Essays on Beliefs about Catastrophic Risks

by

Timothy Hyde

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics

Tepper School of Business

Carnegie Mellon University

Dissertation committee:
Nicholas Muller, CMU
Saurabh Bhargava, CMU
Dennis Epple, CMU
Lala Ma, University of Kentucky

Acknowledgements

This dissertation would not have been possible without incredible support from people in the CMU community and beyond. First and foremost, immense thanks to Nick Muller and Saurabh Bhargava for their guidance, advice, and encouragement. Nick's ability to kick-start the brainstorming process helped bring so many of my half-formed ideas to fruition, and his relentless optimism about my work helped me persevere at the lowest points of my graduate study. Through Saurabh's close and attentive mentoring, I learned so much about the research process, and his willingness to have me collaborate on his projects provided invaluable hands-on experience. I also thank Carrie Hopkins, who coauthored my third chapter and whose work inspired many of the ideas in this dissertation. Thanks also to Dennis Epple and Lala Ma for their service on my dissertation committee, and their encouragement, guidance, R code recommendations, and constructive criticism.

I also acknowledge invaluable assistance, constructive comments, data facilitation, and suggestions from Jay Apt, Linda Babcock, James Best, Justin Gallagher, Kareem Haggag, Tim Houlihan, Alex Imas, Karam Kang, David Kelly, Jennifer Kirby, Rebecca Lessem, Yanjun Liao, George Loewenstein, Stephen Spear, Lowell Taylor, Margaret Walls, Randall Walsh, and many more. In addition, Stephanie Rifai, Cassandra Taylor, and Sana Lakdawala provided excellent research assistance and data management.

Finally, thanks to my partner Micol and my son Milo, who have shown incredible patience and grace during this journey.

for Milo: may you enjoy a lifetime of learning

Contents

1 How do homebuyers adapt after experiencing a natural disaster? 1

1.1 Introduction . . . 1
1.2 Literature review . . . 2
1.2.1 Belief updating after a disaster shock . . . 2
1.2.2 Disaster shocks in coastal real estate markets . . . 3
1.2.3 Climate adaptation in real estate . . . 3
1.3 Data . . . 4
1.3.1 ZTRAX database . . . 4
1.3.2 Determining buyer origin . . . 8
1.3.3 Hydrological data . . . 9
1.3.4 Flood event data . . . 10
1.3.5 Disaster pressure index . . . 12
1.4 Retreat from the water: methodology and results . . . 13
1.4.1 Model specification . . . 14
1.4.2 Results from count models . . . 15
1.4.3 Falsification tests . . . 21
1.5 Sorting model and welfare analysis . . . 21
1.5.1 Model construction . . . 21
1.5.2 Instrumental variables approach . . . 23
1.5.3 Genetic matching procedure . . . 24
1.5.4 Model results . . . 25
1.5.5 Welfare analysis . . . 26
1.6 Amenity erosion: methodology and results . . . 28
1.6.1 Model specification . . . 30
1.6.2 Results from hedonic models . . . 30
1.6.3 Property value erosion calculation . . . 34
1.7 Conclusion . . . 36

2 Risk aversion in the field: evidence from an employee rewards program 37

2.1 Introduction . . . 37
2.2 Background . . . 41
2.2.1 Institutional background . . . 41
2.2.2 Reward program overview . . . 41
2.2.3 Goal and reward structure . . . 42
2.3 Theoretical framework for goal choice . . . 43
2.3.1 Generalized expected utility framework . . . 43
2.3.2 Conservatism with EU and rational expectations ((s) = s, v(x_n, .) = u(x_n)) . . . 43
2.3.3 Conservatism due to non-standard beliefs ($\hat{s}_n$ 6= E(s_nj)) . . . 44

2.3.4 Conservatism due to non-standard decision weights ((s) 6 s) . . . 44
2.3.5 Conservatism due to loss aversion (u(.) = u(x_n,)) . . . 45
2.4 Data and sample construction . . . 45
2.4.1 Primary employee sample . . . 45
2.4.2 Goal choice and employee productivity . . . 46
2.4.3 Employee beliefs . . . 48
2.5 Characterization of goal choice by benchmark model . . . 50
2.5.1 Expected utility with risk neutrality . . . 50
2.5.2 Expected utility with risk aversion . . . 53
2.5.3 Non-linear decision weights . . . 54
2.5.4 Gain-loss utility . . . 56
2.6 Goal choice in experimental paradigm . . . 59
2.6.1 Research design and protocol . . . 59
2.6.2 Comparison of goal choice in the lab and the field . . . 60
2.6.3 Characterization of choice . . . 62
2.7 Conclusion . . . 63

3 Do risk perceptions change after a disaster? Evidence from flood-prone counties 65
3.1 Introduction . . . 65
3.2 Background . . . 66
3.3 Survey and data sources . . . 67
3.3.1 Survey construction . . . 67
3.3.2 Additional data sources . . . 69
3.3.3 Map exposure experimental design . . . 69
3.3.4 Sample composition . . . 70
3.3.5 Drawbacks of the internet survey approach . . . 72
3.4 Approach and hypothesis . . . 73
3.4.1 Cross sectional analysis . . . 73
3.4.2 Map exposure treatment . . . 73
3.4.3 Hurricane exposure treatment . . . 74
3.4.4 Threats to causal inference . . . 76
3.5 Results . . . 76
3.5.1 Cross sectional analysis . . . 76
3.5.2 Map exposure treatment . . . 77
3.5.3 Hurricane exposure treatment . . . 82
3.5.4 Robustness checks . . . 82
3.6 Conclusion . . . 85

Appendix materials 95
A1 Observations excluded from analysis . . . 95
A2 Same-address listing behavior . . . 96
A3 Identifying ZIP codes with physical flood damage . . . 96
A4 Distinctive origin states . . . 97
A5 Imputation of expected performance for online GoalQuest participants . . . 97
A6 Survey instrument . . . 97
A7 Appendix figures and tables . . . 101

