedonic equilibrium to study housing externalities related to amenities, such as being close to a river or park, or of disamenities, such as proximity to hazardous waste sites and polluted water bodies (McCluskey and Rausser, 2001; Bin et al., 2008; Leggett and Bockstael, 2000). However, there are multiple issues of interpretability and identification with regards to hedonic price equilibria. Leggett and Bockstael (2000) illustrate the presence of omitted variable bias in a number of studies using a hedonic equilibrium, since risk factors, such as emitters of hazardous waste or pollutants, are likely correlated with other disamenities, for example noise or unpleasant view of a nearby factory. By ignoring this issue, one can overstate the magnitude of the studied externality effect. Conversely, if risk factors are correlated with positive amenities, the externality will be biased downwards. For example, a house location close to a river increases the risk of flooding, but provides access to amenities associated with rivers, such as parks or fishing.

In their chapter on spatial econometrics, Anselin and Bera (1998) discuss another major concern with hedonic equilibria: houses close to each other are frequently similar, for example due to shared amenities or zoning restrictions. Ignoring this issue brings endogeneity to the model: If the matrix W defines the neighbor relationship between different houses (such as a one for each house within 100 meters), a possible model for prices is $y = \rho W y + X\beta + \epsilon$, where ρ measures the strength of the spatial relationship, X are control variables and ϵ is a vector of error terms. By construction, the model suffers from endogeneity and ignoring W will imply inconsistent estimators of β. This problem can be overcome by using repeated sale transactions or market valuations over time for the same houses, so that all time-invariant characteristics including spatial interactions are accounted for as fixed effects (Hallstrom and Smith, 2005).

Due to the identification issues related to hedonic equilibria, natural experiments are used to identify housing externalities: Rossi-Hansberg et al. (2010) exploit a government urban policy, in which the state of Virginia (USA) subsidized repairs for houses in selected disadvantaged neighborhoods. The authors find that a house undergoing repairs increases the value of nearby housing by 10 percent, a surprisingly large figure. They also show that the externality approximately halves every 300 meters. In a more recent paper, Currie et al. (2015) study the effect of opening a polluting plant near a residential neighborhood and find a negative externality of 11 percent extending to a distance of 800 meters. Diamond and McQuade (2016) exploit the building of affordable housing for low-income groups by the government. They find externalities heterogeneous by neighborhood wealth: while the project increased the value of nearby houses in low-income neighborhoods by 6.5 percent, it decreased it in high-income neighborhoods by 2.5 percent.

3 Hurricane Harvey as a Natural Experiment

Following the recent literature, I use hurricane Harvey as a natural experiment and exploit the exogenous variation in the city's flooding to identify housing externalities. Hurricane Harvey stalled above the city of Houston on August 25th, 2017 and remained there for several days, delivering massive amounts of rainfall. In the span of three days, Harris County (which comprises most of the Houston agglomeration) received 74 centimeters of rain on average, with some rain gauges recording as much as 107 centimeters of rain. On average, Houston receives about 114 centimeters of rainfall *per year* (Harris County Flood Control District, 2017). Since the amount of rain within the city varied greatly, some residents had to be rescued by helicopters from the roofs of their houses, while others stayed completely dry. While the suburbs and outskirts of the city received the

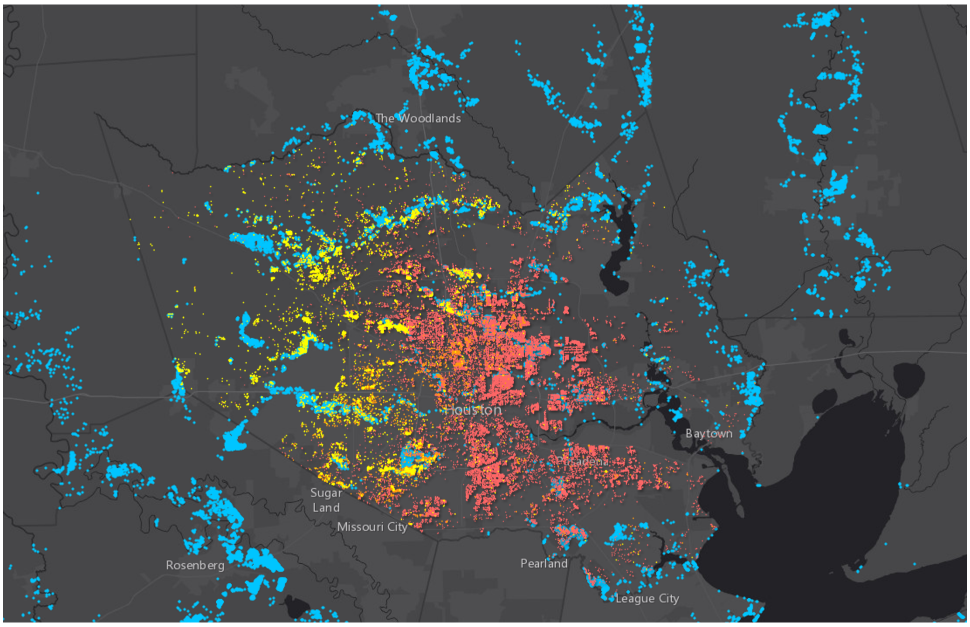

Figure 1: Past Flooding

most damage, more than half of all zip codes submitted more than 2,500 applications for federal disaster assistance (Federal Emergency Management Agency, 2017b). Although it is not yet clear how much damage was inflicted, preliminary estimates range between 25 and 37 billion dollars (6-8 percent of Houston's GDP), making it the second costliest disaster in US history only surpassed by hurricane Katrina (Core Logic, 2017).

Hurricane Harvey is a rare large-scale natural experiment, which can shed light on flood damage externalities. However, in order to identify externality effects using OLS, it is required the flooding is fully exogenous and unanticipated. I argue the hurricane brought an exogenous wave of flooding for several reasons: First, 78 percent of flooding happened outside of government-designated 100-year flood zones, where the chance of flooding should be minimal. One may worry owners received additional information from past floods, several of which ocurred in recent history. A map from the Kinder Institute (2017) in Figure 1 shows flooded houses during four large flooding events in Houston: Tropical Storm Allison in 2001 (red), Memorial Day Flood in 2015 (orange), Tax Day Flood in 2016 (yellow) and hurricane Harvey in 2017 (blue). It is clear that while previous storms mostly affected central Houston, hurricane Harvey largely affected suburbs and outskirts of the city, which were safe during all previous flood events.

Figure 2: Satellite Image of Flooding

Second, all the flooding was due to large amounts of rainfall, which is induced by exogenous weather events – as mentioned previously, some parts of the city received 107 centimeters rain, while others none at all. And finally, since many neighborhoods were flooded for the first time in decades or first time ever (many suburbs were newly built), homeowners were not aware if their house was slightly more elevated than the rest, or if the slant of the street was such that the flooding would not reach their home. In fact, almost every flooded neighborhood proved to have a few lucky houses, which remained dry, even though they appeared similar to their neighbors. An example can be seen in Figure 2, where a few houses at the bottom of the image remained dry, while the rest received flooding.

4 Methodology

To measure externality effects caused by nearby flooded commercial and residential buildings, I propose to use the universe of updated market valuations from HCAD in a difference-in-differences framework. As my main source of variation, I exploit the changes in a house's surroundings in terms of the proportion of buildings flooded. I choose to work with the HCAD data instead of the true transactions from Zillow in order to avoid selection sample problems, where houses sold before a hurricane are likely different from those sold right after. The disadvantage of using HCAD data is their questionable reliability. In section 5, I show that the HCAD valuations are highly informative of the true transactions before the hurricane as well as after it. As for the house pricing model, I assume a semi-log specification based on previous literature, such as Bin et al. (2008) and Hallstrom and Smith (2005):

$$\log P_{it} = x_i'\gamma + \sum_k \delta_k F_{Res,k} + \sum_l \theta_l F_{Com,l} + \alpha s_{it} + \beta_t e_i + \pi_t d_i + w z_i + \eta_i + \epsilon_{it} \tag{1}$$

Here, P_{it} are house prices, x_i represents time-invariant house characteristics (such as square footage or number of bedrooms), $F_{Res,k}$ represents the proportion of residential buildings flooded in a radius of k meters, while $F_{Com,l}$ is the proportion of commercial buildings flooded in a radius l meters. s_{it} represents damage caused by flooding (for houses directly hit), e_i stands for a house's elevation relative to the sea level and d_i for the distance to the nearest major river. z is a binary variable which takes the value of one if a house is within a government-designated flood zone. Finally, η_i stands for all other time-invariant house fixed effects and ϵ_{it} represents a house-specific, time-varying iid error term. The proportion flooded measure has a "donut-like" shape, so that the flooding in an area of a larger radius is considered, but the area of the smaller radius, which had already been taken into account, is excluded (this method makes interpretation more intuitive).

