Title
Essays in Environmental Economics

Author
Sutula, Deirdre B.

Publication Date
2019

Peer reviewed|Thesis/dissertation

Essays in Environmental Economics

by

Deirdre B. Sutula

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Agricultural and Resource Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor James Sallee, Chair
Associate Professor Meredith Fowlie
Associate Professor Solomon Hsiang

Spring 2019

Essays in Environmental Economics

Abstract

Essays in Environmental Economics

by

Deirdre B. Sutula

Doctor of Philosophy in Agricultural and Resource Economics

University of California, Berkeley

Associate Professor James Sallee, Chair

Climate change and habitat loss are at the forefront of policy discussions, but can be difficult to evaluate using an economist's toolkit because the associated damages are external to most markets. Nevertheless, there is a robust body of economics literature on climate change and on ecosystem services. This dissertation provides methodological and policy contributions to these bodies of literature; I provide a few small answers to the broad questions "What are we are missing as economists?" and "How can we positively influence human behavior?". I focus on two issues in particular: flooding and deforestation. These are two topics that are globally relevant and highly measurable, but difficult to study. Despite the fact that flood risk tends to be well-documented, there are many complications that affect the interpretation of flooding research. It is not always clear which homes are flooded, nor whether property sale prices should reflect true costs. It is also not clear how we can mitigate these costs or how we can help consumers make informed choices. As for deforestation, although there are excellent data on forest cover and forest loss, studying deforestation presents its own set of difficulties. Namely, the costs of deforestation are largely social costs and not private costs, and thus are not included in any market price. I present some ways in which deforestation has relatively local (within-country) effects, and so may be able to be valued in the market via government intervention. Overall, I argue that (1) there is particular complexity at the nexus of public policy, human behavior, and ecological systems, and (2) despite this fact, political and research progress can be made through careful work.

In chapter 1, co-authored with Rachel Baker, we tackle the question of how policymakers can influence human behavior, specifically in regard to floods and the property market. In most property markets, it is not clear whether homeowners and buyers understand or respond to the risk of floods; in fact, there is evidence that home buyers have difficulty understanding risk information, and forget about actual events within a few years. Although consumers may not respond to risk information, prices provide another potential signal to induce market response. In this chapter, we provide some of the first empirical estimates of whether home buyers are attentive to disaster insurance prices. We take advantage of an exogenous shock to insurance prices: a policy change that subsidizes flood insurance for high-risk homes in the

UK. Using a dataset of all residential property sales in England, we find that home buyers are quite sophisticated in their response to the price of flood insurance. Homes that receive flood insurance subsidies rise in value by 9.8%, implying full or close to full capitalization of insurance prices. Our results suggest that risk-based disaster insurance pricing can influence the housing market and promote adaptation to climate change.

In chapter 2, I examine some previously understudied impacts of deforestation, and their consequences for both policymakers and applied econometricians. Contrary to what is commonly believed among economists, there is evidence from the atmospheric science literature that local precipitation can be significantly affected by human activity. In particular, changes in forest cover in tropical regions has been found to affect precipitation by as much as a factor of two. However, these analyses do not fully account for the reverse causality issues associated with a regression of rainfall on forest cover. In this chapter, I use an instrumental variables approach to estimate the effect of exogenous changes in forest cover on rainfall, using global agricultural prices as instruments. I find that forest cover affects rainfall to a greater degree than previously estimated, and I argue that both economists and social planners need to think carefully about the unintended impacts human activity can have on all aspects of environmental services, including the local climate.

Chapter 3 of this dissertation once again explores methodological issues within economics, but this time in regards to natural disaster cost estimation, There is a rich literature that estimates the costs of and reactions to natural disasters using property sale prices. Post-disaster property sale prices reflect changes in the physical state of buildings (physical damage net repairs) as well as changes to perceived risk. These effects can work in opposite directions, as property owners may receive insurance payouts and renovate their homes before putting them on the market. In this paper, I use high-resolution flood outline data in England to examine the effects of a flood on the price of homes just inside versus just outside of a flood boundary, which should face similar changes in consumer perception of local risk. I find that homes that are sold after being flooded *increase* in value relative to homes just outside a flood boundary, which indicates that the net effect of physical damage and post-flood repairs is positive. These results suggest that using post-disaster property sale prices will lead to an underestimation of the true costs of the natural disaster.

Dedicated to Peter Berck

Contents

List of Figures

List of Tables

Acknowledgments

I am so thankful for the many professors, family members, colleagues, and friends who have supported me throughout graduate school. In particular, the completion of this dissertation would not have been possible without Jim Sallee or the late Peter Berck. Jim's support, guidance, and ability to consistently positively reframe any setback with a brilliant economic model was what allowed me to finish my final year of writing the dissertation. I will forever miss Peter's warm, sunny office with his always-open door and constant stream of visitors. Thank you Peter for supporting me for so many years with your perfect blend of kindness, incisive advice, and free espresso. I am also indebted to Max Auffhammer and Meredith Fowlie for their continuous help and guidance.

I gratefully acknowledge the support and helpful comments from Michael Anderson, Danamona Andrianarimanana Tamma Carleton, Karl Dunkle-Werner, Dina Gorenshteyn, Jonathan Kadish, Larry Karp, Erin Kelley, Greg Lane, Ethan Ligon, John Loeser, Dave McLaughlin, Sol Hsiang, Jeremy Magruder, Jeff Perloff, Elizabeth Ramirez-Ritchie, Betty Sadoulet, Hilary Soldati, Itai Trilnick, Sofia Villas-Boas, Reed Walker, Brian Wright, and Matt Woerman. Particular thanks is due to Rachel Baker for her coauthorship, including the inspiration for two of these chapters and her brilliant knowledge of the issues and data. I am grateful to Sophie Andrews, Sirjan Kafle, and Paul Shao for their suberb research assistance. Thank you to my ARE graduate cohort for six years of research advice, moral support, and the occasional bout of fun.

Finally, I want to thank my family and friends. In particular, I would like to thank my parents for teaching me to overcome setbacks and focus on what is important. Thank you to Elise Epner and Irwin Feintzeig for their encouragement and generosity over the years. Finally, I am eternally grateful to my partner Jacob Feintzeig for his unwavering support and for always believing in me.

Chapter 1

Does the Housing Market Respond to Flood Insurance Prices? Evidence from England

Joint with Rachel Baker

1.1 Introduction

The risks we face from natural disasters are both difficult to understand and expensive to insure against. Evidence shows that homeowners have trouble finding and interpreting natural disaster risk information (Chivers and Flores, 2002). While actual events provide a signal of risk, the effects of natural disasters on the housing market fade within a few years (Atreya et al., 2013; Bin and Landry, 2013a; McCoy and Walsh, 2018). Ideally, as the climate changes, there would be decreased investment in areas where the risk of catastrophe is increasing. For this to happen, the market needs to respond to risk information, price information, or both. Although researchers have found that consumers are inattentive to flood *risk*, flood insurance *prices* might be more salient. However, buying a home is a very complicated transaction, so it is not obvious whether consumers respond to changes in insurance prices. There are many papers that find that consumers are not well informed when making important financial decisions (Bucks and Pence, 2008; Chan and Stevens, 2008; Bhargava et al., 2015; Agarwal et al., 2017; Ho et al., 2017), but there are few empirical studies of whether home buyers are attentive to the price of disaster insurance.

In this paper, we provide one of the first estimates of the effect of flood insurance prices on the value of a home. Using 1.8 million residential property transactions from 2012 - 2018, we analyze the price and quantity of homes sold in England when insurance prices change exogenously. We take advantage of a natural experiment in which a policy change in the UK effectively subsidizes flood insurance for high risk homes. We exploit the multiple criteria that determine whether a home receives a subsidy by using a triple-differences design. To qualify for subsidized insurance, homes must be built before 2009. To benefit from the subsidized insurance offerings, homes must be in high-risk areas. We use these two characteristics, in addition to pre- versus post-policy implementation, to find the effects of subsidized insurance on the prices of homes that are eligible for the subsidy.

We find that homes that qualify for the subsidy increase in value by approximately 9.8%. If home buyers and sellers fully understood and paid attention to flood insurance prices at the time of sale, then the price of the home should rise by present discounted value of the subsidy. This is referred to as "full capitalization," where the value of the subsidy is reflected in the value of the home. With a 9.8% increase in home values, we cannot rule out full capitalization, and we can rule out no capitalization. Thus we conclude that the housing market is quite responsive to changes in the price of flood insurance.

We also examine the distributional effects of this flood insurance subsidy. Although policymakers justified the policy by touting its impact on the lower-middle and middle class, a subsidy that protects home values has the potential to be highly regressive (Department for Environment Food & Rural Affairs, 2014). This is particularly true in this case, as there is no home value cutoff for eligibility. We find that, although low-income and high-income homeowners qualify for the subsidy at similar rates, the majority of the subsidy will flow to high-income homeowners.

The paper most similar to ours is Gibson et al. (2017), which studies the effects of the Biggert-Waters Flood Insurance Reform Act, Hurricane Sandy, and changes in floodplain maps on housing prices in New York City. The Biggert-Waters Act does provide exogenous variation in flood insurance prices, which should allow Gibson et al. (2017) to estimate the causal effect of flood insurance prices on home values. However, the authors are unable to reject the null hypothesis that flood insurance prices do not affect home prices, perhaps because Hurricane Sandy and floodplain map updates both happened concurrently with this policy change. Fortunately, we have found a policy change in the UK with fewer confounding events. We find consistently statistically significant effects of flood insurance prices on home values.