List of Figures

1.1 Map of 818 main Florida origin counties . . . 7
1.2 Tree map of "buyer mail" variable in Sarasota County transactions . . . 9
1.3 Hydrological characteristics of coastal properties in Naples, Florida . . . 10
1.4 Flood hazard by property distance to water . . . 11
1.5 The disaster pressure index in four MSAs over time . . . 13
1.6 Retreat effect among 6-12 month treatment group: distance from water . . . 18
1.7 Retreat effect among 6-12 month treatment group: flood zones . . . 18
1.8 Transience of the retreat effect in the 6-ft sea-level-rise zone . . . 19
1.9 Retreat from the water under different disaster definitions . . . 20
1.10 Results of genetic matching on data from Tampa MSA . . . 25
1.11 Most distinctive origin state for each Florida destination county . . . 29
1.12 Change in amenity value induced by rising disaster pressure . . . 32
1.13 Hedonic meta-analysis of flood zone disamenity . . . 33
1.14 Coastal amenity erosion across different submarkets . . . 35

2.1 GoalQuest goal selection webflow . . . 38
2.2 Cumulative distribution of unrealized rewards: ex-post attainment . . . 52
2.3 Distribution of rational expectations and subj. beliefs of goal attainment . . . 53
2.4 Optimal Choice under EU by risk preference and information regime . . . 55

3.1 Location of 1,582 survey responses . . . 68
3.2 Hurricane recollections . . . 75
3.3 Effect of exposure to FEMA flood maps . . . 79
3.4 Calibration of flood risk perceptions to FEMA flood maps . . . 79
3.5 Natural experiment: hurricane exposure . . . 83
A1 FEMA flood map for Miami, Florida . . . 101
A2 Illustration of treatment definition . . . 103
A3 QQ plots for Tampa MSA . . . 104
A4 Principal component analysis of Florida MSA catchment areas . . . 105
A5 Example addition task from lab paradigm . . . 106
A6 Relationship between disaster pressure and same-address listing . . . 109

List of Tables

1.1 Descriptive statistics . . . 5
1.2 Transactions by metropolitan statistical area (MSA), origin data only . . . 6
1.3 U.S. flood disaster frequency, 2005-2017 . . . 12
1.4 Sales count models . . . 16
1.5 Sales count model robustness checks . . . 17
1.6 Falsification tests for count models . . . 20
1.7 Sorting model estimates . . . 27
1.8 Statewide hedonic price models . . . 31
1.9 Amenity erosion with influx of shocked buyers . . . 35

2.1 Summary of sample, group, and employee characteristics . . . 47
2.2 Goal choice, employee productivity, and goal attainment . . . 48
2.3 Employee beliefs and confidence of goal attainment . . . 49
2.4 Goal choice characterization for expected utility benchmarks . . . 51
2.5 Goal choice characterization for non-standard benchmarks . . . 57
2.6 Descriptive accuracy of gain-loss utility models by candidate reference point . . . 58
2.7 GoalQuest-style menus offered in lab paradigm . . . 59
2.8 Goal choice, beliefs, and goal attainment: online paradigm . . . 60
2.9 Goal choice characterization for non-standard benchmarks: online paradigm . . . 61

3.1 Sample composition . . . 71
3.2 Distribution of key response variables . . . 71
3.3 Problematic observations in survey sample . . . 72
3.4 Recent storm exposure and response variables . . . 77
3.5 Effect of flood map exposure . . . 80
3.6 Interaction between map exposure and hydrological characteristics . . . 81
3.7 Treatment 1: Florence impact zone . . . 83
3.8 Treatment 2: Flooded between waves . . . 84
3.9 Treatment 3: Exposed in ’16 or ’17, not ’18 . . . 84
A1 Robustness of hedonic models . . . 102
A2 Placebo test of flood map exposure . . . 106
A3 Effect of flood map exposure with municipality-level flood data . . . 107
A4 Robustness check: Effect of map exposure with attention screen . . . 107
A5 Robustness check: Map interaction effect with attention screen . . . 108
A6 Robustness check: Effect of map exposure without imputed elevations . . . 109
A7 Robustness check: Map interaction effect without imputed elevations . . . 110

Chapter 1

How do homebuyers adapt after experiencing a natural disaster? Evidence from the Florida real estate market

1.1 Introduction

One of the biggest questions surrounding the issue of climate change is adaptation: will people adapt to rising temperatures and extreme weather events, how costly will these adaptations be, and will they be made before it is too late? There is an enormous literature studying climate adaptation around the world (Dell, Jones and Olken 2014; Mendelsohn, Nordhaus and Shaw 1994; Deschênes and Greenstone 2007; Kaiser et al. 1993). A key ingredient in adaptation is the initial understanding that a change needs to be made. Researchers have shown that first-hand experience with factors like excessive heat and natural disasters can promote belief in climate change (Deryugina 2013; Konisky, Hughes and Kaylor 2016) and spur adaptation in the context of agricultural investments or flood insurance takeup (Mase, Gramig and Prokopy 2017; Gallagher 2014). However, there is still relatively little empirical evidence on how exposure to climate change affects real estate purchases (Kahn 2016; Bunten and Kahn 2014; Boustan, Kahn and Rhode 2012). Any such response has major implications for coastal areas. Adverse demand shocks concentrated in these markets would destroy personal and commercial wealth, thereby shrinking municipal tax bases and reducing funding for schools and other public goods.

In this study, I use the timing and geographic extent of flooding events as an exogenous informational and psychological treatment on the people living nearby, and measure the effects of this "disaster shock" on subsequent participation in non-local real estate markets. I aim to shed light on two important questions: how do people translate experience with disasters into adaptive behavior, and how destructive are these adaptations for coastal real estate markets? The answers to these questions bear on whether coastal markets are overbuilt with properties that will likely be submerged within a few generations.