The main objects of interest in this study are δ_k and θ_l. Based on the previous literature, they could take on negative values due to the negative amenity effect, or a positive value due to the shortage effect, depending on the strength of each effect. The negative amenity effect should be decreasing in distance, while this relationship is less clear for the shortage effect. I attempt to control for the risk-updating effect by including risk-related measures of elevation, distance to river and flood-zone designation for each house.

In order to find out the rate of dissipation of the externality effect, I consider increasing radii of 200 and 250 meters for surrounding residential and commercial buildings, respectively (the results are relatively robust to choosing different radii, but with smaller radii differences, many data points have to be thrown out as there are no additional residential or commercial buildings added in at least one "donut"). To estimate the repeated-price model, I take first differences in order to eliminate time invariant effects of x_i and η_i:

$$\Delta \log P_{it} = \sum_k \delta_k \Delta F_{Res,k} + \sum_l \theta_l \Delta F_{Com,l} + \alpha \Delta s_{it} + \Delta\beta_t e_i + \Delta\pi_t d_i + \Delta w z_i + \Delta\tau + \Delta\epsilon_{it} \quad (2)$$

Even though a house's distance to a river, elevation or flood zone designation remains constant over time, I hypothesize the *perception* of these variables can change, due to the risk updating effect, for example if people now value houses far away from rivers more due to their higher safety from flooding. Note the variable $\Delta\tau$ accounts for city-wide price changes, such as due to price growth related to national economy. An important problem here is the lack of information on damages Δs_{it}. Given knowledge of which houses were flooded, it is still difficult to assess how much water got inside and to estimate damage levels. A possible proxy would be the number of FEMA applications for the purposes of Shelter or Emergency, which would suggest serious flooding damage. However, this information only comes at the zip code level, which is insufficient. Ignoring damages could bias the results, since some owners are forced to sell damaged homes far below their previous market value; others with more cash may use flooding repair works as an opportunity for a larger house renovation, after which the market value can easily increase, despite the previously sustained damages (for example some owners spends tens of thousands to elevate the entire home by several meters). Due to lack of information about damages and repairs, I only consider houses which were not flooded for my main specification. The main regression, run on the subset of non-flooded homes is the following:

$$\Delta \log P_{it} = \sum_k \delta_k \Delta F_{Res,k} + \sum_l \theta_l \Delta F_{Com,l} + \Delta\beta_t e_i + \Delta\pi_t d_i + \Delta w z_i + \Delta\tau + \Delta\epsilon_{it} \quad (3)$$

To ensure unbiasedness of the results using OLS, independence of $\Delta\epsilon_{it}$ from all the explanatory variables is required (this is ensured by the exogenous nature of flooding). To ensure the validity of standard errors, independent and identically distributed errors are required. One worry may be the dependence of errors for houses in the same neighborhood or due to geographical proximity. In a robustness check, I therefore run a cluster-robust least squares regression.

In a secondary specification, I also include the flooded homes, with the caveats of underpriced sales due to existing damages or overpriced sales due to additional investments during reconstruction. The coefficients are allowed to vary depending on whether flooding is present (the time effect and the risk-controls such as elevation are omitted here for sake of readability, but they are included in the regression). The function $G(i)$ "checks" the flooding status of every observation.

$$\Delta \log P_{it} = \sum_{a=f,nf} \mathbb{1}\{G(i) = a\} \left(\sum_k \delta_{k,a} \Delta F_{Res,k} + \sum_l \theta_{l,a} \Delta F_{Com,l} + \Delta \beta_{t,a} e_i \right) + \Delta \epsilon_{it} \qquad (4)$$

5 Data

I use data from the Harris County Appraisal District (HCAD), a political subdivision of the State of Texas, which sets property values for more than 1.7 million parcels of property every several years. Because of constrained resources, the office updates prices for a random portion of houses every year (in my analysis, I use the parcels which have been updated). This evaluation is essential for the city's functioning – since Texas has no income tax, services such as schools, city administration and sewage systems are funded from property taxes. In 2016, each property owner paid on average 2.54 percent of the value of their property in tax. The timing of the valuations is as follows: owners receive the first preliminary evaluations in April, after which they have the right to appeal the decision in case they consider the proposed valuation to be incorrect. For example, they can argue that a neighbor just sold a house for a low price, which is why the valuation of their own house is too high. After all cases are settled, the appraisal office sends a final valuation for each property by mid-August. For my main dataset, I use HCAD values from April 2017 (about 4 months before the hurricane) and April 2018 (about 8 months after the hurricane), which gives me repeated market values for $155,653$ houses in the area of analysis for which price was updated. The HCAD data also contains detailed geolocation data, information about the square footage of the house and of the land, as well as whether the home owner contested the proposed market valuation.

Table 1 shows the summary statistics for this dataset. On average, the overall price of an average house in Houston grew by 9 percent despite the hurricane. An average house would be elevated more than 19 meters above the sea level and 2.3 kilometers away from the nearest river. The share of homes inside of 100-year flood zones was 7 percent, which was the same as the share of houses which received flooding (however, the overlap between flood zones and flooding was small as shown in Table 3). Finally, the proportion of nearby flooded residential homes for various radii ranges from 0 to 1 with an average standard deviation of 0.12. The proportion of nearby flooded commercial buildings also varies from 0 to 1 and has an average standard deviation of 0.14.

Table 1: Summary Statistics HCAD

	Mean	Standard Deviation	Minimum	Maximum
Square Footage House	1919.94	(980.83)	252.00	16284.00
Square Footage Land	7073.28	(2999.08)	92.00	22019.00
Log Price Before	11.89	(0.71)	10.13	15.55
Log Price After	11.97	(0.69)	10.15	15.57
Elevation	19.27	(10.34)	0.07	64.73
Distance to River (km)	2.30	(1.50)	0.00	8.81
100-Year Flood Zone	0.07	(0.25)	0.00	1.00
Flooded	0.07	(0.26)	0.00	1.00
Prop Flooded Res 200m	0.07	(0.14)	0.00	1.00
Prop Flooded Res 600m	0.07	(0.11)	0.00	1.00
Prop Flooded Res 1000m	0.08	(0.10)	0.00	1.00
Prop Flooded Res 2000m	0.07	(0.08)	0.00	0.82
Prop Flooded Com 250m	0.09	(0.20)	0.00	1.00
Prop Flooded Com 500m	0.07	(0.15)	0.00	1.00
Prop Flooded Com 1000m	0.07	(0.12)	0.00	1.00
Prop Flooded Com 2000m	0.07	(0.09)	0.00	1.00
Observations	138177			

To find out which houses were flooded, I use a flooding model from the US Army Corps of Engineers and researchers from University of Central Florida (USACE, 2017). This model is based on rain, elevation and soil permeability data and offers information detailed enough to classify individual houses as flooded or dry. I cross-checked a number of the model's predictions with detailed satellite images purchased and released by the National Oceanic and Atmospheric Administration (2017), and the results largely matched. I also considered an alternative flooding dataset by the FEMA (2017), but I found it insufficient to classify individual houses in terms of flooding.

As a secondary data source, I gather data on house prices in Houston by web-scraping the website zillow.com. Zillow.com is searchable database of real estate data in the US, offering information on actual sales of housing. Due to a Houston ordinance, every house purchase has to be reported to the HCAD office, and this information is available to anyone on request, making it an ideal source of data for Zillow. I use Python code based on Muir (2018) with some of my own additions. The resulting dataset contains $18,562$ units, with 17,145 units sold before September 1, 2017 (when the hurricane and flooding had occurred), and $1,417$ units sold after. Since most of the houses sold before and after are different units, there are issues with unobserved heterogeneity in their internal characteristics. Because of the possible sample selection and limited sample size, I only use this data to check the robustness of the data acquired from HCAD. The summary statistics for the Zillow dataset are included in the Appendix in tables 7 and 8.