This paper contributes to the literature on homeowner understanding of and attention to the risk of natural disasters. There is a long history of literature showing that people have difficulty interpreting the risk they face from low-probability events (Camerer and Kunreuther, 1989; Kunreuther et al., 2001; Keller et al., 2006). Within the natural disasters literature, many papers have found that consumers do care about disasters in the period immediately following an event, but the salience of the risk fades within a few years (Atreya et al., 2013; Bin and Landry, 2013a; Gallagher, 2014; McCoy and Walsh, 2018). The problem of limited attention is likely exacerbated by that fact that, even in high risk areas, home buyers may not see information on the flood risk until late in the purchasing process (Chivers and Flores, 2002). Our work contributes to this literature by examining whether consumers pay attention to a price signal, which may be more salient and easier to understand than a risk signal. Because there is rarely exogenous variation in insurance prices, this is one of the few papers that causally estimates the effect of flood insurance prices on home values.

In addition, this work informs the literature on attention when making important financial decisions. We know that people are not well-informed in regard to their pensions (Choi et al., 2002; Gustman and Steinmeier, 2001; Chan and Stevens, 2008) and health insurance choices (Bhargava et al., 2015; Ho et al., 2017). When purchasing cars, consumers exhibit left-digit bias[1] (Lacetera et al., 2012) and are influenced by factors as trivial as the weather at the time of purchase (Busse et al., 2015). In regards to housing, borrowers often do not know the terms of their mortgage (Bucks and Pence, 2008), and make suboptimal mortgage refinancing decisions (Agarwal et al., 2015; Keys et al., 2016; Agarwal et al., 2017). Due to the complexity of the transaction, we might expect that homebuyers are unable to perfectly incorporate the costs of flood insurance into the decision of what price (if any) to offer for a home. On the other hand, there is evidence that consumers are quite good at incorporating estimated future fuel costs when buying homes (Myers, 2018) or cars (Busse et al., 2013; Allcott and Wozny, 2014; Sallee et al., 2016; Grigolon et al., 2018). Our results are more similar to this latter set of papers; we find that consumers are sophisticated in regard to flood insurance prices.

Finally, this paper contributes to the literature on effective means to promote climate change adaptation. There has been research on human adaptation to increasing cyclone risk (Hsiang and Narita, 2012), extreme heat events (Barreca et al., 2016) and rising agricultural damages (Auffhammer and Schlenker, 2014; Burke and Emerick, 2016). The literature on means to promote and guide adaptation, however, is mostly theoretical (Fankhauser et al., 1999; Pizer, 2002; Grothmann and Patt, 2005; Thomalla et al., 2006; Adger et al., 2009). There are a few closely related papers to ours; Tompkins et al. (2010) find that most climate adaptation in the UK takes place via large government investment in infrastructure. Similarly, Ford et al. (2011) find that most climate adaptation worldwide happens via government intervention. Ours is one of the first papers to examine the potential for price signals to create climate adaptation[2].

The rest of this paper will proceed in the following manner: In section 1.2, we discuss the policy change that took place in the UK. Section 1.3 describes the data that we use, and section 1.4 describes our empirical strategy. Section 1.5 calculates what full capitalization of the subsidy would be before we present the results in section 1.6. Section 1.7 presents results from various robustness checks on our analysis. Section 1.8 discusses the distributive effects of this policy, and section 1.9 presents our conclusions about our findings.

1.2 Background

The policy change that we study takes place in 2016, precipitated by events taking place over the previous two decades. Before 2016, the flood insurance market was entirely privately provided, with government interference only via the building of flood defenses. Private flood

[1]Left-digit bias is the tendency to trim rather than round when interpreting numbers, such as interpreting \$1.99 as \$1.

[2]There are a number of papers that examine the effectiveness of taxes as a means for *mitigation*: Fullerton and West (2002); Davis and Kilian (2011); Martin et al. (2014); Rivers and Schaufele (2015).

insurance has been required to obtain a mortgage in England since the 1970s, but, due to inability to accurately determine risk, low-risk and high-risk homes paid similar rates through the early 1990s (Arnell et al., 1984; Penning-Rowsell et al., 2014). Starting in the 1990s, insurance providers gained access to better and better maps of flood risk factors, and increasingly set prices based on risk (Department for Environment Food & Rural Affairs, 2014). In 2001, insurers reached an agreement with the UK government that allowed insurers to refuse coverage to homes with annual flood risk above 1.3%, or 1 in every 75 years (Association of British Insurers, 2002). Although very few homeowners were actually unable to obtain flood insurance offers after 2002 (Lamond et al., 2009), there was increasing outcry and media coverage of rising flood insurance prices, particularly after the major flood events in the spring of 2012 and the winter of 2013-2014, which together caused nearly £1 billion in damages (Environment Agency, 2013, 2016).

Due to concern among policymakers about the affordability of flood insurance, and pressure to protect insurance companies from high levels of risk, the UK government published a draft of proposed flood insurance legislation in September 2013, and passed the legislation as part of the Water Act on May 14, 2014. The Water Act called for the creation of a highly regulated flood reinsurance scheme, called Flood Re. Under this scheme, flood insurance companies must offer coverage to all owner-occupied residential properties built before 2009 in buildings with three or fewer units. (Thus new properties, rented properties, and large apartment buildings are excluded.) Insurance companies pay into Flood Re, and Flood Re reimburses them when participating homes are flooded.

This reinsurance scheme is somewhat atypical in that insurers can choose which policies to reinsure. When an insurer offers coverage to a homeowner, the insurer decides whether to pass on the risk of that particular property to Flood Re. If the insurer chooses to pass on the risk, they must pay Flood Re a fixed rate. If the property is flooded and the homeowner submits a claim, the insurer will be reimbursed by Flood Re. If the insurer chooses not to pass the risk on to Flood Re, no premium is paid to Flood Re and Flood Re will not reimburse the insurer in the event of a flood to that property. This will clearly lead to adverse selection into Flood Re, further exacerbated by the fact that premiums to Flood Re are not based on risk. Insurers pay Flood Re based on the value of the home, but not based on the flood risk of the home (see Table 1.1). This premium is higher than most homeowners would otherwise pay, but much lower than what homeowners were paying in risky areas (Flood Re, 2016). This is the mechanism via which Flood Re reduces premiums for high-risk homes.

We expect that Flood Re will act as a price ceiling, limiting the maximum price that higher-risk homes will pay for flood insurance. Figure 1.1 depicts a stylized version of this market. Insurance companies can attract low-risk customers by charging prices below the Flood Re rates. Thus the price ceiling should not bind for low-risk customers, and their prices will only change slightly, if at all. For high-risk customers, insurers that choose to pass the risk on to Flood Re will be able to offer lower prices than insurers that take on the risk themselves. Because the market for flood insurance is fairly competitive[3], the Flood

[3]According to a 2013 report by Her Majesty's Treasury, the UK insurance market is one of the most

Re rates (plus administrative costs to the insurer) should act as a price ceiling for high-risk homes.

In addition, the Water Act added a £180 million annual tax on flood insurers, so that Flood Re can charge low premiums despite adverse selection into the program. The exact impact this £180 million tax has on flood insurance prices is currently unknown. Importantly, we do *not* have data on flood insurance prices, so we cannot directly observe the level and method of pass-through. Most analyses of Flood Re assume that the tax has 100% pass-through to consumers, at a flat rate of £10.50 per policy[4] (Department for Environment Food & Rural Affairs, 2014; Edmonds, 2017). At the very least, it is reasonable to assume that flood insurance prices for low-risk homes are the same or slightly higher than they would be without the policy change. Flood insurance prices for high-risk homes should decrease, and should now be only slightly higher than the premiums that insurers pay to Flood Re. All evidence from industry and news reports confirms that prices dropped substantially for homeowners in high-risk areas (Edmonds, 2017; Association of British Insurers, 2017; Rudgard, 2016).

The goal of Flood Re is to lower flood insurance prices for high-risk homes in order to soften the blow of the transition to risk-based pricing. The policy is designed to be slowly phased out over 25 years, at which point it will expire. In order to estimate the efficacy of this policy, one might want to measure the difference between previous insurance prices and prices under Flood Re. Although we do not observe private insurance prices, we do not need to do so in order to evaluate the effect of the policy. According to the hedonic method developed by Rosen (1974), the value of the changes in insurance prices can be inferred using the change in the price of affected homes. We use a hedonic regression to estimate the treatment effect of Flood Re on eligible homes. In Section 1.5, we interpret the economic meaning of the treatment effect using estimates of the expected savings under Flood Re.

1.3 Data

Her Majesty's Land Registry publishes all residential property transactions that are sold for market value[5] in England and Wales since 1995. Currently, data are available through June 2018, for 23.5 million total observations. The data do not include property-level characteristics such as square feet and number of bedrooms, but they do include: full address; date of sale; sale price; whether the property is a detached house, a semi-detached house, or an apartment; whether the property is newly built; and whether the sale is permanent ("freehold") or for a period of time ranging from 40-120 years, at which point the property will be transferred to the original owner ("leasehold"). We use the sale price, adjusted for

competitive in the world (HM Treasury, 2013). As of 2017 there were approximately 65 flood insurance companies it the UK, 60 of which were participating in Flood Re (Edmonds, 2017).