To analyze adaptive behavior after disasters, I pair data on over 4 million property sales from a database of Florida real estate transactions with information on over 400 flood disaster events across the United States. By focusing on the behavior of buyers who participate in Florida markets but hail from other states, I disentangle the psychological impact of disaster exposure from the direct physical damage impacts that can roil local real estate markets (Graff Zivin, Liao and Panassie 2020). Florida is a natural candidate for study as the state with the most real estate at risk from sea level rise over the next century (Union of Concerned Scientists 2018), as well as the most popular destination for out-of-state home buyers (Kerns and Locklear 2019).

I use the random variation in the timing of flood disaster declarations to identify the causal effect of recent flood exposure on home purchases in coastal Florida markets, and purchases of homes near the water in particular. I find significant evidence of retreat from the coast among buyers coming from counties with recent flood disasters,

with shocked buyers purchasing 20-30% fewer homes within 1 km of the water (relative to farther inland) than under the counterfactual where they are not shocked. I use a random-utility sorting model to estimate how much shocked buyers devalue flood-prone properties, and find suggestive evidence that shocked buyers value properties in flood zones $500 -$3,000 less than comparable non-shocked buyers. I then employ hedonic price models to quantify the impact of disaster-shocked buyers on the marginal implicit price of coastal proximity. I find that a large influx of shocked buyers to a Florida market can significantly erode the amenity value of access to the water, reducing property values in the most affected parts of the market by 1-2%.

In Section 1.2, I survey the fast-growing literature on questions of climate belief, adaptation, and real estate markets. In Section 1.3, I describe ZTRAX, the real estate database, as well as other public data sources that allow me to complete this analysis. I then proceed to attack the question of adaptive behavior in this context from three directions. In Section 1.4, I introduce simple transaction count models that allow me to identify retreat behavior, and show significant and robust evidence of retreat in Florida markets among shocked buyers. In Section 1.5, I introduce a residential sorting model that I estimate in an attempt to value the different preferences of shocked buyers in dollars terms, and measure the welfare gain to shocked buyers of the information imparted by their past storm experience. In Section 1.6, I introduce a “meta” hedonic price approach that I argue allows me to estimate the threat that retreat behavior poses to the value of the coastal amenity currently capitalized into Florida homes. In Section 1.7, I discuss the implications of my findings and propose avenues for future research.

1.2 Literature review

Below, I survey the literature on the impacts of disaster exposure at both the individual and housing market level, along with past attempts to document adaptive behavior in the wake of disasters. This study contributes to this literature by demonstrating the transfer of disaster shock in one location to markets in other locations, and also identifying a plausible climate adaptation strategy – retreat from the water – undertaken by these shocked home buyers. To my knowledge, this study is the first to analyze the differential behavior of home buyers in the same market who have, and have not, recently experienced informational disaster shocks.

1.2.1 Belief updating after a disaster shock

There is a robust literature focused on measuring belief updating in the wake of weather events like excessive heat or natural disaster that implicate climate change. Konisky, Hughes and Kaylor (2016) and Shao and Goidel (2016) measure the effect of extreme weather events on climate attitudes, combining historical disaster data with survey data from opinion polls. Both studies find some effect of past local experience on concerns about climate change, while the latter finds that political affiliations are still a much more important determinant of beliefs. Konisky, Hughes and Kaylor (2016), in particular, find that the effect is only significant for extreme weather that occurred in the four months prior to the survey, and that less recent disasters have no statistically significant link with attitudes. Botzen, Kunreuther and Michel-Kerjan (2015) conduct a case study in New York City in the years after the disastrous Hurricane Sandy flooding in 2012. They find that people tend to overestimate likelihood of flood, and underestimate damage conditional on flood, but do not have the analogous before-flood data needed to attribute this to Sandy in particular. Some researchers have studied the link between more mild climate shocks, like increased temperatures, and climate opinion. Egan and Mullin (2012) and Deryugina (2013) both find that exposure to abnormally high temperatures in one’s local area leads to increased belief that climate change is happening or will happen. Finally, in Chapter 3 of this work, we show that in the area affected by torrential flooding due to Hurricane Harvey in 2017, pessimism about flood risk faded over time in the wake of the disaster.

Another strand of the literature has tried to infer changes in beliefs from economic decisions. Gallagher (2014) finds that flood insurance uptake increases markedly in counties affected by flood disaster declarations and remains elevated for several years. Dessaint and Matray (2017) study the behavior of corporate managers, who tend to adopt more conservative portfolios and increase cash holdings when hurricanes strike nearby. Similarly, Cameron and Shah (2015) find that survivors of floods and earthquakes in Indonesia perceive higher general

levels of risk and behave in a more risk-averse manner. Gibson, Mullins and Hill (2019) look at changing real estate markets in New York City to infer belief changes attributable to both Sandy and flood zone reclassification. However, neither is able to connect these responses to explicit belief measures. The present study fits most closely into this latter literature, as I am not able to measure changes in beliefs directly, but observe substantial changes in investment decisions (specifically house purchases) that are directly attributable to disaster shocks.

1.2.2 Disaster shocks in coastal real estate markets

Several studies have documented and quantified the effects of local natural disasters, but few have been able to decouple local flood damage from possible changes in beliefs, and only one has taken advantage of information on heterogeneous contemporary buyers to explain purchase decisions in the context of climate change.

The New York City real estate market in the wake of Hurricane Sandy has attracted particular researcher interest. Gibson, Mullins and Hill (2019) find that the publication of new flood maps and the arrival of Hurricane Sandy each separately had large depressive effects on real estate prices in flood-vulnerable parts of New York. Similarly, Ortega and Taspinar (2018) find that property values for non-damaged structures in areas affected by Hurricane Sandy declined 8% by 2017, five years after the storm. These papers are closely related to this study, in that they measure the effect of disaster shocks on real estate markets, but they can not disentangle the psychological or informational impact of the disaster from the physical damage to New York City neighborhoods.