To provide more evidence about the reliability of HCAD data, I provide a comparison between true transaction prices observed via Zillow and market value estimates from HCAD. I find that the correlation between the true prices and the HCAD estimate to be 91.3 percent and 89.8 percent for prices before and after the hurricane, respectively, suggesting a very strong correspondence between HCAD valuations and true transaction prices. A standard least square regression of the logarithm of true price (from Zillow transactions) on the logarithm of the HCAD estimate gives coefficients of 0.994 and 0.973, where the hypothesis, that the coefficient is 1, cannot be rejected at the 95 percent confidence level (the constant term is not significantly different from zero in either regression). A scatterplot of the two variables after the hurricane is shown in Figure 3 This offers evidence that the HCAD estimates are relatively accurate, which is in line with

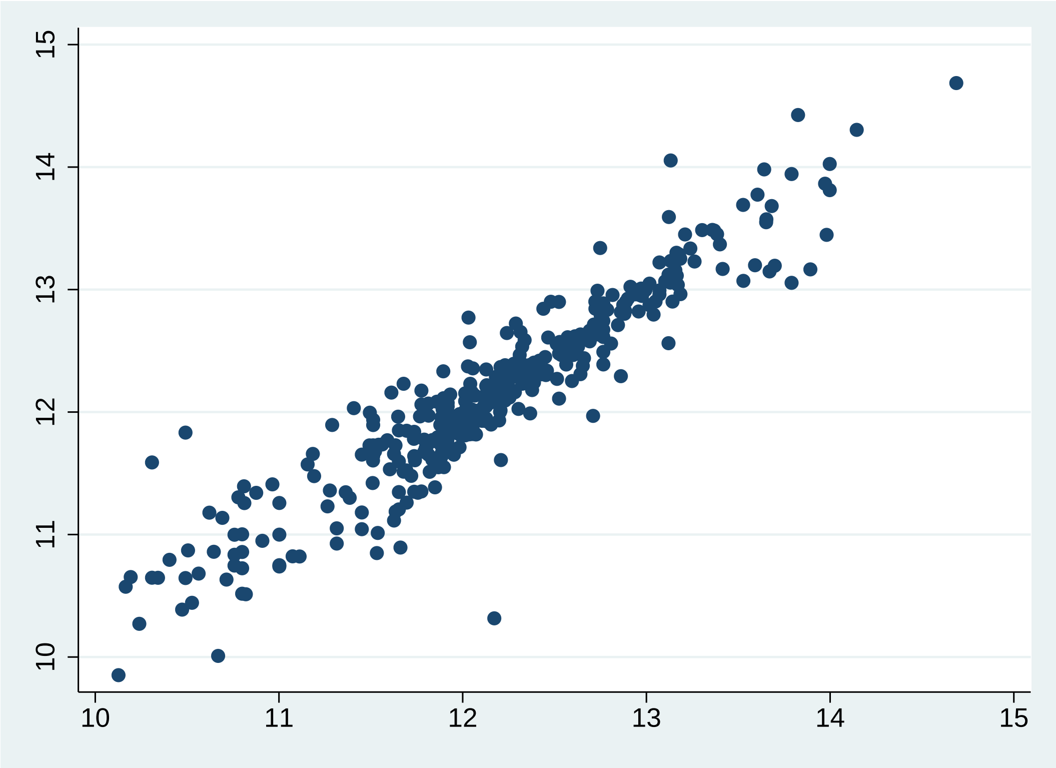

Figure 3: True Prices (Zillow) vs. HCAD Market Valuations in Log Terms

their claim that actual transactions are taken into account in their price-setting model.

To account for the risk-updating effect, I control for elevation, distance to a river and flood-zone designation at the individual house level. Houston-Galveston Area Council (2018) provides data for elevation and a river shapefile. The official flood zone shapefile is provided by Federal Emergency Management Agency (2017a).

Figure 9 shows a graphic representation of price changes due to hurricane Harvey using the main (HCAD) dataset for non-flooded homes. Specifically, it shows deviations from the average growth of market values between 2017 and 2018 in order to show a cleaner picture of changes due to within-city effects. Each bar's height is proportional to the magnitude of the deviation in unit of percent. There are several interesting observations to be made: There is a large number of green bars (positive deviations from mean), which are frequently surrounded by red bars – this suggests large within-neighborhood variation, likely due to some lucky houses which remained dry and which now receive a premium due to the shortage of usable housing stock. Figure 10 shows only negative price deviations. Most of them are located very close to flood zones, and given these are houses which remained dry, but which fell in value, it is likely due to the combination of the risk-updating and the externality effects.

Before the hurricane, flooded and non-flooded houses are relatively similar as can be seen in Tables 2 and 3, with those flooded actually on average *more expensive* than those that remained dry. The reason is much of the flooding happened in affluent suburbs in the western part of Houston. On average, the flooded houses are also closer to a major river by 450 meters and elevated lower by 3 meters than the non-flooded homes. This gives some ground to using the measures of elevation and distance to rivers to control for the risk-updating effect.

Table 2: Summary HCAD Non-Flooded

	Mean	Standard Deviation	Minimum	Maximum
Square Footage House	1918.12	(969.20)	252.00	16284.00
Square Footage Land	7140.88	(2981.89)	92.00	22019.00
Log Price Before	11.89	(0.70)	10.13	15.55
Elevation	19.49	(10.57)	0.07	64.73
Distance to River (km)	2.33	(1.51)	0.00	8.81
100-Year Flood Zone	0.06	(0.23)	0.00	1.00
Observations	128330			

Table 3: Summary HCAD Flooded

	Mean	Standard Deviation	Minimum	Maximum
Square Footage House	1943.55	(1121.18)	392.00	15256.00
Square Footage Land	6192.22	(3082.13)	718.00	21900.00
Log Price Before	11.90	(0.83)	10.13	15.16
Elevation	16.41	(5.80)	3.48	39.59
Distance to River (km)	1.88	(1.30)	0.03	6.66
100-Year Flood Zone	0.18	(0.38)	0.00	1.00
Observations	9847			

6 Empirical Results

In this section, I present the empirical results of the repeated market valuation model in context of related literature and interpret the findings. Table 4 presents the main regression table of this analysis. The first column presents an OLS regression of price changes versus proportions of flooded buildings nearby (residential and commercial). The second regression controls for the risk-updating effect by including elevation, distance to a major river and flood zone designation as control variables, as defined in Equation 3. Finally, the third regression controls for risk as well as fixed effects at the neighborhood level. The results are largely robust to the different specifications, which is why the following interpretation applies to all three specifications. Each regression controls for whether a home-owner protested against the proposed valuation by the city.[1]

The regressions shows an interesting non-monotonic externality effect with regards to residential buildings, as shown in figure 4. The figure shows that in the area immediately surrounding a house (radius of 200 meters), the effect of flooded residential buildings is negative and significant at the 1 percent level with a magnitude of approximately 5 percent if the entire surrounding area is flooded. This effect is comparable to the related literature, where Currie et al. (2015) finds the negative effect of a polluting plant to be 11 percent in a radius of 800 meters, or Diamond and McQuade (2016) who show that low-income housing decreases value of nearby expensive homes by 2.5 percent. The most comparable study is that of Rossi-Hansberg et al. (2010), which also has changes in house appeal at the individual level, albeit it is improvements rather than damages like in this paper. Rossi-Hansberg et al. (2010) find the effect from a nearby repaired house to be approximately 10% of total value, which is almost twice as much as the 5 percent value found in this analysis. One reason for this is that in the natural experiment of subsidized repairs for low-income neighborhoods, residents may have expected the appeal of the repaired homes to remain in place for many years before deterioration would occur (unless the entire neighborhood would switch to a higher-valued equilibrium, in which case it would not). On the other hand, Houston did not experience significant departures of its residents after Harvey, and so the market may expect the disamenity from flooded housing to be only temporary. Secondly, the negative externality effect measured in this analysis could be partially weakened by the positive shortage effect.

[1]Figure 6 shows the results hold when controlling for risk-updating, including fixed effects and clustering at the neighborhood level.