[4]This number is calculated by dividing £180 million by the number of policies in the UK.

[5]Properties are defined as not sold at market value if they are transferred at low or zero price in events such as a divorce or inheritance.

inflation, as our main variable of interest. We winsorize the data to remove the top and bottom 2% of the data by inflation-adjusted price. Summary statistics for these data can be seen in the first ten rows of Table 3.1, and overall trends of price and quantity are presented in Figure 1.2.

For our analysis, we take advantage of the fact that home sales need to fulfill multiple criteria in order to be affected by Flood Re, as described in Table 1.2). These criteria are: (1) the sale must occur after April 4, 2016, when the policy takes effect, (2) the home must be in a high risk area in order for the price ceiling to bind (see Figure 1.1), (3) the home must be built before 2009 to qualify for Flood Re, (4) the home must be in a building with fewer than three residential units, and (5) the home must be owner-occupied. In our data, we observe or estimate criteria (1) - (3), but not criteria (4) and (5). Criterion (1), the date of sale, is reported in our sales data. Criterion (2), the risk level of the home, is estimated using flood risk data, as described in the next paragraph. Our sales dataset does not contain criterion (3), the build date of the home, only whether the property is newly built and being sold for the first time. By using the address to match properties and note repeat sales, we can find the age of most homes that are built and sold since 1995, since they will show up as "new" at some point in our data. Homes that never appear as "new" in the data could be built before 1995, but could also have a typo in the address or could have originally been sold for less than market value. For homes that never appear as "new" in the data, we try both dropping these observations and assuming that they are older than 1995.

In order to estimate criterion (2) for each home, we combine our sales data with data on risk of flooding. The Environment Agency publishes maps of flood risk in England, measured in probability of flooding per year. There are five levels available: less than 1 in 1000, between 1 in 1000 and 1 in 100, between 1 in 100 and 1 in 75, between 1 in 75 and 1 in 30, and greater than 1 in 30. This data is at the postcode level, and reports the number of residential and non-residential properties that fall into each category, within each postcode. A postcode is much smaller than a United States zipcode; the average postcode contains just 15 homes. In our main analysis, we designate postcodes as "high risk" if they contain at least one residential property at greater than 1 in 75 risk. Because this dataset does not include risk for properties in Wales, the 5% of properties that are in Wales are excluded from our analysis. We use these flood risk maps to estimate the risk of each home based on its postcode, and this "high risk" variable serves as the final binary variable for our triple-differences strategy.

Although we do not directly observe criteria (4) and (5), our data contain variables that are correlated with the number of units per building and owner occupancy. We observe whether a property is an apartment, so we try both including and excluding apartments from our analysis. We also observe whether a property is "freehold" or "leasehold." Freehold properties are analogous to the property ownership model in the United States. A leasehold is an institution in the UK that allows a person to "rent" a property for 40 - 120 years, with full control over upkeep and renovation. Leasehold properties can qualify for Flood Re if the leaseholder is responsible for buildings insurance. However, most leasehold properties are insured by the property owner. Our main analysis excludes leasehold properties and apartments.

The UK government also publishes shapefiles of actual flood events, which gives us an unusual level of precision in the ability to estimate whether a home has actually been flooded. These data provide the outline of actual floods in England since 1946, as verified in person or by aerial photography. These Recorded Flood Outlines also include the start and end date of the flood, and the flood source (fluvial, tidal, or coastal), along with the name of the local source. We intersect these shapefiles with the GPS coordinates of each home to find when (and if) the exterior of a home has been flooded, and how many times. Of course, exterior flooding will not lead to interior flooding in every case, but this is our best estimate of property-level flood history. There are very few actual flood event datasets worldwide with this level of precision; for example, in the United States, flood events are often measured through the existence of a Presidential Disaster Declaration, which is at the county level.

Finally, in order to increase the precision of our estimates, and better control for any systematic differences between groups in our triple-differences design, we include house-level characteristics from the dataset of Energy Performance Certificates (EPC). Since 2008, almost all homes in England are required to undergo an energy performance assessment when built, sold, or leased. The data collected during these assessments is published by The Ministry of Housing, Communities & Local Government. These data contain house-level variables that we include as controls in our analysis: total floor area; energy efficiency (numeric from 1 to 100); and built form (based on the number of walls, ceiling, and floor shared with other units). Because not all homes in our sales data have received an energy performance assessment, when we restrict our sample to homes that appear in the EPC data we have approximately 13.9 million observations.

1.4 Methods

The purpose of this study is to estimate whether the housing market responds to changes in flood insurance prices. Typically, changes in insurance prices are endogenous to changes in home values and other factors that affect home values, such as the updating of risk maps. The Water Act provides useful exogenous variation in the flood insurance industry that we can exploit to find the true effect of insurance prices on the housing market.

Looking at all home prices before and after the policy would provide a vast underestimate of the effect of the policy, as most homes will not be affected, and will miss the effect of general trends in the housing market. To account for these trends, we use a triple-differences design. In order to qualify for the subsidized insurance, homes must be built before January 1, 2009. In order to benefit from the subsidized insurance offerings, homes must be in high-risk areas. We use these two characteristics, in addition to whether the sale occurred before of after the policy implementation, to find the effects of subsidized insurance on homes that can benefit from the subsidy. The identifying assumption is that there are no other systematic differences that cause changes in home price trends between high risk and low risk homes, and older (pre-2009) versus newer (post-2009) homes, before and after the policy.

Formally, we run the following regression:

$$\begin{aligned} ln(p_{it}) = \beta_1 Post_{it} + \beta_2 Old_i + \beta_3 Risk_i + \beta_4 Post_{it} \cdot Old_i + \beta_5 Post_{it} \cdot Risk_i \\ + \beta_6 Old_i \cdot Risk_i + \beta_7 Post_{it} \cdot Old_i \cdot Risk_i + \gamma_s + \tau_t + X_i' \Delta + \varepsilon_{it} \end{aligned} \tag{1.1}$$

where $ln(p_{it})$ is the log of the sale price of a home, $Post_{it}$ is a dummy variable which equals one after the insurance policy has been implemented, Old_i is a dummy variable equal to one if the home was built before 2009, and $Risk_i$ is a dummy variable equal to one if the home is in a high flood risk area. γ_s are spatial fixed effects, typically at the district level, and τ_t are time fixed effects, typically the month of sample. X_i' represents property-level controls, which include home vintage fixed effects, the type of property, type of sale, whether the property is newly built, total floor area, energy efficiency, and built form. β_7 is the coeffiicient of interest, which measures the effect of flood insurance subsidization on housing prices for high-risk homes that qualify for the policy.

Summary statistics for our data are presented in Table 3.1. We divide the sample into the homes that are eligible and those that are ineligible for Flood Re, according to the criteria that we observe. The first three columns of this table show summary statistics for the full dataset of all homes in Her Majesty's Land Registry sold since 2012, and the final three columns present statistics for those homes that were matched with an EPC certificate. In both cases, there are no eligible sales for "new" homes, because to be eligible a sale must be for an older home (built before 2009) that is sold after April 4, 2016. Rows 3-10 of this table show statistically significant differences between eligible and ineligible homes, but the statistical significance is mechanical, driven by the fact that we have a large dataset and these are standard deviations for binary variables. These differences do not appear to be economically significant.

Theoretically, all homes sold since 2009 should have an EPC, so any unmatched properties likely had discrepancies in the address string between the two datasets, such as a typo in one version of the address. Many homes in the UK use a name rather than a house number[6], and some can use either, adding to the difficulty of matching. It appears that there are systematic differences between the matched and unmatched homes, whether by chance or due to the characteristics of homes with typos in the address, or the characteristics of homes that are referred to by name in one dataset and by number in another.

The vast majority (96.6%) of the homes in the data were built before 2009, and most (91%) of the sales occurred before Flood Re was in place. Because few homes (2.7%) are in the high risk category, there are only 40,032 sales of homes that should qualify and benefit from the policy, which is 0.2% of total sales. To help balance the sample in our main specification, we look only at sales that occur in 2012 or later, for a total of 5.3 million observations in the full data set, and 4.1 million observations that are matched with energy performance certificates. We are able to use the three eligibility factors to estimate the causal effect of Flood Re on home values, assuming that there are no other factors that only affect sales of high-risk older homes after April 4, 2016.

[6]Eg: "Bosley, Manor Road, Abbots Leigh, UK."

1.5 Capitalization

Our estimated coefficient of interest gives us the average percent change in the price of homes that qualify and benefit from this policy. In order to put this coefficient into context, we need to compare it to what we might expect to see. If home buyers and sellers pay no attention to flood insurance prices during the transaction process, we would expect β_7 to be close to zero and not statistically significant. If home buyers and sellers pay full attention to flood insurance prices, we should expect to see the price of qualifying homes to rise by the present discounted value (PDV) of the subsidy. This would reflect full capitalization, where the value of the asset of interest (in this case, a home) changes by the discounted expected decrease in future costs (in this case, flood insurance prices).