A host of studies of other specific disasters suffer from the same limitation. Studies of the 1993 Mississippi River floods in St. Louis (Kousky 2010), a major flood in Georgia in 1994 (Atreya and Czajkowski 2014) and repeated flooding in Fargo, North Dakota (Zhang and Leonard 2019), as well as a series of hurricanes in Florida (Graff Zivin, Liao and Panassie 2020), all find significant negative effects of disasters on property values, including differentially worse effects for the most flood-prone properties. McCoy and Walsh (2018) find a similar result in the context of wildfires at the wildland-urban interface near Denver, Colorado. However, a common limitation of all these studies is that they cannot separate the psychological and informational impact of disasters on the people living there from the impact on the physical infrastructure in the market. The present study surmounts this problem by focusing on residents who are shocked by an event in one location, and then enter a real estate market in Florida, far from the original disaster.

1.2.3 Climate adaptation in real estate

Few studies have specifically addressed the question of how ongoing climate change will affect long-term migration patterns, as opposed to the immediate migrations induced by mass destruction (Kahn 2016). Albouy et al. (2016) use hedonic models of U.S. cities' climate characteristics to predict that excessive heating will drive more people to migrate to the Pacific Northwest, but do not attempt to predict how the real estate markets within metropolitan areas will be affected. Bernstein, Gustafson and Lewis (2019) use the same database as the present study to analyze homes at risk of sea level rise (SLR) and find that these homes sell for less than homes that are similar on observables. The effect is concentrated among second-home purchasers, who the authors infer are more sophisticated than other buyers based on the education and income levels in their home ZIP codes. Similarly, Baldauf, Garlappi and Yannelis (2018) find that the SLR discount is more pronounced in "believer" neighborhoods where more survey respondents indicate a belief in that climate change is happening. Keys and Mulder (2020) find that buyers in Florida real estate markets seem to have become more sensitive to sea-level-rise zones over the past two decades – although sellers' attitudes have not changed – and postulate that this could be due to rising awareness of climate change generally. Murfin and Spiegel (2020), by contrast, find that what they term "relative sea level rise," projections that account for land subsidence, is not properly capitalized by markets.

Another set of studies takes advantage of changes and updates to official flood maps that occur periodically, even in the absence of any local disaster. These studies detect responses in real estate markets to determine the extent to which flood risk information is capitalized in this market, generally finding that there is a response (Hino and Burke 2021; Mulder 2021; Shr and Zipp 2019).

I am aware of two studies that analyze the effect of non-local shocks that could plausibly have affected real estate market participants only through their information content. Muller and Hopkins (2019) study the New Jersey coastal real estate market and use non-local disasters like Hurricane Katrina that garnered national media attention to measure the role of information in coastal real estate markets, finding a significant negative effect on price for homes in flood zones. Hallstrom and Smith (2005) study the effect of Hurricane Andrew – a massively destructive storm that ravaged Miami-Dade County and other parts of southern Florida in 1992 – on nearby Lee County, which narrowly escaped damage. They also find that homes in flood zones lost significant value, which they attribute to buyers and sellers recognizing the information conveyed by the storm and adapting their market behavior.

In sum, these studies explore purchasing behavior in areas impacted by disasters, but few considerd how buying behavior differs across heterogeneous buyer types (Baldauf, Garlappi and Yannelis 2018; Keys and Mulder 2020). To my knowledge, no study examines the differential buying behavior of individuals who have experienced disasters far from a given real estate market. This geographic remove is crucial to establish the informational and psychological aspects of disaster exposure apart from the equilibrium effects of local disaster damage.

1.3 Data

To study the question of climate adaptation in real estate markets, I rely on comprehensive Florida real estate transaction and assessment data from ZTRAX, a database provided by Zillow Inc (Zillow 2021).[1] This database provides detailed information on each property transacted and, crucially, geographic characteristics of buyers including their designated mailing addresses, which in the basis for my assignment of experienced flood history to buyers.

I complement the ZTRAX database with hydrological data on each property from U.S. government sources. With precise latitude and longitude coordinates from ZTRAX, I match each property to a flood zone classification, a sea-level-rise submersion zone, and calculate a straight-line distance to the water. Geographic data on flood zone classifications as designated by the Army Corps of Engineers is available from the U.S. Federal Emergency Management Agency (2021), while sea-level-rise projections are available from the National Oceanic and Atmospheric Administration (2021). I use QGIS mapping software to calculate the distance from each property location to the water with a detailed shoreline map from the University of Florida GeoPlan Center (2019).

Additionally, I draw on the Florida Cooperative Land Cover map maintained by the Florida Fish and Wildlife Conservation Commission (Florida FWC 2021) to identify neighborhood amenities such as parks and recreational sites. Finally, to assign a treatment measure to each observation, I use county-level historical disaster data maintained by the U.S. Federal Emergency Management Agency (Federal Emergency Management Agency 2019).

In the remainder of this section, I describe these sources in more detail.

1.3.1 ZTRAX database

The ZTRAX database provides information on real estate transactions and assessments across more than 2,700 U.S. counties dating back to the 1990s. In this study, I restrict focus to data on properties sold in Florida between January 2005 and December 2017, the period for which the transaction data was the most comprehensive at the time I accessed it. Zillow compiles data from county authorities and claims essentially universal coverage of transactions during this period (Zillow 2021). In approximately 48% of cases, I can identify the buyer's county of origin, as indicated by the buyer's mailing address, and in nearly every case I can identify the precise geographic location of the property transacted. ZTRAX data contains many observations with missing or seemingly erroneous information. I am forced to drop a small percentage of observations without listed prices or with listed prices of $0. I also exclude properties with listed sale prices above $10 million, because it appears that some of these are likely data entry errors (these account for less than 0.5% of all transactions). Additionally, some properties are

[1] Data provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author and do not reflect the position of Zillow Group.