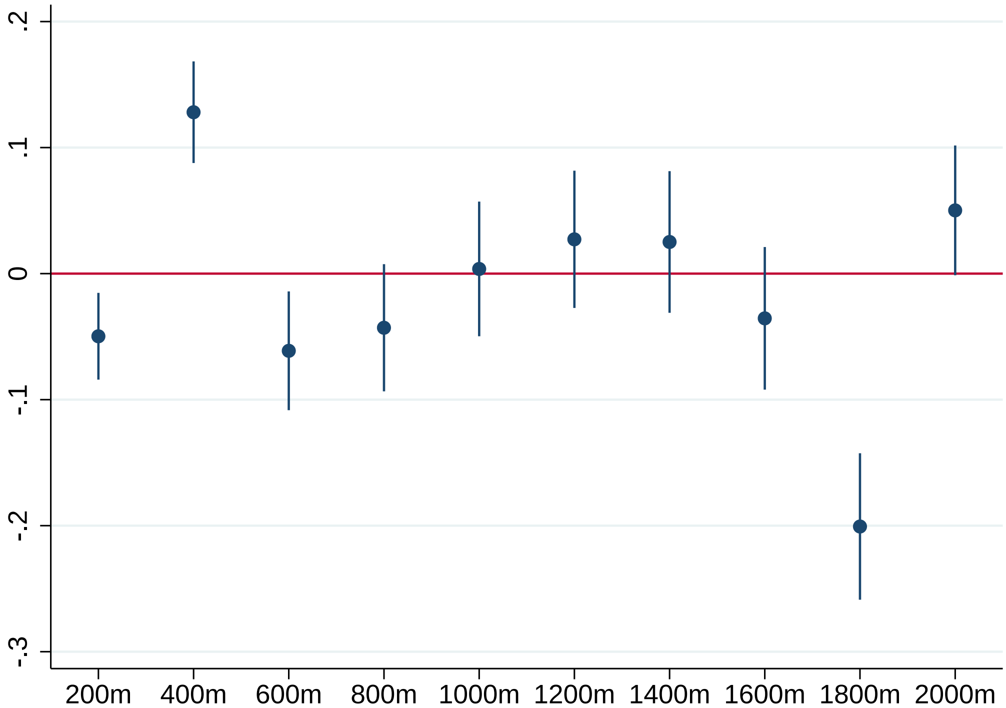

Figure 4: Flooding Nearby, Residential

The negative sign of the coefficient associated with flooded residential buildings in a 200-meter radius is expected due to the disamenity from having damaged, abandoned or under-construction housing surrounding one's dwelling. On the contrary, none of the studies I am aware of find a non-monotonic effect as is observed in the regression: if the additional area spanning from 200 meters to 400 meters from a house is flooded, this actually increases the value of one's non-flooded home by 13 percent in case the whole area is flooded. This effect happens due to the shortage effect and strongly depends on whether the neighborhood with flooded houses is appealing, so that residents are willing to pay a premium in order to stay living in it (the appeal could be due to a high-quality school district or a nearby job, for example). The effect for the area spanning 400 to 600 meters is negative and statistically significant with a magnitude of seven percent, while virtually all other effects further than 600 meters are statistically non-significant.[2]

In order to support the presence of the shortage effect by more evidence, I run the same

[2]The area between 1600 and 1800 meters is negative and statistically significant according to the analysis, even though it is not line with the rest of the results. I do not find an economic explanation for why this specific radius would have such a large negative effect.

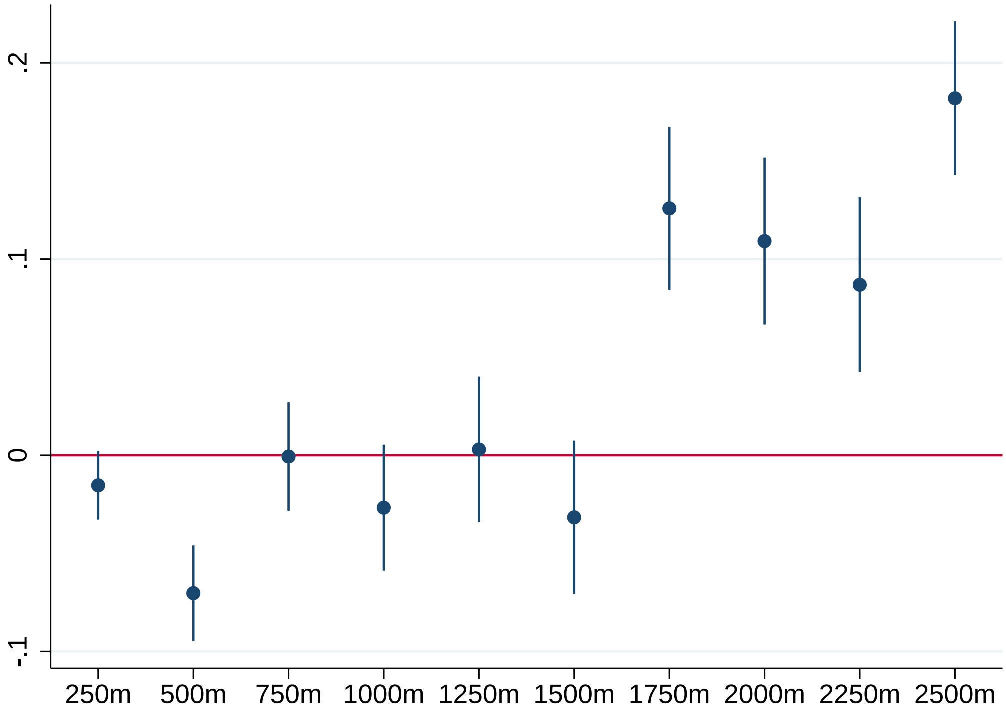

Figure 5: Flooding Nearby, Commercial

specification as in Equation 3 on two different subsamples: highly appealing and highly unappealing neighborhoods. If the hypothesis is correct, the shortage effects should be increasing in the level of neighborhood attractiveness. I use the ranking of attractiveness by Niche.com (2018), which uses 15 dimensions – including school quality, crime rate, employment and local amenities – to create a composite score of zip code appeal. I choose the 10 most and the 10 least appealing zip codes in Houston for 2018 and re-run the main regression on these subsamples. Figures 6 and 7 show the resulting coefficients for externalities from residential buildings. Even though the noise increases since the sample sizes are significantly smaller, the positive shortage effect is clearly stronger for the appealing neighborhoods with statistically significant positive values of 0.13 and 0.46 for the 400 and 600 meter radii, as opposed to 0.13 and -0.07 for the entire sample. On the other hand, the positive externality is non-existent for the unappealing neighborhoods for any radius, suggesting the amenity effect strongly dominates over the shortage effect (the coefficients for 400 and 600 meter radii are -0.15 and -0.23.

Figure 5 shows the results from Table 4 (using the entire sample of non-flooded homes)

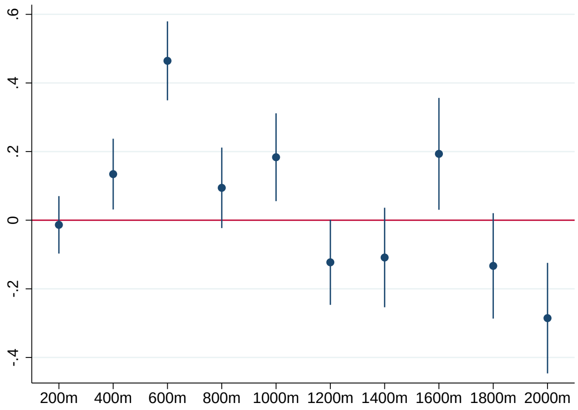

Figure 6: Flooding Nearby, Residential – Appealing Neighborhoods

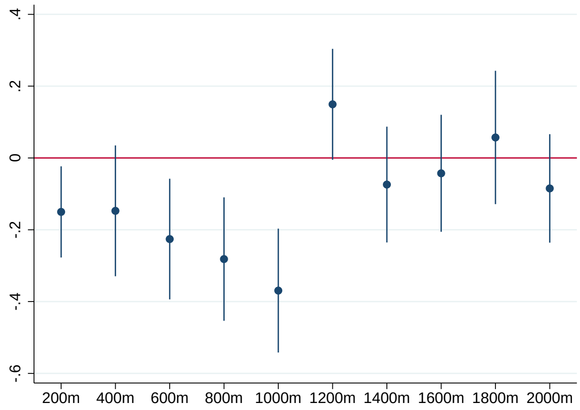

Figure 7: Flooding Nearby, Residential – Unappealing Neighborhoods

with regards to externalities produced by commercial buildings. Due to the much smaller incidence of commercial buildings, the radii additions are larger (250 meters), so that large amounts of data does not have to be discarded due to no commercial buildings in an area (which introduces division by zero in the proportion of flooded commercial buildings). The pattern for commercial buildings is very different from that of residential buildings – the effect for a 250 meter area is negative, but loses statistical significance when controlling for risk-updating. It has a magnitude of approximately 2 percent. The effect for the next area up to 500 meters is negative and statistically significant with a magnitude of 7 percent, so that if all commercial buildings in the region between 250 and 500 meters are flooded, one's house value declines by 7 percent. The dissipation of the negative effects happens already at the 750 meter level and remains indistinguishable from zero until the 1500 meter mark. Areas between 1500 and 2500 meters all show high positive and statistically significant effects, which I attribute to the shortage effect, since one likely does not visit commercial buildings which are almost two kilometers away by air distance, so that the distance when driving is much larger. The negative sign for flooded commercial areas close to one's home can be attributed to the loss of amenities such as nearby grocery stores, cafes or gyms. If these buildings receive flooding, it may take long time for them to return to the previous condition and there is a risk of their permanent closure. The negative effect only seems to last only until the 500 meter distance. While external effects from residential buildings dissipate at 800 meters, the disamenity effects of commercial buildings only last up to 500 meters. The disamenity effects of residential and commercial externalities are of comparable magnitudes, although the shortage effect is much larger for commercial buildings.