In order to calculate the present discounted value of the subsidy, we need an estimate of (1) the expected change in flood insurance prices, (2) the appropriate discount rate, and (3) the appropriate time horizon that consumers care about. We do not have data on flood insurance prices, so we use news articles and impact assessments to estimate the annual savings. A 2013 report by the Department for Food & Rural Affairs estimates that the median high-risk home will save \$465 annually under Flood Re (Department for Environment Food & Rural Affairs, 2013). An article in The Telegraph quotes a customer with a mid-range home saving £3,075 per year, and another interviews a homeowner whose annual premiums rose by £1,982 annually when her insurer flagged her home as at risk of flooding (Blackmore, 2015; Rudgard, 2016). We calculate the present discounted value (PDV) of the subsidy using these savings estimates and the standard formula for the PDV of an annuity:

$$PDV = C \cdot \frac{1 - (1 + i)^{-n}}{i} \tag{1.2}$$

where C is the annual savings from the subsidy, i is the discount rate, and n is the number of years over which the savings are realized. Flood Re is designed to be phased out after 25 years, so we use an n of 25 for this calculation. Figure 1.3 presents the results over a range of assumptions. In this figure, we allow the annual savings to range from £465 - £3,075, and allow the discount rate to range from 0% - 15%. The green ribbon represents the range of present discounted values of the subsidy, transformed into the implied percent increase in home value for a qualifying home under full capitalization. At a 5% discount rate, the PDV of the subsidy is approximately £6,500 - £43,000, depending on the assumed annual savings. For the homes in our data that are high risk, built before 2009, and sold after April 4, 2016, the mean sale price is £265,000. Thus, at a 5% discount rate, full capitalization would imply a 2.5% - 16.4% increase in the price of the mean eligible home in our data.

1.6 Results

The results from estimating equation 1.1 are presented in Table 1.4. All results include district and month of sample fixed effects, housing characteristic controls as described in section 1.3, and are two-way clustered by postcode and month of sample. Column (1) shows results from the housing sales data, without additional EPC variables. In row 6, we see that the homes sold after April 4, 2016 are worth slightly less than other homes. This is likely because of a concurrent policy change: the tax on residential property sales increased in the UK in April 2016. Thus, a simple pre- versus post- analysis would misestimate the effects of the policy. In rows 2 and 4, we see that high risk homes are generally decreasing in value during the study period. However, in row 1, we see that homes that are eligible for Flood Re increase in value by 8.3% after the policy takes effect.

Column 2 of Table 1.4 shows results from the same regression on the subset of properties that were matched to an EPC. We see that the estimate for our coefficient of interest (row 1) rises to 12.5%, which is a signal that this subsample of data has slightly different characteristics than the data used for column 1. In column 3, we add the EPC controls (size, built form, and energy efficiency). Our coefficient of interest drops to 9.8%. Because column 3 contains important house-level controls, this is our preferred specification.

One omitted variable that may be biasing our results is the effect of a flood itself. Homes in high risk areas are clearly more likely to have been flooded. A flood can damage a home, and it can provide a signal of risk to potential home buyers, but it can also result in an insurance payout and subsequent renovation. In Chapter 3, I use the recorded flood outlines data provided by the UK government to examine the effect of a flood on the sale price of an affected home. The precision of the flood outline maps allows for a regression discontinuity design, comparing homes just inside versus just outside of a flood. Looking at sales within 1, 5, or 10 years after a flood, I find that homes just inside a flood boundary are worth more than homes that were just outside a flood boundary. It appears that the impact of a renovation is stronger than the impacts of the risk signal or any remaining damage.

In order to account for the fact that actual flood history is correlated with one of our eligibility variables, we run regressions that include whether a home has been flooded in the 10 years prior to the sale. We calculate the flooding history of the home using the recorded flood outline maps. In Table 1.5, we interact this binary variable with our other eligibility variables. We see that our coefficient of interest (row 1) only changes slightly from Table 1.4.

1.7 Robustness Checks

We include robustness checks to see whether our results are sensitive to any of our specification choices. First, in Table 1.6, we try including all apartments in our regressions. In our main specification, we excluded apartments because homes in a building with more than three residential units do not qualify for Flood Re. We see that our coefficient of interest, presented in the first row of the table, does not change for all homes (column 1), but rises slightly for homes that are matched to an EPC (columns 2 and 3). The coefficient on high risk homes (row 5) is now statistically significant for all three specifications, implying that apartments in flood-prone postcodes are relatively more valuable than other types of flood-prone property, perhaps because floods are unlikely to reach upper-level units.

Next, we test whether the years we include in the pre-period affect our results. Table 1.7 displays results for regressions with first fewer pre-period years than in our main regression (column 1), and then an increasing number of pre-period years (columns 2 through 5). We see that, as we include more and more data before Flood Re was in place, the size and significance of our coefficient of interest decrease. However, our estimate remains statistically significant at the 5% level even when we include all years in our data (column 5).

1.8 Distributional Effects

The Flood Re subsidy is available regardless of the value of the property or homeowner income. The costs of Flood Re are partly paid for by premiums, but of course these premiums are not based on risk and so cannot cover the full costs of damage to insured homes, due to adverse selection into the program. The remaining costs are paid for by a £180 million annual tax on insurers. This tax, which averages £10.50 per household, is generally expected to be passed on to consumers (Edmonds, 2017).

In the UK, more expensive homes tend to face higher risk, particularly riverfront and coastal properties. For this reason, Flood Re has the potential to be quite regressive. Before passing the Water Act, the UK government ran an impact assessment which estimated that the majority of the subsidy would flow to the poorest households, but this was partly due to the assumption that the most expensive homes would not be eligible for Flood Re (Department for Environment Food & Rural Affairs, 2014). The Water Bill proposal was later changed to include all homes. We conduct our own analysis to examine the distributional effects of this policy.

First, we examine whether the policy targets certain income groups via its eligibility criteria. Figure 1.4 displays our calculations of the proportion of homes that should qualify for and benefit from Flood Re, by income price decile. The income deciles are drawn from the 2016 release of the Small Area Income Estimates dataset from the Office of National Statistics. These data are at the middle layer super output area (MSOA) geographic level, which on average have a population of 7,200. We match the income estimates to homes by address, and then find the deciles of income in our dataset. These deciles by construction contain less variation than the actual population, due to averaging at the MSOA level. We classify a property as “eligible” if it is in a postcode with risk equal to or greater than 1 in 75, if it is built before 2009, and if the property is not an apartment. Figure 1.4 shows the mean value and 95% confidence interval for this eligibility variable. We bound the lower half of the confidence interval at zero. As can be seen in the figure, although there is some variation across house price deciles, the vast majority of the variation occurs within deciles; that is, if this policy is progressive or regressive, it is unlikely to be due to the eligibility criteria that we observe.

Next, as seen in Figure 1.5 we estimate the value of the subsidy to homes that qualify for Flood Re. For this calculation, because we do not have data on actual insurance prices, we must estimate the price homeowners would pay for insurance without Flood Re. We assume actuarily fair pricing, and find the annual insurance cost by multiplying the value of the home by the average damage from a flood (which is 15% of the value of the home), and by the home's annual probability of flooding. To find the price that eligible homeowners pay after Flood Re, we estimate their council tax band and find the corresponding insurer premium under Flood Re. Because we do not have information on how the tax on insurers is passed on to consumers, we follow the procedure of other Flood Re analysts and assume a flat tax of £10.50 per household. We subtract the estimated premium under Flood Re and the £10.50 flat tax from the estimated premium without Flood Re to find the estimated value of the subsidy. Figure 1.5 displays these estimates. We find that the value of the subsidy does vary by income, with wealthier deciles benefiting more. Combining Figure 1.4 with Figure 1.5, it can be seen that most of the difference between deciles comes from the change in the value of a home, and not from differences in flood risk.

Finally, we estimate the value of the subsidy for all homes, including homes that do not qualify for Flood Re, such as apartments or newer homes. For homes that do qualify for Flood Re, we use the value of the subsidy as estimated for Figure 1.5, described in the preceding paragraph. For homes that do not qualify for Flood Re, we estimate the value of the subsidy to be -£10.50, because, as is shown in Figure 1.1, we do not expect the policy to raise prices for non-qualifying homes. The only impact to non-qualifying homes should be via the new tax on insurers, which averages to £10.50 per household. Figure 1.6 shows these results. We restrict the lower bound of the 95% confidence interval to -£10.5. The mean of these estimates does not vary much by decile, because most homes do not receive the subsidy. However, the standard deviation rises as the income decile rises, because the maximum value of the estimated subsidy depends on the house price. Overall, we find that, although diferent income deciles qualify for Flood Re at similar rates, the majority of program funds will flow to the owners of more expensive homes, even after accounting for the fact that they will pay higher premiums. This is especially problematic because it is unclear how the tax on insurers will be passed on to policyholders. If the tax is not applied in a way that is proportional to home value, owners of inexpensive homes could be subsidizing ownership of expensive homes in high risk areas.

1.9 Conclusion

There are few other papers that examine whether consumers respond to flood insurance prices, because there are few instances of exogenous changes to the price of insurance. We take advantage of a policy change in the UK that effectively subsidizes insurance for a subset of homes. Using a triple differences design, we find that homes that qualify for the subsidy increase in value by 9.8%. This is consistent with full capitalization, and so it appears that home buyers are actually quite sophisticated when incorporating this add-on cost. This is reassuring, as housing markets are not consistently responsive to other signals of flood risk. Our paper suggests that risk-based flood insurance pricing could work as a strategy for flood adaptation under climate change.