Table 1.1 Descriptive statistics

	(1) **Full sample, 2005-2017** *(n = 4,764,491)*	(2) **Matched to assessment data** *(n = 4,451,961)*	(3) **Known origin only** *(n = 2,273,503)*	(4) **Buyers from 818 main origin counties** *(n = 963,412)*	(5) **Buyers from main origin counties, frontage only** *(n = 84,634)*
Property characteristics					
Sales Price ($k)	247	245	227	265	504
	(350)	(335)	(378)	(429)	(695)
Square footage	1,953	1,945	1,844	1,901	2,071
	(1,023)	(1,019)	(1,003)	(983)	(1,329)
Number of stories	1.2	1.2	1.18	1.17	1.27
	(0.48)	(0.48)	(0.46)	(0.5)	(0.57)
Year built/remodeled	1989	1989	1989	1993	1989
	(19)	(19)	(20)	(17)	(17)
Single-family home	61%	62%	53%	48%	37%
Has garage	30%	30%	25%	27%	21%
Has pool	19%	20%	17%	20%	32%
Has boatslip	2%	2%	2%	3%	27%
Hydrological characteristics					
Distance to water (km)	3.55	3.47	3.3	3.2	.02
	(4.86)	(4.65)	(4.84)	(4.63)	(.01)
Frontage (<50 m to water)	6%	6%	7%	9%	100%
100-year flood zone	19%	19%	22%	23%	67%
500-year flood zone	29%	29%	31%	32%	76%
6-ft SLR submersion zone	14%	14%	16%	15%	54%
Transaction characteristics					
Floreclosure sale	21%	21%	24%	20%	13%
Indeterminant origin county	52%	53%	0%	0%	0%
Buyer origin outside FL	16%	16%	33%	74%	83%
Months since last flood	122	122	118	104	104
in buyer's county	(54)	(54)	(59)	(69)	(74)
Floods in last ten years	3.33	3.34	3.21	2.99	2.76
in buyer's county	(2.06)	(2.07)	(2.01)	(1.92)	(1.81)

Mean values and standard deviations for key variables in the sample. Column 1: the full sample include all residential home purchases in Florida between January 2005 and December 2017 for which reliable sales price information, transaction details, and precise geographic location data are available. Column 2: this sample excludes transactions for which thorough assessment data was not available, internally contradictory, or otherwise not reliable. Column 3: a transaction is included in the "known origin only" sample if the buyer lists a correspondence address that is not identical to the property address. Column 4: a county is included in the 818-county group if there are at least 100 purchases in the full sample associated with that county. Column 5: a subset of the sample in Column 4, inluding only properties within 50 meters of the water. Source: Author's analysis of Zillow ZTRAX database, and GIS calculations based on University of Florida coastline maps, FEMA flood zone maps, and NOAA sea-level-rise maps.

Table 1.2 Transactions by metropolitan statistical area (MSA), origin data only

Metropolitan area	# of Ctys	Sales count (k)	Mean sales price ($k)	Non-FL buyer	New York buyer	Commonest non-NY origin	Condo	Flood zone	Sea level rise zone
Daytona Beach	2	118	176	34%	7%	New Jersey	19%	12%	6%
Destin	2	66	389	53%	1%	Georgia	28%	10%	2%
Ft Myers	1	259	202	47%	6%	Illinois	23%	30%	8%
Homosassa	1	41	89	31%	6%	New Jersey	2%	9%	7%
Jacksonville	5	159	177	24%	2%	California	15%	10%	6%
Miami	3	676	269	22%	7%	New Jersey	48%	19%	9%
Naples	1	104	463	61%	7%	Illinois	50%	16%	10%
Panama City	2	45	223	53%	1%	Georgia	36%	13%	6%
Pensacola	2	50	149	28%	1%	Alabama	9%	12%	4%
Port St. Lucie	2	85	184	33%	8%	New Jersey	18%	5%	3%
Punta Gorda	1	72	130	47%	6%	Massachusetts	11%	44%	19%
Sarasota	2	152	271	47%	6%	Ohio	28%	18%	6%
Space Coast	1	84	146	31%	6%	California	19%	8%	4%
Tampa	4	365	167	27%	4%	California	21%	12%	6%
Vero Beach	1	32	275	40%	7%	New Jersey	22%	19%	8%
Full sample	30	2,307	228	33%	6%	New Jersey	30%	17%	7%

Source: Zillow ZTRAX database and U.S. Census metropolitan area definitions. Information on properties in flood zones and sea-level-rise zones is provided by FEMA and NOAA respectively. Figures include all transactions between January 2005 and December 2017 for which reliable sales price data and buyer origin data are available. As such, they reflect only 48% of total transactions during this period.

listed with latitude and longitude coordinates that are inconsistent with county or ZIP code information provided; these are dropped where they can be identified. See Appendix A1 for more details about the myriad reasons that some transactions had to be dropped from the main sample.