Comparing the "Standard" reduced-form regression with the one controlling for risk in Table 4, one can observe the magnitude of the risk-updating effect. When controlling for risk-updating, virtually all externality effects become larger, although the difference is small enough, so that it affect statistical significance only for the 250 meter area for commercial buildings. The increase in coefficients is expected, since the total negative effect was hypothesized to be a combination of the disamenity and the risk-updating effect. Both of these effects go in the same direction, since after Harvey, the market increased the probability of future flooding for most houses. As for the measures of risk themselves, the regression shows houses which are more elevated increase in value after Harvey as people value safe houses more. Specifically, a house which is elevated by 10 more meters then average is now worth 3 percent more. A similar effect is observed for distance to river, where an additional distance of 10 kilometers from a major river increases a home's value by 4 percent. Finally, the regression shows that despite their limitations, government flood zones carry some information as houses within them declined in value by 2.5 percent after the hurricane.

Table 5 shows regression results using both non-flooded and flooded houses as specified in Equation 4. The regression shows that a flooded house on average rose in value by 24 percent, which strongly suggests that majority of homeowners who received flooding invested heavily in mitigating damages as well as improving the house at the same time. Since information on damages or home investment is not available, the results presented here might not be robust. When including flooded homes in the sample, the results for non-flooded homes do not significantly change. Table 5 shows that distance to a major river has a different effect on flooded homes than is the effect on non-flooded ones. Flooded homes further away from a river decline in value, although the effect is relatively small at 1 percent for additional 10 kilometers. The reason is likely that houses far away from the river had been considered safe from flooding before Harvey, which is why they carried a safety premium, which is now lost. The coefficients for elevation and flood zone designation have the same sign for all houses, although the magnitude is slightly larger for flooded ones. The pattern of externality effects on flooded houses does not seem to be highly informative, as the effects are non-monotonic and frequently change sign, making it difficult to differentiate between the disamenity and the shortage effects.

7 Robustness Checks

7.1 Parallel Trends for Flooded and Non-flooded Homes

Figure 8 graphically shows that homes flooded during hurricane Harvey are, on average, approximately 30 percent more expensive than non-flooded homes. However, both categories of homes follow virtually identical growth trend before Harvey: very little growth between 2010 and 2012, strong growth between 2012 and 2015 and a slowdown in in market value appreciation between 2015 and 2017. This is additional evidence that the externality effects measured in this analysis are not due to systematic differences between flooded and non-flooded homes before hurricane Harvey.

7.2 Heterogeneous Effects by Internal House Characteristics

There is a possibility that externality effects affect houses differently depending on the house characteristics or type. For example, large houses may suffer from flooding nearby more than small ones. In this robustness check, I interact the flooding externalities with internal house characteristics, specifically house square footage and land square footage. I find that when including these interactions, they are virtually all statistically insignificant and small as can be seen in Tables 9 and 10 in the Appendix. When I perform an F-test of the join significance of all the interaction terms, I also find they are jointly non-significant. The non-interacted flooding externalities (as defined in Equation 3) remain robust to including these interactions.

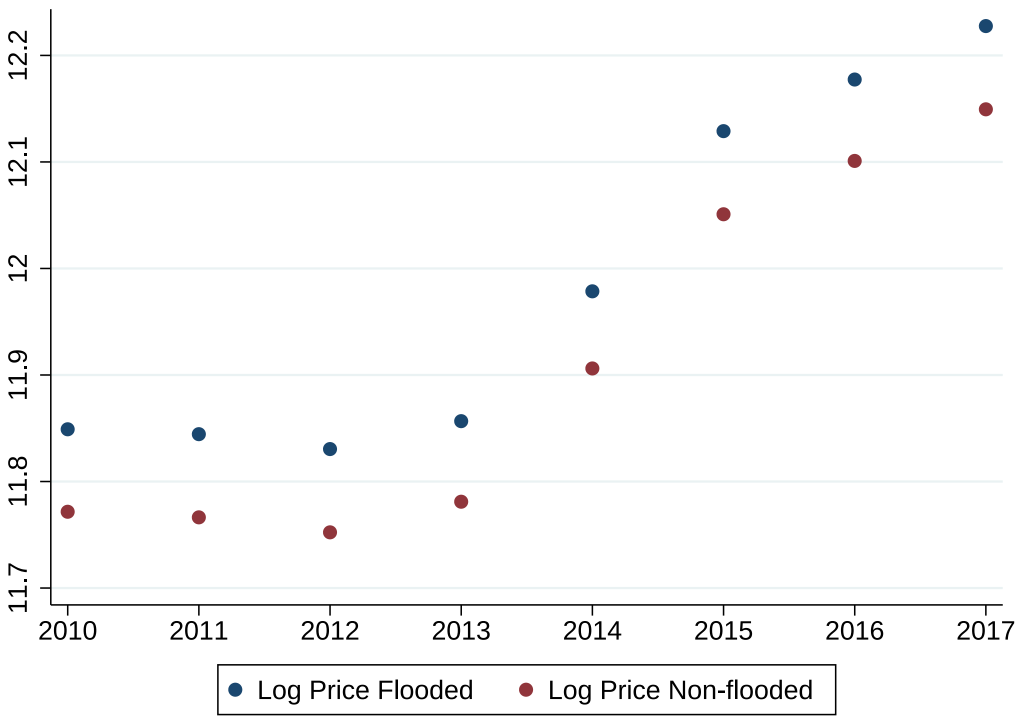

Figure 8: Parallel Trends in Prices Before Harvey

8 Conclusion

In this analysis, I find that housing externalities are large, and they arise from both residential and commercial buildings. The disamenity effect from nearby damaged buildings ranges between negative 5 and negative 7 percent, which is comparable to other findings in the related literature. Moreover, I find large shortage effects associated with both types of buildings, showing evidence they are driven by residents' strong preference for some appealing neighborhoods. I find the shortage effects to be of magnitudes around 12 percent. Finally, the analysis shows the disamenity effects from commercial buildings dissipate at 500 meters, while they stay present until 800 meteres for residential buildings. On the other hand, the shortage effect for commercial buildings does not occur even up to 2,500 meters. Understanding these externalities is important in evaluating any urban policy, such as subsidizing home repairs, building a new park, or understanding how home closures (e.g. Detroit) affect house prices. This question has economic significance, as most individuals hold majority of savings in their property, and spend a large portion of their income to finance their housing via a mortgage. This analysis also shows that even residents whose houses are not flooded may receive significant financial damages due to the negative externalities from flooded buildings. Finally, residents whose houses receive flooding and prefer to stay in the same neighborhood may suffer two-fold, since buying or renting a house there may be significantly more expensive due to the shortage effects.

This paper contributes to the related literature in several ways. As discussed in section 3, the variation caused by flooded buildings is easier to defend in terms of exogeneity compared to many of the previously mentioned natural experiments. For example, Rossi-Hansberg et al. (2010) exploit the government subsidy program for house repairs. However, the program targeted especially poor neighborhoods and possibly especially damaged houses with the largest possible gains from investment, calling into question the external validity of the results. Secondly, to my knowledge, no other paper has looked at the housing externalities with relation to an external change in commercial buildings, which I find to be significant and large. Finally, this analysis could be interesting simply as a case study of a large US city (Houston is the 4th largest agglomeration by population). In most of the literature, repeated sales models are used, so that the entire universe of transactions over many years has to be purchased to partially alleviate the sample selection major problem. Given the data's large price tag (it is ordinarily bought and used in business by real estate brokers), this constrained previous analysis to small cities, such as Richmond with 200,000 people used by Rossi-Hansberg et al. (2010) or a county in North Carolina with only several thousand people used by Hallstrom and Smith (2005).

To extend this analysis, it may be worthwhile to consider the effect of schools damaged by flooding on house market values, since school quality is a strong determinant of neighborhood appeal. Secondly, the heterogeneity in housing externalities may be large, as was suggested by the differential effects of flooded buildings for appealing and unappealing neighborhoods, or in the paper by Diamond and McQuade (2016), where the externality of newly built affordable housing is positive for low-income and negative for high-income neighborhoods. Given the relatively large sample size, this heterogeneity could be described in a more systematic manner. Finally, a caveat of the paper is the measurement of the net effect of the disamenity and the shortage effect – ideally, one could find an instrument to disentangle these two.