We also study the distributional impacts of the policy change, which subsidizes flood insurance for high risk homes. We find that flood risk is not correlated with the mean income of an area; there is much more risk variation within income deciles than between deciles. However, because wealthier people tend to have more expensive homes, most of the subsidy will flow to higher income groups. This is true despite the fact that homeowner payments into the program are based on home value. Depending on the method of pass-through, the tax on insurers may add to or reduce the regressive nature of the policy. Overall, although the UK government was transparent about the choice to help high-risk areas at the expense of low-risk areas, the policy design also disproportionately helps the wealthy, at the expense of the poor.

Our analysis underscores the need for further research on attentiveness to disaster insurance. It is not clear how valid these results are for other countries such as the United States with slightly different processes for purchasing a home. In particular, in the United Kingdom, flood insurance is required for all homes to obtain a mortgage, whereas in the United States insurance is only a requirement for high-risk homes. It is also not clear whether the effectiveness of a price signal varies based on the salience of flooding. Floods in the UK tend to receive extensive news coverage, particularly in the years leading up to the policy change that we study. More research like the paper by Gibson et al. (2017) is needed to understand the how flood salience impacts consumer price response. There is evidence that flood salience matters for the housing market (Atreya et al., 2013; Bin and Landry, 2013a; McCoy and Walsh, 2018) and insurance takeup (Gallagher, 2014), so it is quite possible that salience also matters for home buyer response to insurance prices.

Table 1.1: FloodRe Premiums in England, 2016

Council Tax Band	A	B	C	D	E	F	G	H
Buildings	£132	£132	£148	£168	£199	£260	£334	£800
Contents	£78	£78	£98	£108	£131	£148	£206	£400
Combined	£210	£210	£246	£276	£330	£408	£540	£1200

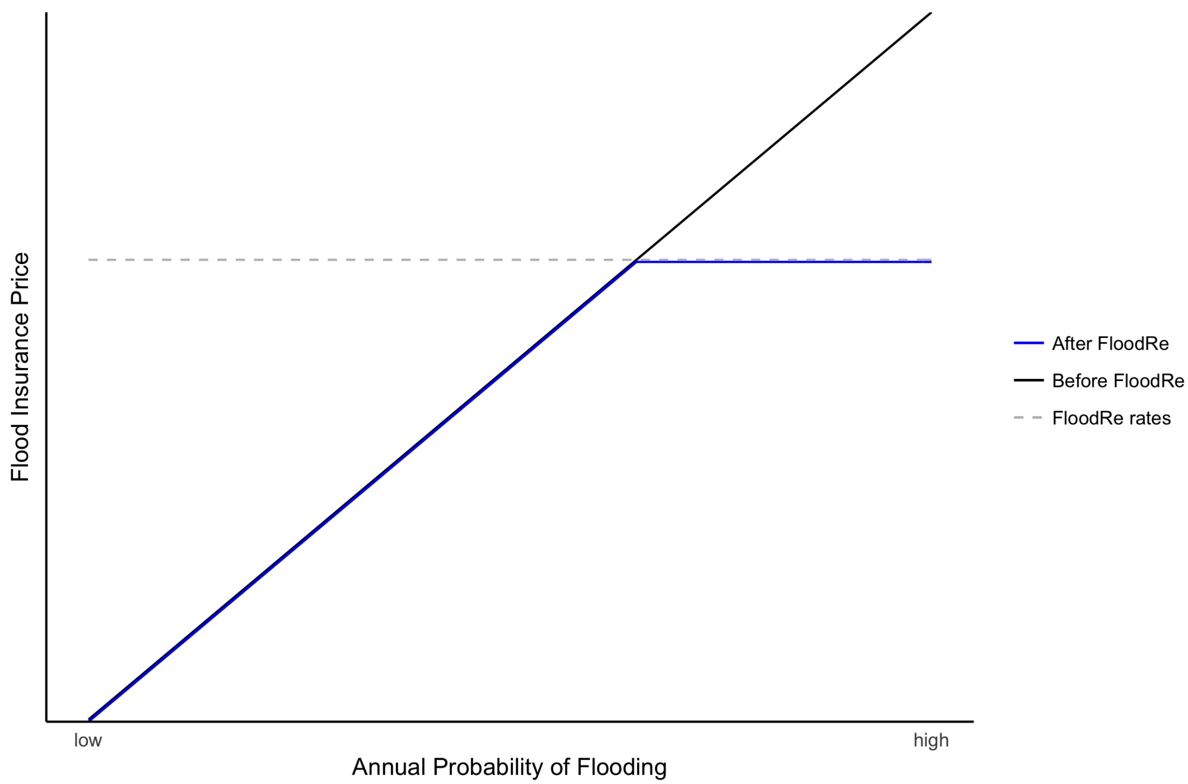
Notes: This table lists the premiums that insurers must pay to FloodRe if the insurer chooses to reinsure the flood risk of the home via FloodRe (FloodRe Transition Plan, 2016). The prices vary by council tax band, which is based on the estimated value the property would have sold for on April 1, 1991. The homeowner and insurer can agree to insure the property structure only ("buildings"), the internal possessions only ("contents"), or both ("combined"). Home buyers are required to have buildings insurance in order to be approved for a mortgage.

Table 1.2: Criteria used in Analysis

	Criteria to benefit from FloodRe		Observed
(1)	Sale date	After April 4, 2016	Y
(2)	Risk level	Greater than 1 in 75	Y
	Criteria to qualify for FloodRe		**Observed**
(3)	Vintage	Built before Jan 1, 2009	Y
(4)	Property type	Building with 1-3 residential units (no large apartment buildings)	N
(5)	Rental?	Owner-occupied only	N

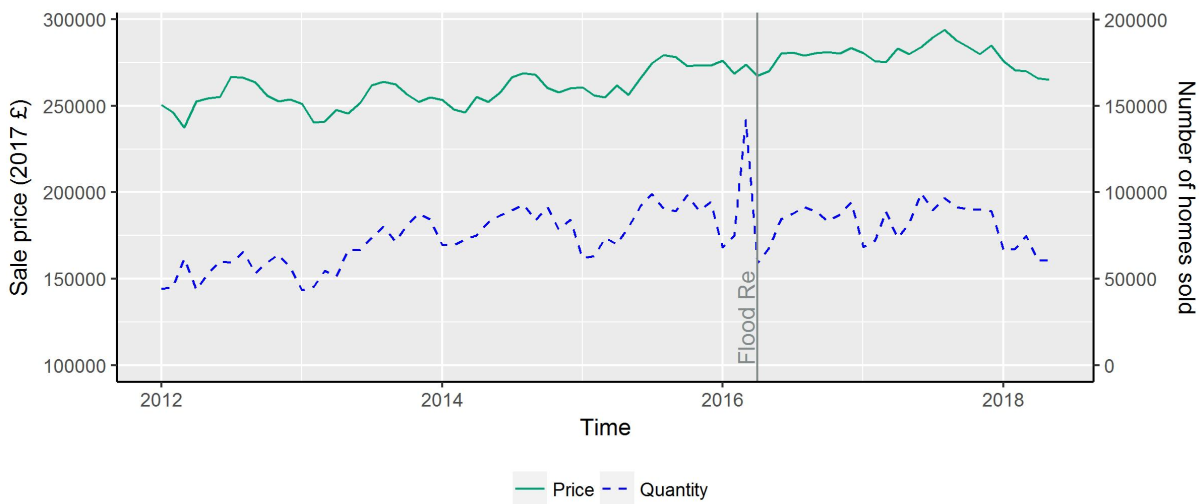
Notes: This table lists the criteria required for a home sale to be what we call "eligible" for FloodRe (defined as both qualifying for the program and being able to benefit from it). FloodRe became available on April 4, 2016. We do observe the date of sale, so we designate all sales after April 4, 2016 as able to benefit for this category. FloodRe should not affect the insurance prices of low-risk homes (see Figure 1.1). We designate homes in a postcode with at least 1 in 75 risk of flooding per year as able to benefit based on flood risk. Only homes built before 2009 can have their risk passed on to FloodRe. We do not observe the exact build date of the home, but we designate any home sold as "new" after December 31, 2008 as unable to qualify for FloodRe due to vintage. Homes in an apartment building with more than three residential units do not qualify for FloodRe. We do not observe the number of units per building, so we run our specification both including and excluding all apartments. We do not observe whether a property is owner-occupied.

Figure 1.1: Insurance Prices as a Function of Risk, before and after Flood Re



Notes: This figure is a stylized representation of the price of flood insurance in England as a function of flood risk, to emphasize that low-risk homes will not be affected by the policy. The blue line represents pricing before Flood Re came into place on April 4, 2016, and the black line (covered by the blue line at lower risk levels) represents pricing after Flood Re is in effect. The dashed horizontal line represents the price ceiling set by Flood Re. Under Flood Re, insurers are allowed to charge any price they wish, but they must send a specified amount to Flood Re if they are to be reinsured for the risk on that particular home. Because Flood Re prices are not based on the probability of flooding, insurers that pass risk of high-risk homes on to Flood Re can charge lower prices than insurers that take on the risk of high-risk homes. Because the home insurance market is quite competitive, we suspect that insurers will charge approximately the Flood Re payments plus administrative costs, and so the price ceiling will bind at that level. This figure abstracts away from the fact that, in reality, the price ceiling differs by the council tax band of the home (which is based on the value of the home). Homes in England will face one of seven price ceilings, depending on their council tax bands.