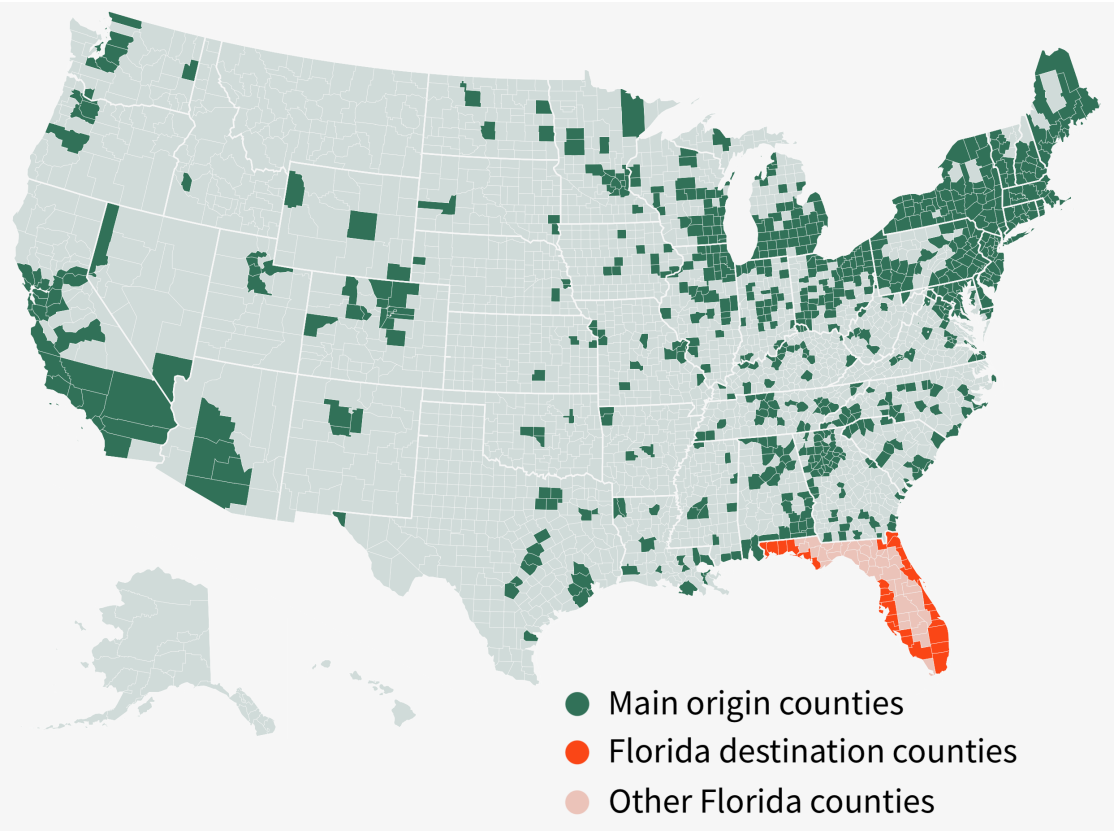
Condominium properties present an additional challenge for my analysis. People who live in condominiums or high-rise buildings are effectively decoupling access to the water from flood risk, which indeed may be part of the appeal. I cannot reliably observe what floor a property is located on in a large building in the ZTRAX data. In some analysis, I treat condos separately from other observations as the geocoding of condo properties to flood zones may not represent an accurate assessment of that property's risk, even abstracting away from any inaccuracies in the flood maps themselves. On a similar note, I cannot distinguish properties where the living areas are raised on stilts or otherwise elevated, leading to some inevitable measurement error in my flood risk measures.

ZTRAX assessment data drawn from government sources has a rich complement of structural variables for each property, although for about 17% of transactions I am unable to successfully match the underlying property to assessment data, usually due to a lack of assessment record for a particular parcel or incomplete records without key data on square footage or building age. Unmatched observations are still used in some parts of the analysis, but cannot be included in models that rely on property characteristics like square footage.

Table 1.1 displays descriptive statistics for the sample. The full sample (Column 1) include all residential home purchases in Florida between January 2005 and December 2017 for which reliable sales price information, transaction details, and precise geographic location data are available. Columns 2 and 3 show two distinct sub-samples of this full sample. In Column 2, the sample excludes transactions for which thorough assessment data were not available, internally contradictory, or otherwise not reliable. This is the core sample for the hedonic analysis in Section 1.6. A transaction is included in the "known origin only" sample (Column 3) if the buyer lists a correspondence address that is not identical to the property address (see Section 1.3.2 below for more on this issue). Column 4 presents a subset of the known-origin sample that includes only non-local buyers (i.e. not from the same MSA as the property purchased) and is restricted to the 818-county group with at least 100 purchases in the full sample associated with that county. This is the core sample used in the count models in Section 1.4. Column 5 displays a subset of the sample in Column 4, including only properties within 50 meters of the water,

Figure 1.1 Map of 818 main Florida origin counties

Counties qualify as origin counties if there are at least 100 transactions in the sample between 2005 and 2017 where the buyer address is associated with that county. The counties comprising the 15 metropolitan areas in Florida included in the study are highlighted in orange.



which I term "frontage" properties.

The descriptive statistics reveal some distinctive features of homes purchased by out-of-MSA buyers (Column 4), relative to the broader group of known-origin homes in Column 3. While these homes are nearly 16% more expensive than the average home sold, they are less than 4% larger. They do, however, enjoy significantly better access to the water, being 4% closer to the water on average, and slightly more likely to be in the frontage zone (within 50 meters of the water) and the 100-year flood zone. These differences provide suggestive evidence that out-of-state buyers prioritize access to the water more highly than local buyers when making purchase decisions.

The housing market in the frontage zone, which may be most radically affected by changing beliefs about disaster risk, is distinctive from other parts of the market. Frontage homes are nearly twice as expensive as the average property, and are significantly larger, although slightly older on average. These deluxe properties likely occupy a separate market apart from the general coastal real estate market, which might explain the relatively resilient demand for these homes among shocked buyers that I find in some models.

Table 1.2 breaks down the sample by metropolitan area (MSA). I restrict my focus to these MSAs, which are the 15 coastal MSAs in Florida as defined by the U.S. Census Bureau, and contain over 95% of coastal county observations.[2] The MSAs range in size, average home price, and the share of buyers coming from out of state,

[2] A sixteenth Florida MSA, Tallahassee, includes two coastal counties, but I exclude it for two reasons. First, the coastal areas in this MSA are relatively sparsely populated, as they are mostly wetlands covered by National Wildlife Refuge status and other state and federal protections. Second, the central MSA area around Tallahassee, where most of the population lives, is not itself located near the coast. In all other cases, the coastal MSAs I do include are centered around cities that are themselves coastal.

reflecting the fact that some areas are more popular retirement or second-home markets, while others may be attracting mostly people relocating for work or family. Out-of-state buyers are most prevalent in Fort Myers, Sarasota, Naples, and Punta Gorda, which are clustered together on the Gulf Coast in the southwestern part of the state.

Figure 1.1 displays a map of the 818 counties that I include in my count models. Counties qualify as origin counties if there are at least 100 transactions in the main sample between 2005 and 2017 where the buyer address is associated with that county. Unsurprisingly, there is a noticeable skew toward the eastern half of the United States, particularly the most populous urban and suburban areas. Metropolitan areas in the northeast are especially over-represented. I limit myself to these counties to reduce the number of observations with zero counts in the negative binomial regressions. This group includes over 94% of out-of-state buyers in the nationwide sample.