This paper goes in line in the relatively new literature describing housing externalities and supports the notion that they must be taken into account when considering the cost-benefit analysis of any urban design policy.

References

Anselin, L. and Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. *Statistics Textbooks and Monographs*, 155:237–290.

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

Core Logic (2017). Corelogic analysis estimates flood loss for hurricane harvey. *http://www.businesswire.com/news/home/20170831006476/en/*.

Coulomb, R. and Zylberberg, Y. (2016). Rare events and risk perception: evidence from fukushima accident.

Currie, J., Davis, L., Greenstone, M., and Walker, R. (2015). Environmental health risks and housing values: evidence from 1,600 toxic plant openings and closings. *American Economic Review*, 105(2):678–709.

Diamond, R. and McQuade, T. (2016). Who wants affordable housing in their backyard? an equilibrium analysis of low income property development. *National Bureau of Economic Research*.

Federal Emergency Management Agency (2017a). Fema flood map service center. *https://msc.fema.gov/portal/*.

Federal Emergency Management Agency (2017b). Individual assistance open disaster statistics. *https://www.fema.gov/media-library/assets/documents/132213*.

FEMA (2017). Hurricane harvey coastal modeled. *https://data.femadata.com/NationalDisasters/HurricaneHarvey/Data/DepthGrid/FEMA/*.

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

Harris County Flood Control District (2017). Harris county flood warning system. *https://www.harriscountyfws.org/*.

Houston-Galveston Area Council (2018). Gis datasets. *http://www.h-gac.com/rds/gis-data/gis-datasets.aspx*.

Kinder Institute (2017). Comparing harvey to houston's past storms.

Leggett, C. G. and Bockstael, N. E. (2000). Evidence of the effects of water quality on residential land prices. *Journal of Environmental Economics and Management*, 39(2):121–144.

McCluskey, J. J. and Rausser, G. C. (2001). Estimation of perceived risk and its effect on property values. *Land Economics*, 77(1):42–55.

Muir, C. (2018). Zillow scraping with python. *https://github.com/ChrisMuir/Zillow*.

National Oceanic and Atmospheric Administration (2017). Hurricane harvey imagery. *https://storms.ngs.noaa.gov/storms/harvey/index.html19/29.77061/-95.62298*.

Niche.com (2018). 2018 best zip codes to live in the houston area. *https://www.niche.com/places-to-live/search/best-zip-codes-to-live/m/houston-metro-area/*.

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, 82(1):34–55.

Rossi-Hansberg, E., Sarte, P.-D., and Owens III, R. (2010). Housing externalities. *Journal of political Economy*, 118(3):485–535.

US Census Bureau (2017). Median value of assets for households, by type of asset owned and selected characteristics: 2013. *Survey of Income and Program Participation, 2014 Panel, Wave 1*.

USACE (2017). Flooding model hurricane harvey.

Vigdor, J. (2008). The economic aftermath of hurricane katrina. *Journal of Economic Perspectives*, 22(4):135–54.

Appendix

Table 4: Flooding Externalities, Non-flooded Houses Only

	Standard		Control for Risk		Control for Risk and FE	
Prop Flooded Res 200m	-0.053**	(0.018)	-0.042*	(0.018)	-0.040*	(0.018)
Prop Flooded Res 400m	0.128***	(0.021)	0.130***	(0.021)	0.126***	(0.020)
Prop Flooded Res 600m	-0.068**	(0.024)	-0.062**	(0.024)	-0.069**	(0.024)
Prop Flooded Res 800m	-0.046	(0.026)	-0.044	(0.026)	-0.042	(0.026)
Prop Flooded Res 1000m	-0.010	(0.027)	0.003	(0.027)	0.004	(0.027)
Prop Flooded Res 1200m	0.019	(0.028)	0.025	(0.028)	0.029	(0.028)
Prop Flooded Res 1400m	0.023	(0.029)	0.029	(0.029)	0.028	(0.029)
Prop Flooded Res 1600m	-0.043	(0.029)	-0.037	(0.029)	-0.030	(0.029)
Prop Flooded Res 1800m	-0.210***	(0.030)	-0.203***	(0.030)	-0.197***	(0.029)
Prop Flooded Res 2000m	0.026	(0.026)	0.049	(0.026)	0.049	(0.026)
Prop Flooded Com 250m	-0.019*	(0.009)	-0.016	(0.009)	-0.018*	(0.009)
Prop Flooded Com 500m	-0.070***	(0.012)	-0.069***	(0.012)	-0.073***	(0.012)
Prop Flooded Com 750m	-0.003	(0.014)	0.001	(0.014)	0.001	(0.014)
Prop Flooded Com 1000m	-0.026	(0.016)	-0.024	(0.016)	-0.033*	(0.016)
Prop Flooded Com 1250m	0.004	(0.019)	0.004	(0.019)	-0.007	(0.019)
Prop Flooded Com 1500m	-0.037	(0.020)	-0.033	(0.020)	-0.028	(0.020)
Prop Flooded Com 1750m	0.114***	(0.021)	0.128***	(0.021)	0.128***	(0.021)
Prop Flooded Com 2000m	0.093***	(0.022)	0.111***	(0.022)	0.100***	(0.022)
Prop Flooded Com 2250m	0.074**	(0.023)	0.088***	(0.023)	0.081***	(0.023)
Prop Flooded Com 2500m	0.147***	(0.020)	0.176***	(0.020)	0.163***	(0.020)
Elevation			0.003***	(0.000)	0.002***	(0.000)
Distance to River (km)			0.004***	(0.001)	0.008***	(0.001)
Flood Zone			-0.025***	(0.006)	-0.023***	(0.006)
Observations	74560		74560		74549	
F	23.947		40.784		46.465	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