Figure 1.2: Property Sale Price and Quantity, 2012-2018



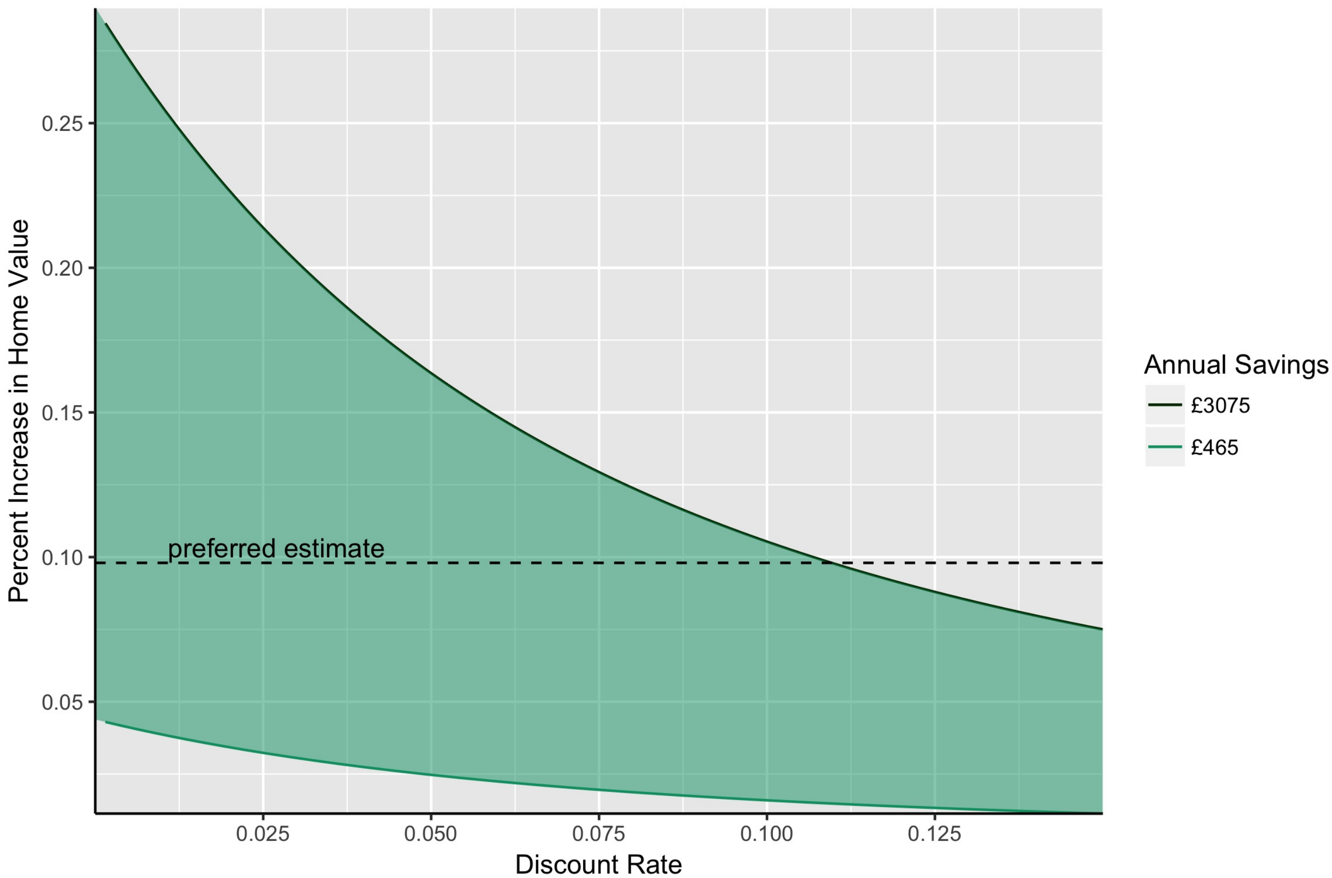
Notes: This figure presents the overall trends for price and quantity in the years immediately before and after the implementation of Flood Re. The solid green line depicts the mean price of all homes sold each month, with units on the left vertical axis. The dashed blue line depicts the total number of housing units sold each month, with units on the right vertical axis. Property owners had access to subsidized insurance through Flood Re starting on April 4, 2016, as depicted by the grey vertical line labeled "Flood Re." The spike in quantity right before April 2016 is due to the fact that property sales taxes (called "stamp duty land tax") rose on April 1, 2016.

Table 1.3: Summary Statistics of Properties Sold

Property Characteristics	**Full Sample** Eligible	Ineligible	Difference in Means	**Matched with EPC** Eligible	Ineligible	Difference in Means
Price (2017 £)	257,451.51	244,532.11	12,919.41	242,255.10	229,152.39	13,102.71
	(150,226.00)	(143,542.08)	(688.32)	(125,295.70)	(119,150.92)	(733.14)
Vintage (year)	2003.94	2008.26	-4.33*	2004.22	2007.36	-3.14*
	(5.52)	(7.03)	(0.06)	(5.38)	(6.78)	(0.08)
New Sale	0.00	0.10	-0.10***	0.00	0.07	-0.07***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Mortgage Type						
Freehold	0.77	0.75	0.03***	0.82	0.82	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Leasehold	0.23	0.25	-0.03***	0.18	0.18	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Property Type						
Detached	0.24	0.23	0.02***	0.25	0.24	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Semi-Detached	0.22	0.27	-0.04***	0.25	0.31	-0.06***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Terraced	0.30	0.29	0.01***	0.35	0.33	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Flat	0.18	0.20	-0.02***	0.13	0.12	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Other	0.05	0.02	0.03***	0.02	0.00	0.02***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Eligibility Variables						
Built before 2009	1.00	0.90	0.10***	1.00	0.93	0.07***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
High Risk	1.00	0.02	0.98***	1.00	0.02	0.98***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Sold after FloodRe	1.00	0.35	0.65***	1.00	0.28	0.72***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
EPC Variables						
Floor Area (m^2)	–	–	–	90.99	89.90	1.10
	–	–	–	(42.75)	(39.23)	(0.25)
Energy Efficiency	–	–	–	59.49	61.86	-2.38*
	–	–	–	(15.13)	(13.65)	(0.09)
Number of Observations	48,032	5,249,208	–	29,412	3,826,204	–

Notes: This table compares property characteristics between eligible and non-eligible home sales since 2012, by data set. We define a home sale as "eligible" if it meets all of three characteristics: (1) it must be built before 2009 to qualify for FloodRe, (2) it must be in a high risk area in order for the price ceiling to bind (the policy works as a price ceiling; see Figure 1.1), and (3) the sale must occur after April 4, 2016, when the policy takes effect. The left three columns, under "Full Sample," include all residential property sales in England from January 1995 until June 2018, after winsorization. The right three columns, under "Matched with EPC," contain only those properties that successfully merged with the Energy Performance Certificates (EPC) data. Homes that have been sold or renovated since before 2008 are not in the EPC data. The rows for mortgage type, property type, and eligibility variables, as well as for new sales, describe the proportion of homes with that characteristic. The other rows describe continuous variables. Standard deviations are in parentheses. Significance: *p< 0.1; **p< 0.05; ***p< 0.01.

Figure 1.3: Implied Percent Change in Home Value under Full Capitalization



Notes: This figure displays the percent increase in home value that we would expect for the mean home affected by Flood Re, under a range of discount rates. We do not have insurance prices, so we estimate the annual savings for an affected home based on impact assessments and news articles (Defra 2013; Defra 2014; Blackmore 2015; Rudgard 2016), and find a range of savings from £465 - £3075 per year. Our preferred estimate is $\beta_7 = 9.8\%$ (see Table 1.4), which is represented by the dashed line in the figure above. The green ribbon represents the present discounted value of the subsidy, over a range of discount rates and savings amounts. The present discounted value is divided by the mean price of an affected home in our sample (£265,000), making the y-axis the value of the subsidy as a percent of the home value, which is the implied percent increase in the value of a home under full capitalization.