1.3.2 Determining buyer origin

To apply my treatment variable, which is the amount of time since the most recent flood disaster in a buyer's home county, I need to reliably determine where each buyer was living in the months and years prior to the transaction. The ZTRAX data for each transaction contains information on the buyer's mailing address as recorded by the local government for future tax correspondence purposes. I use this variable to identify the buyer's home ZIP code and county, and in turn to define the treatment variable for each observation.

In cases where a buyer is purchasing a home in Florida and local officials record a residential address in Wisconsin or Massachusetts as the buyer's mailing address, it seems reasonable to assume that this buyer was previously living – and possibly is continuing to live – at the out-of-state address, and the new property is an investment or a second home.

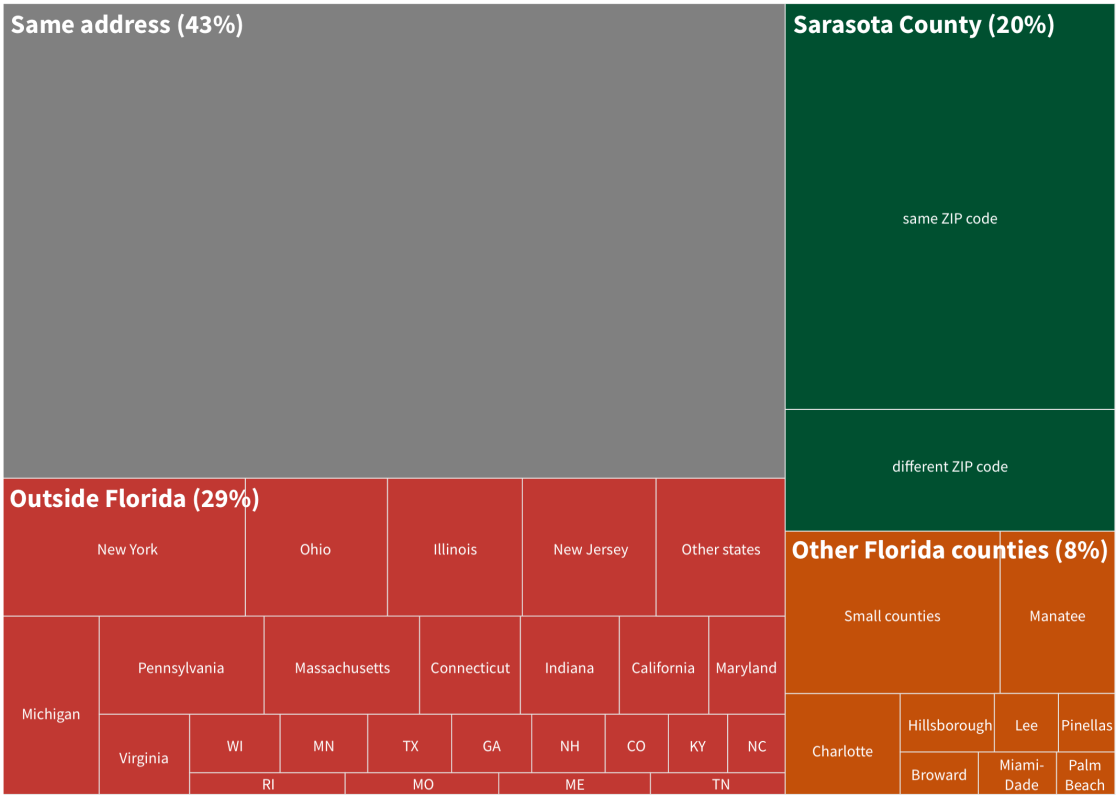
However, in approximately 52% of cases, the buyer's mailing address exactly matches the address of the property purchased, which I term the "same-address listing" phenomenon (see Figure 1.2 for an example breakdown of the buyer address variable in Sarasota County). In these cases, the buyer's address does not seem like a reasonable basis for determining a buyer's previous location. The buyer could plausibly have been previously living in Wisconsin, Massachusetts, some other part of Florida, or right down the street. In these cases, I have no way to identify the origin of these buyers, and no way to impute their recent flood history.

Previous literature has interpreted the same-address listing behavior as intent to immediately occupy the home, and to receive official correspondence at the new address going forward (Chinco and Mayer 2016; Graff Zivin, Liao and Panassie 2020). This would imply that non-same-address listers, by contrast, comprise some combination of investment companies, second-home purchasers, and people who intend to relocate to Florida permanently or semi-permanently in the intermediate future but not right away. Assessment data from ZTRAX includes information about whether properties were assessed as owner occupied and whether owners are claiming the homestead tax credit, which implies owner occupancy in most cases. These data sources are somewhat incomplete and mutually contradictory, and do not correspond closely to same-address listing behavior either.

Regardless of the root causes of this behavior, my analysis in Sections 1.4 and 1.5 will necessarily be limited to comparisons within the large minority of buyers where this address data does not match the property in question. I argue that this analysis is still valid and important. First, Florida is a major market for investment and second-home purchases, and this market is important in its own right. Second, previous research shows that non-owner-occupiers tend to be more savvy about climate risk (Bernstein, Gustafson and Lewis 2019), and detecting changes in their behavior in response to disaster experience may be of particular interest. Finally, I show evidence in Appendix A2 that while there is a statistically significant relationship between same-address listing behavior and incoming buyers' disaster experience in some MSAs, the economic significance of the relationship is small. For this relationship to explain my results would require a very peculiar selection effect: people from recently-shocked counties are no less likely to buy homes near the water, but those buying homes near the water are likelier to occupy them immediately.

Figure 1.2 Tree map of "buyer mail" variable in Sarasota County transactions

A tree map of the buyer mailing address variable from ZTRAX transactions in Sarasota County, 2005-2017. Each rectangle represents a mutually exclusive group of transactions, with the size of the rectangle representing the group's relative size. Buyer mailing address is the key variable I use to determine each buyer's disaster shock treatment. Buyers who list a mailing address that matches the address of the property purchased ("same address" buyers) comprise a large minority of the sample in most counties, and their origins can not be identified. These purchases are excluded from portions of the analysis that rely on identifying buyers' origin counties. See text for more details about my interpretation of the same-address listing behavior.