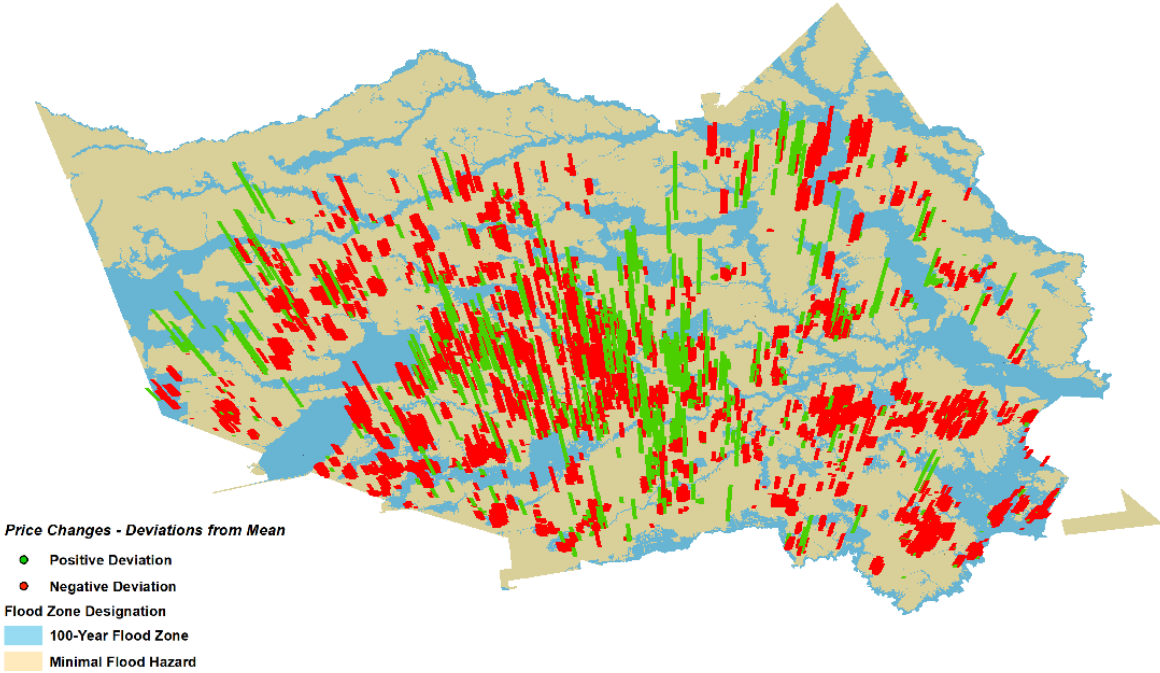

Figure 9: Price Deviations from Mean

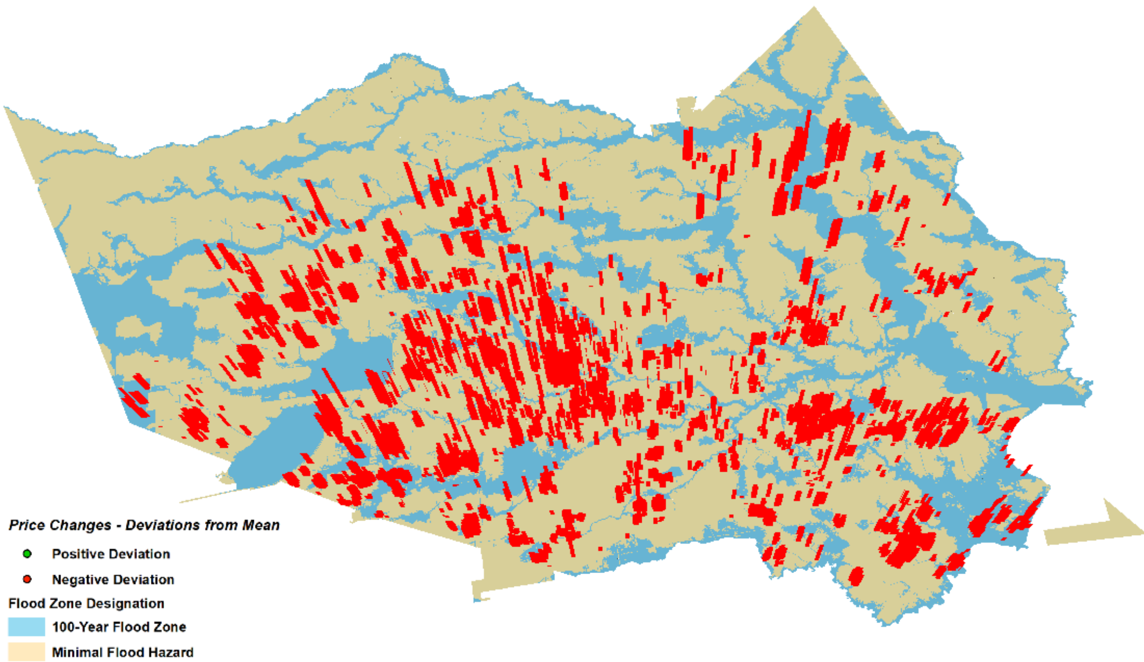

Figure 10: Price Deviations from Mean (Negative Only)

Table 5: Flooding Externalities, All Houses

		Standard		Control for Risk		Control for Risk and FE	
Flooded		0.242***	(0.072)	0.280***	(0.075)	0.287***	(0.074)
Flooded=0	Prop Flooded Res 200m	-0.053**	(0.018)	-0.042*	(0.018)	-0.041*	(0.018)
Flooded=1	Prop Flooded Res 200m	0.154***	(0.023)	0.159***	(0.023)	0.153***	(0.023)
Flooded=0	Prop Flooded Res 400m	0.128***	(0.021)	0.130***	(0.021)	0.127***	(0.021)
Flooded=1	Prop Flooded Res 400m	-0.249***	(0.034)	-0.248***	(0.034)	-0.249***	(0.033)
Flooded=0	Prop Flooded Res 600m	-0.068**	(0.024)	-0.062*	(0.024)	-0.068**	(0.024)
Flooded=1	Prop Flooded Res 600m	-0.120**	(0.042)	-0.121**	(0.042)	-0.126**	(0.042)
Flooded=0	Prop Flooded Res 800m	-0.046	(0.026)	-0.044	(0.026)	-0.041	(0.026)
Flooded=1	Prop Flooded Res 800m	0.860***	(0.047)	0.850***	(0.047)	0.848***	(0.046)
Flooded=0	Prop Flooded Res 1000m	-0.010	(0.027)	0.003	(0.027)	0.002	(0.027)
Flooded=1	Prop Flooded Res 1000m	-0.553***	(0.053)	-0.584***	(0.053)	-0.581***	(0.053)
Flooded=0	Prop Flooded Res 1200m	0.019	(0.028)	0.025	(0.028)	0.030	(0.028)
Flooded=1	Prop Flooded Res 1200m	0.109	(0.056)	0.084	(0.056)	0.087	(0.056)
Flooded=0	Prop Flooded Res 1400m	0.023	(0.029)	0.029	(0.029)	0.026	(0.029)
Flooded=1	Prop Flooded Res 1400m	-0.327***	(0.069)	-0.327***	(0.069)	-0.326***	(0.068)
Flooded=0	Prop Flooded Res 1600m	-0.043	(0.029)	-0.037	(0.029)	-0.030	(0.029)
Flooded=1	Prop Flooded Res 1600m	0.190**	(0.072)	0.188**	(0.071)	0.193**	(0.071)
Flooded=0	Prop Flooded Res 1800m	-0.210***	(0.030)	-0.203***	(0.030)	-0.197***	(0.030)
Flooded=1	Prop Flooded Res 1800m	-0.183*	(0.081)	-0.187*	(0.080)	-0.188*	(0.080)
Flooded=0	Prop Flooded Res 2000m	0.026	(0.026)	0.049	(0.026)	0.048	(0.026)
Flooded=1	Prop Flooded Res 2000m	-0.297***	(0.078)	-0.282***	(0.078)	-0.278***	(0.078)
Flooded=0	Prop Flooded Com 250m	-0.019*	(0.009)	-0.016	(0.009)	-0.018*	(0.009)
Flooded=1	Prop Flooded Com 250m	0.125***	(0.016)	0.132***	(0.016)	0.132***	(0.016)
Flooded=0	Prop Flooded Com 500m	-0.070***	(0.012)	-0.069***	(0.012)	-0.074***	(0.012)
Flooded=1	Prop Flooded Com 500m	0.075**	(0.027)	0.081**	(0.027)	0.083**	(0.027)
Flooded=0	Prop Flooded Com 750m	-0.003	(0.014)	0.001	(0.014)	0.003	(0.014)
Flooded=1	Prop Flooded Com 750m	-0.097**	(0.034)	-0.070*	(0.034)	-0.072*	(0.034)
Flooded=0	Prop Flooded Com 1000m	-0.026	(0.016)	-0.024	(0.016)	-0.031	(0.016)
Flooded=1	Prop Flooded Com 1000m	-0.098*	(0.040)	-0.087*	(0.040)	-0.085*	(0.040)
Flooded=0	Prop Flooded Com 1250m	0.004	(0.019)	0.004	(0.019)	-0.006	(0.019)
Flooded=1	Prop Flooded Com 1250m	0.136**	(0.045)	0.154***	(0.046)	0.150***	(0.045)
Flooded=0	Prop Flooded Com 1500m	-0.037	(0.020)	-0.033	(0.020)	-0.027	(0.020)
Flooded=1	Prop Flooded Com 1500m	0.029	(0.051)	0.019	(0.051)	0.019	(0.051)
Flooded=0	Prop Flooded Com 1750m	0.114***	(0.021)	0.128***	(0.021)	0.128***	(0.021)
Flooded=1	Prop Flooded Com 1750m	-0.051	(0.055)	-0.047	(0.055)	-0.053	(0.055)
Flooded=0	Prop Flooded Com 2000m	0.093***	(0.022)	0.111***	(0.022)	0.100***	(0.022)
Flooded=1	Prop Flooded Com 2000m	0.550***	(0.059)	0.549***	(0.059)	0.545***	(0.058)
Flooded=0	Prop Flooded Com 2250m	0.074**	(0.023)	0.088***	(0.023)	0.080***	(0.023)
Flooded=1	Prop Flooded Com 2250m	0.051	(0.063)	0.039	(0.063)	0.028	(0.063)
Flooded=0	Prop Flooded Com 2500m	0.147***	(0.020)	0.176***	(0.020)	0.164***	(0.020)
Flooded=1	Prop Flooded Com 2500m	-0.312***	(0.058)	-0.271***	(0.059)	-0.282***	(0.059)
Flooded=0	Elevation			0.003***	(0.000)	0.003***	(0.000)
Flooded=1	Elevation			0.004***	(0.001)	0.003***	(0.001)
Flooded=0	Distance to River (km)			0.004***	(0.001)	0.007***	(0.001)
Flooded=1	Distance to River (km)			-0.012***	(0.004)	-0.008*	(0.004)
Flooded = 0 Flood Zone				-0.025***	(0.006)	-0.023***	(0.006)
Flooded = 1 Flood Zone				-0.059***	(0.012)	-0.054***	(0.012)
Observations		82102		82102		82091	
F		32.619		39.274		42.763	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Flooding Externalities, Cluster at Neighborhood Level