	Dependent variable: Log Price (2017 £)		
	All Homes	Homes with EPC	Homes with EPC
	(1)	(2)	(3)
Built pre 2009× High Risk × Post FloodRe	0.083*	0.125***	0.098**
	(0.048)	(0.047)	(0.043)
Built pre 2009× High Risk	−0.022	−0.028	−0.021*
	(0.016)	(0.017)	(0.012)
Built pre 2009× Post FloodRe	0.064***	0.008	0.004
	(0.021)	(0.012)	(0.006)
High Risk× Post FloodRe	−0.084*	−0.118**	−0.102**
	(0.047)	(0.047)	(0.043)
High Risk	0.024	0.025	0.038***
	(0.015)	(0.017)	(0.012)
Post FloodRe	−0.036**	0.003	−0.005
	(0.017)	(0.010)	(0.005)
Built pre 2009	–	–	–
	–	–	–
Observations	2,189,453	1,823,869	1,823,869
Property FE	Yes	Yes	Yes
EPC Controls	No	No	Yes
Month-of-sample FE	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Clustering	Postcode & Month of sample	Postcode & Month of sample	Postcode & Month of sample
R^2	0.536	0.563	0.831
Adjusted R^2	0.536	0.562	0.827

Notes: This table shows the results of a triple-differences regression of price on eligibility characteristics. The regression in the first column uses all sales after 2012, whereas the regressions shown in the second and third columns only use sales after 2012 that were matched to an Energy Performance Certificate (EPC), which provides more property-level characteristics. For homes matched to EPCs, we can also control for the size and energy efficiency of the home. Column 2 shows results for the subset of homes that are matched to an EPC, but the regression in column 2 does not contain the EPC controls. Thus, the difference from column 1 to 2 is due to a difference in sample characteristics. Column 3 displays results for homes that are matched to an EPC when the size and efficiency of a home are included as controls. The first row of this table displays the coefficient of interest, the estimated percent increase in sale price for homes that are eligible for FloodRe. All regressions include fixed effects for the type of mortgage (freehold or leasehold), the estimated property vintage, whether the property is brand new (newly built and never sold before), the type of property (detached, semi-detached, terraced, flat, or other), the property's district, and the month of sample of the sale. Standard errors are in parentheses, and are two-way clustered by postcode and month of sample. Significance: $^{*}p< 0.1$; $^{**}p< 0.05$; $^{***}p< 0.01$.

Table 1.4: Effects of FloodRe on Housing Prices (Main Specifications)

	Dependent variable: Log Price (2017 £)		
	All Homes	Homes with EPC	Homes with EPC
	(1)	(2)	(3)
Built pre 2009× High Risk × Post FloodRe	0.120**	0.134***	0.097**
	(0.047)	(0.045)	(0.042)
Built pre 2009× High Risk Post FloodRe× Flooded	−0.087	−0.609***	−0.158**
	(0.257)	(0.111)	(0.066)
High Risk× Post FloodRe × Flooded	0.161	0.677***	0.180***
	(0.258)	(0.108)	(0.063)
High Risk× Post FloodRe	−0.118**	−0.127***	−0.101**
	(0.046)	(0.045)	(0.042)
Built pre 2009× Post FloodRe	0.045***	0.008	0.004
	(0.013)	(0.013)	(0.006)
Built pre 2009× High Risk	−0.025	−0.027	−0.016
	(0.016)	(0.018)	(0.012)
Post FloodRe	−0.023**	0.004	−0.005
	(0.011)	(0.011)	(0.005)
High Risk	0.031**	0.027	0.034***
	(0.016)	(0.018)	(0.012)
Flooded	−0.016	0.001	0.009
	(0.052)	(0.057)	(0.032)
Observations	2,103,423	1,816,983	1,816,983
Property FE	Yes	Yes	Yes
EPC Controls	No	No	Yes
Month-of-sample FE	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Clustering	Postcode & Month of sample	Postcode & Month of sample	Postcode & Month of sample
R^2	0.533	0.567	0.832
Adjusted R^2	0.532	0.566	0.828

Notes: This table shows the results of a triple-differences regression of price on eligibility characteristics, when we add an additional interaction term: whether the home was flooded in the 10 years prior to being sold.The regression in the first column uses all sales after 2012, whereas the regressions shown in the second and third columns only use sales after 2012 that were matched to an Energy Performance Certificate (EPC), which provides more property-level characteristics. For homes matched to EPCs, we can also control for the size and energy efficiency of the home.The first row of this table displays the coefficient of interest, the estimated percent increase in sale price for homes that are eligible for FloodRe. All regressions include fixed effects for the type of mortgage (freehold or leasehold), the estimated property vintage, whether the property is brand new (newly built and never sold before), the type of property (detached, semi-detached, terraced, flat, or other), the property's district, and the month of sample of the sale. Standard errors are in parentheses, and are two-way clustered by postcode and month of sample. Significance: *p< 0.1; **p< 0.05; ***p< 0.01.

Table 1.5: Effects of FloodRe on Housing Prices, Flooding Controls

	Dependent variable: Log Price (2017 £)		
	All Homes	Homes with EPC	Homes with EPC
	(1)	(2)	(3)
Built pre 2009× High Risk × Post FloodRe	0.083**	0.170***	0.140***
	(0.039)	(0.048)	(0.037)
Built pre 2009× High Risk	−0.023	−0.035*	−0.031**
	(0.017)	(0.019)	(0.013)
Built pre 2009× Post FloodRe	0.023*	0.002	−0.004
	(0.012)	(0.014)	(0.006)
High Risk× Post FloodRe	−0.088**	−0.165***	−0.144***
	(0.039)	(0.048)	(0.037)
High Risk	0.047***	0.041**	0.051***
	(0.017)	(0.019)	(0.013)
Post FloodRe	−0.015	0.003	−0.001
	(0.010)	(0.011)	(0.005)
Built pre 2009	–	–	–
	–	–	–
Observations	2,927,341	2,130,023	2,130,023
Property FE	Yes	Yes	Yes
EPC Controls	No	No	Yes
Month-of-sample FE	Yes	Yes	Yes
District FE	Yes	Yes	Yes
Clustering	Postcode & Month of sample	Postcode & Month of sample	Postcode & Month of sample
R^2	0.578	0.585	0.824
Adjusted R^2	0.578	0.585	0.820

Notes: This table shows the results of a triple-differences regression of price on eligibility characteristics. The regression in the first column uses all sales after 2012, whereas the regressions shown in the second and third columns only use sales after 2012 that were matched to an Energy Performance Certificate (EPC), which provides more property-level characteristics. For homes matched to EPCs, we can also control for the size and energy efficiency of the home. Column 2 shows results for the subset of homes that are matched to an EPC, but the regression in column 2 does not contain the EPC controls. Thus, the difference from column 1 to 2 is due to a difference in sample characteristics. Column 3 displays results for homes that are matched to an EPC when the size and efficiency of a home are included as controls. The first row of this table displays the coefficient of interest, the estimated percent increase in sale price for homes that are eligible for FloodRe. All regressions include fixed effects for the type of mortgage (freehold or leasehold), the estimated property vintage, whether the property is brand new (newly built and never sold before), the type of property (detached, semi-detached, terraced, flat, or other), the property's district, and the month of sample of the sale. Standard errors are in parentheses, and are two-way clustered by postcode and month of sample. Significance: *p< 0.1; **p< 0.05; ***p< 0.01.

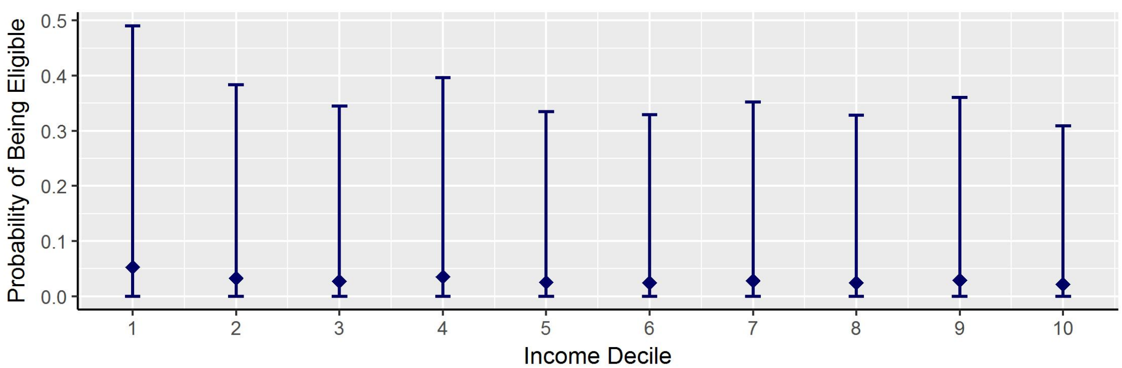
Table 1.6: Effects of FloodRe on Housing Prices, Apartments Included

	Dependent variable: Log Price (2017 £)				
	Since 2014	Since 2010	Since 2005	Since 2000	Since 1995
	(1)	(2)	(3)	(4)	(5)
Built pre 2009× High Risk× Post Flood Re	0.148*** (0.056)	0.109*** (0.037)	0.087** (0.035)	0.086** (0.037)	0.084** (0.038)
Built pre 2009 ×High Risk	−0.094*** (0.023)	−0.029** (0.012)	−0.017 (0.010)	−0.012 (0.011)	−0.016 (0.012)
Built pre 2009× Post Flood Re	0.014** (0.006)	0.018*** (0.005)	0.025*** (0.006)	0.025*** (0.006)	0.026*** (0.006)
High Risk× Post Flood Re	−0.143*** (0.055)	−0.116*** (0.037)	−0.098*** (0.035)	−0.102*** (0.037)	−0.099*** (0.038)
High Risk	0.117*** (0.022)	0.053*** (0.012)	0.042*** (0.010)	0.039*** (0.011)	0.042*** (0.011)
Post Flood Re	−0.008 (0.006)	−0.018*** (0.004)	−0.022*** (0.005)	−0.021*** (0.005)	−0.022*** (0.005)
Built pre 2009	– –	– –	0.090*** (0.006)	0.118*** (0.005)	0.130*** (0.004)
Observations	2,606,160	4,675,892	7,348,139	10,373,555	12,875,636
Property FE	Yes	Yes	Yes	Yes	Yes
EPC Controls	Yes	Yes	Yes	Yes	Yes
Month-of-sample FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes
Clustering	Postcode & Month of sample	Postcode & Month of sample	Postcode & Month of sample	Postcode & Month of sample	Postcode Month of sample
R^2	0.712	0.822	0.796	0.791	0.806
Adjusted R^2	0.712	0.819	0.794	0.788	0.805

Notes: This table shows the results of a triple-differences regression of price on eligibility characteristics, with different pre-periods in each column. For example, the regression used for column 1 includes all sales starting from January 1,2014. The first row of this table displays the coefficient of interest, the estimated percent increase in sale price for homes that are eligible for FloodRe. All regressions include fixed effects for the type of mortgage (freehold or leasehold), the estimated property vintage, whether the property is brand new (newly built and never sold before), the type of property (detached, semi-detached, terraced, flat, or other), the property's district, and the month of sample of the sale. Standard errors are in parentheses, and are two-way clustered by postcode and month of sample. Significance: *p< 0.1; **p< 0.05; ***p< 0.01.