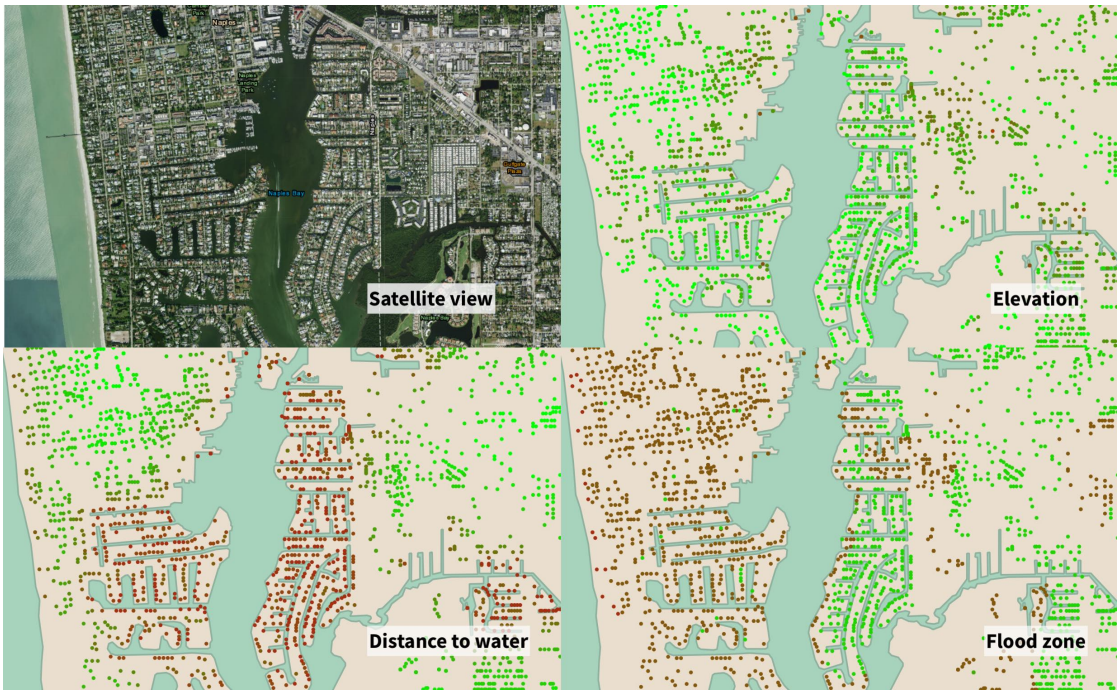
1.3.3 Hydrological data

I complement the ZTRAX database with data that pertain to each property's proximity to water, a factor which is generally understood as an attractive amenity, but which may also repel disaster-shocked buyers in particular. Using precise latitude and longitude information from ZTRAX, I match each property to a flood zone classification from the Army Corps of Engineers using flood maps made available by the U.S. Federal Emergency Management Agency. I categorize properties into five zones: A (corresponding to a 1% or greater annual risk of inundation),[3] VE (areas with 1% or greater risk and additional risk from storm surge flooding), AH (corresponding to a 1% or greater annual risk of shallow flooding or pooling), X (corresponding to an annual risk of inundation between 0.2% and 1%), and the "minimal flood hazard" zone, corresponding to a flood risk lesser than that of the 500-year flood zone. See Appendix Figure A1 for an example flood map as it appears on the FEMA website. Likewise, I match each property to one of seven feet-of-SLR submersion zones identified by National Oceanic and Atmospheric Administration (2021) that correspond to the amount of future sea level rise that would result in property submergence.

[3]The "A" flood zone designation has been replaced by "AE" on updated flood maps, but in either case they designate areas where FEMA is warranting a 1% annual risk of flood or greater.

Figure 1.3 Hydrological characteristics of coastal properties in Naples, Florida

These screen captures from QGIS software illustrate properties from the sample superimposed on a highly-detailed map of the Florida coastline, including inlets, canals, and artificial water features like marinas. In each capture, properties are color-coded according to a different measure of flood vulnerability, with red properties facing the greatest risk and green the least. Quantities calculated using property latitude and longitude data mapped to flood zone data from FEMA, coastal boundary data from the University of Florida GeoPlan Center, and sea-level-rise data from NOAA.



Obtaining information on each parcel's distance to the water is more involved. I use GIS mapping software to calculate the straight-line distance from each property location to a detailed shoreline map from the University of Florida GeoPlan Center (University of Florida GeoPlan Center 2019). This allows me to observe homes that are relatively far from the open ocean but close to features like inlets, canals, and marinas that could be a source of floodwater, and code them accordingly (see Figure 1.3).

Figure 1.4 illustrates the distribution of flood zones for all properties in the 15-MSA sample within 10 km of the water. Owing to the state's famously flat topography, properties in Florida face significant flood risk even when they are relatively far from the water. While naturally a large percentage of properties in the "frontage" zone (50 m from the water or less) are in the 100-year flood plain, a non-negligible fraction of properties even further inland are also so labeled. For example, over 15% of homes between 2 and 4 km from the water are in one of the two flood zones. This has implications for where I would expect retreat behavior to arise, not simply from the immediate frontage areas to the areas a bit inland, but even from areas more than one kilometer from the water. Below, I show results indicating that retreat is detected in the region 50 m to 1 km from the water.

1.3.4 Flood event data

My source of identification in the analysis below is the random timing of major flood events in U.S. counties over the 2005-2017 period, as determined by Presidential Disaster Declarations (Gallagher 2014). I use historical county-level disaster data maintained by the U.S. Federal Emergency Management Agency. Each disaster is listed with a geographic range and an official description, such as "severe winter storms, heavy rains, and flooding." A