	Standard		Control for Risk		Control for Risk and FE	
Prop Flooded Res 200m	-0.053***	(0.007)	-0.041***	(0.008)	-0.040***	(0.005)
Prop Flooded Res 400m	0.128***	(0.004)	0.130***	(0.003)	0.126***	(0.004)
Prop Flooded Res 600m	-0.068***	(0.006)	-0.062***	(0.008)	-0.069***	(0.005)
Prop Flooded Res 800m	-0.046***	(0.005)	-0.044***	(0.006)	-0.042***	(0.006)
Prop Flooded Res 1000m	-0.010*	(0.004)	0.003	(0.005)	0.004	(0.005)
Prop Flooded Res 1200m	0.019	(0.013)	0.025**	(0.008)	0.029***	(0.002)
Prop Flooded Res 1400m	0.024**	(0.007)	0.030*	(0.012)	0.028*	(0.012)
Prop Flooded Res 1600m	-0.044	(0.024)	-0.038*	(0.018)	-0.030**	(0.009)
Prop Flooded Res 1800m	-0.210***	(0.007)	-0.203***	(0.008)	-0.197***	(0.014)
Prop Flooded Res 2000m	0.026	(0.021)	0.049*	(0.019)	0.049**	(0.013)
Prop Flooded Com 250m	-0.019***	(0.005)	-0.016**	(0.005)	-0.018***	(0.002)
Prop Flooded Com 500m	-0.070***	(0.006)	-0.069***	(0.003)	-0.073***	(0.003)
Prop Flooded Com 750m	-0.003	(0.013)	0.001	(0.008)	0.001	(0.004)
Prop Flooded Com 1000m	-0.025	(0.017)	-0.023	(0.013)	-0.033***	(0.004)
Prop Flooded Com 1250m	0.004	(0.020)	0.004	(0.016)	-0.007	(0.006)
Prop Flooded Com 1500m	-0.038*	(0.014)	-0.033***	(0.007)	-0.028*	(0.013)
Prop Flooded Com 1750m	0.114***	(0.004)	0.129***	(0.007)	0.128***	(0.008)
Prop Flooded Com 2000m	0.093***	(0.005)	0.111***	(0.005)	0.100***	(0.012)
Prop Flooded Com 2250m	0.075***	(0.012)	0.089***	(0.009)	0.081***	(0.007)
Prop Flooded Com 2500m	0.147***	(0.029)	0.176***	(0.025)	0.163***	(0.015)
Elevation			0.003**	(0.001)	0.002***	(0.000)
Distance to River (km)			0.004	(0.004)	0.008*	(0.004)
Flood Zone			-0.025	(0.013)	-0.023	(0.015)
Observations	74549		74549		74549	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Summary Statistics Zillow Before

	Mean	Standard Deviation	Minimum	Maximum
Square Footage House	2128.42	(949.79)	283.00	15000.00
Price Zillow	12.12	(0.75)	10.13	15.52
Number of Bedrooms	3.29	(0.87)	0.00	16.00
Number of Bathrooms	2.42	(0.90)	1.00	16.00
Elevation	22.53	(11.53)	0.08	61.60
Distance to River (km)	2.09	(1.51)	0.00	8.48
Flooded	0.08	(0.27)	0.00	1.00
Observations	15049			

Table 8: Summary Statistics Zillow After

	Mean	Standard Deviation	Minimum	Maximum
Square Footage House	2109.04	(891.84)	256.00	7838.00
Price Zillow	12.10	(0.75)	10.13	14.69
Number of Bedrooms	3.31	(0.84)	0.00	6.00
Number of Bathrooms	2.49	(0.96)	1.00	7.00
Elevation	23.72	(12.25)	1.74	59.71
Distance to River (km)	2.06	(1.52)	0.01	8.39
Flooded	0.08	(0.27)	0.00	1.00
Observations	1234			

Table 9: Flooding Externalities with House Square Footage Interactions, Non-flooded Houses Only

Res 200m	-0.048**	(0.018)
Res 400m	0.139***	(0.021)
Res 600m	-0.069**	(0.025)
Res 800m	-0.039	(0.026)
Res 1000m	0.001	(0.028)
Res 1200m	0.022	(0.028)
Res 1400m	0.029	(0.029)
Res 1600m	-0.036	(0.029)
Res 1800m	-0.211***	(0.030)
Res 2000m	0.051	(0.027)
Com 250m	-0.015	(0.009)
Com 500m	-0.071***	(0.013)
Com 750m	0.001	(0.014)
Com 1000m	-0.024	(0.017)
Com 1250m	0.003	(0.019)
Com 1500m	-0.034	(0.020)
Com 1750m	0.123***	(0.021)
Com 2000m	0.119***	(0.022)
Com 2250m	0.089***	(0.023)
Com 2500m	0.181***	(0.020)
Square Footage House Res 200m	0.000	(0.000)
Square Footage House Res 400m	-0.000	(0.000)
Square Footage House Res 600m	0.000	(0.000)
Square Footage House Res 800m	-0.000	(0.000)
Square Footage House Res 1000m	0.000	(0.000)
Square Footage House Res 1200m	0.000	(0.000)
Square Footage House Res 1400m	-0.000	(0.000)
Square Footage House Res 1600m	-0.000	(0.000)
Square Footage House Res 1800m	0.000	(0.000)
Square Footage House Res 2000m	-0.000	(0.000)
Square Footage House Com 250m	-0.000	(0.000)
Square Footage House Com 500m	0.000	(0.000)
Square Footage House Com 750m	0.000	(0.000)
Square Footage House Com 1000m	-0.000	(0.000)
Square Footage House Com 1250m	0.000	(0.000)
Square Footage House Com 1500m	0.000	(0.000)
Square Footage House Com 1750m	0.000	(0.000)
Square Footage House Com 2000m	-0.000*	(0.000)
Square Footage House Com 2250m	-0.000	(0.000)
Square Footage House Com 2500m	-0.000	(0.000)
Elevation	0.003***	(0.000)
Distance to River (km)	0.004***	(0.001)
Flood Zone	-0.025***	(0.006)
Observations	74560	
F	23.336	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 10: Flooding Externalities with Land Square Footage Interactions, Non-flooded Houses Only

Res 200m	-0.050**	(0.018)
Res 400m	0.143***	(0.021)
Res 600m	-0.072**	(0.025)
Res 800m	-0.035	(0.027)
Res 1000m	-0.004	(0.028)
Res 1200m	0.027	(0.029)
Res 1400m	0.024	(0.029)
Res 1600m	-0.035	(0.030)
Res 1800m	-0.215***	(0.031)
Res 2000m	0.051	(0.027)
Com 250m	-0.017	(0.009)
Com 500m	-0.071***	(0.013)
Com 750m	0.000	(0.014)
Com 1000m	-0.022	(0.017)
Com 1250m	0.003	(0.019)
Com 1500m	-0.032	(0.021)
Com 1750m	0.127***	(0.022)
Com 2000m	0.119***	(0.022)
Com 2250m	0.094***	(0.023)
Com 2500m	0.182***	(0.021)
Square Footage Land Res 200m	0.000	(0.000)
Square Footage Land Res 400m	-0.000*	(0.000)
Square Footage Land Res 600m	0.000	(0.000)
Square Footage Land Res 800m	-0.000	(0.000)
Square Footage Land Res 1000m	0.000	(0.000)
Square Footage Land Res 1200m	0.000	(0.000)
Square Footage Land Res 1400m	0.000	(0.000)
Square Footage Land Res 1600m	-0.000	(0.000)
Square Footage Land Res 1800m	0.000	(0.000)
Square Footage Land Res 2000m	-0.000	(0.000)
Square Footage Land Com 250m	0.000	(0.000)
Square Footage Land Com 500m	0.000	(0.000)
Square Footage Land Com 750m	0.000	(0.000)
Square Footage Land Com 1000m	-0.000	(0.000)
Square Footage Land Com 1250m	0.000	(0.000)
Square Footage Land Com 1500m	-0.000	(0.000)
Square Footage Land Com 1750m	0.000	(0.000)
Square Footage Land Com 2000m	-0.000	(0.000)
Square Footage Land Com 2250m	-0.000	(0.000)
Square Footage Land Com 2500m	-0.000	(0.000)
Elevation	0.003***	(0.000)
Distance to River (km)	0.004***	(0.001)
Flood Zone	-0.025***	(0.006)
Observations	74560	
F	23.306	

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$