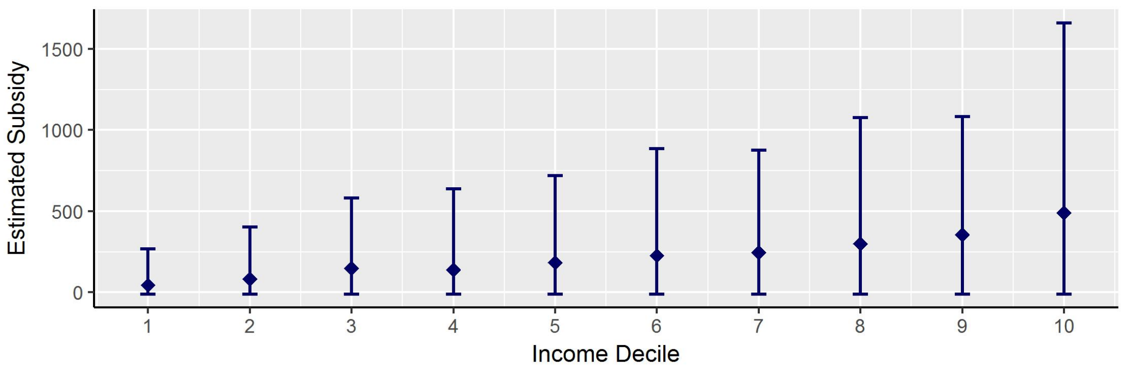
Table 1.7: Effects of FloodRe on Housing Prices, Different Pre-Periods

Figure 1.4: Percent of homes that are eligible for Flood Re, by income decile



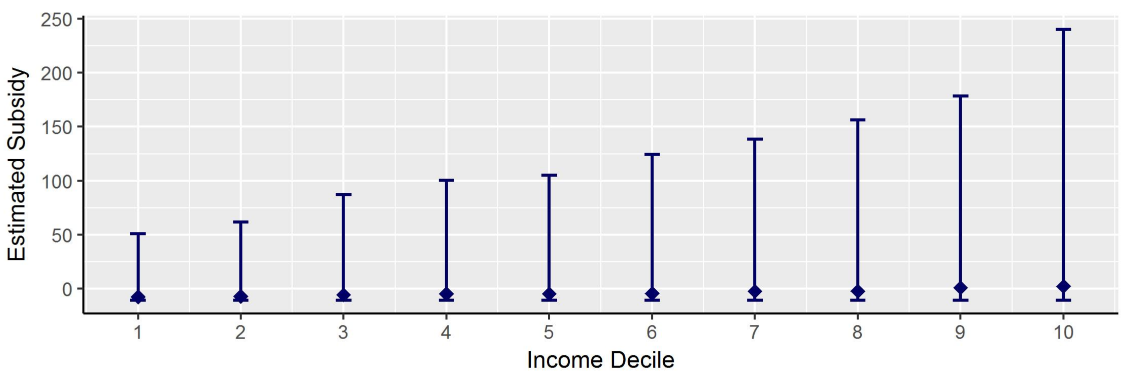
Notes: The figure above depicts the mean proportion of homes that are eligible for Flood Re, along with the 95% confidence interval of this proportion, by income decile. For this figure, we define a home as "eligible" if is built before 2009 and is in a postcode with greater than 1 in 75 annual risk of flooding. The income deciles are drawn from the 2016 release of the Small Area Income Estimates dataset from the Office of National Statistics. These data are at the middle layer super output area (MSOA) geographic level, which on average have a population of 7,200. We match the income estimates to homes by address, and then find the deciles of income in our dataset. These deciles by construction contain less variation than the actual population, due to averaging at the MSOA level.

Figure 1.5: Estimated value of Flood Re subsidy for eligible homes, by income decile



Notes: The figure above depicts the mean estimated annual subsidy for homes that are eligible for Flood Re, along with the 95% confidence interval of this measure, by income decile. For this figure, we define a home as "eligible" if is built before 2009 and is in a postcode with greater than 1 in 75 annual risk of flooding. We estimate the annual subsidy by subtracting the estimated current flood insurance price from the estimated flood insurance price before Flood Re. The current flood insurance price is found using the schedule of Flood Re insurer premiums by council tax band. The price of insurance before Flood Re is found by multiplying the lower bound of the risk level for the property (eg: 1 in 75 or above is estimated to be 1 in 75) by the value of the home and the average percent of a home value that is lost in a flood (15%). We truncate the lower bound of the subsidy standard deviation at -£10.50, as homes are not expected to lose money from Flood Re except via the tax on insurers. The income deciles are drawn from the 2016 release of the Small Area Income Estimates dataset from the Office of National Statistics. These data are at the middle layer super output area (MSOA) geographic level, which on average have a population of 7,200. We match the income estimates to homes by address, and then find the deciles of income in our dataset.

Figure 1.6: Estimated value of Flood Re subsidy for all homes, by income decile



Notes: The figure above depicts the mean estimated annual subsidy all homes in England, along with the 95% confidence interval of this measure, by income decile. We estimate the annual subsidy by subtracting the estimated current flood insurance price from the estimated flood insurance price before Flood Re. The current flood insurance price is found using the schedule of Flood Re insurer premiums by council tax band. The price of insurance before Flood Re is found by multiplying the lower bound of the risk level for the property (eg: 1 in 75 or above is estimated to be 1 in 75) by the value of the home and the average percent of a home value that is lost in a flood (15%). We truncate the lower bound of the subsidy standard deviation at -£10.50, as homes are not expected to lose money from Flood Re except via the tax on insurers. The income deciles are drawn from the 2016 release of the Small Area Income Estimates dataset from the Office of National Statistics. These data are at the middle layer super output area (MSOA) geographic level, which on average have a population of 7,200. We match the income estimates to homes by address, and then find the deciles of income in our dataset.

Chapter 2

Endogenous Rainfall: The Effect of Forest Cover on Local Rainfall in Brazil

2.1 Introduction

Though economists have known for decades that humans can have an impact on global climate, we have not yet come to understand all of the ways that human activity can affect local weather. Precipitation in particular is a variable that is assumed to be exogenous, but in truth depends largely on air moisture and temperature, both of which can be affected by local anthropogenic forces. Beginning with Miguel et al. (2004), many economists have taken advantage of the presumed exogeneity of rainfall to use deviation from mean precipitation as an instrument in numerous regressions where it seems unlikely to violate the exclusion restriction. Rainfall has been used to examine the effects of "exogenous" variation in resource availability on child health (Maccini and Yang, 2009), remittances (Yang and Choi, 2007), economic growth (Barrios et al., 2010), conflict (Miguel et al., 2004; Bohlken and Sergenti, 2010; Fjelde and von Uexkull, 2012; Maystadt et al., 2013), and political regime changes (De Figueiredo et al., 2010; Chaney, 2013). However, there is significant evidence from the scientific literature that human activity, including deforestation, can affect local precipitation. A forested nation experiencing changes in economic growth, civil conflict, or natural resource policy might see simultaneous changes in precipitation. This previously unexamined exclusion restriction violation could be problematic for research in areas with large amounts of moist forest cover, which includes Central Africa, Southeast Asia, Central America, and parts of South America, all of which are popular areas for development economics research.

For decades, scientists have explored the ways in which agriculture and rural land use can affect the hydrological cycle, but this has typically been done via simulations and modelling exercises (Shukla and Mintz, 1982; Henderson-Sellers and Gornitz, 1984; Lean and Warrilow, 1989; Eltahir and Bras, 1994; Xue and Shukla, 1996; Wang et al., 2000; Taylor et al., 2002;

Ramos da Silva et al., 2008; Garcia-Carreras and Parker, 2011). The mechanism for endogenous rainfall is fairly well understood; forests release more moisture than other vegetation types, and also reduce albedo and solar heating (Makarieva et al., 2006; Negri et al., 2004; Sheil and Murdiyarso, 2009). This should increase precipitation in two ways: by increasing air pressure, and by creating atmospheric pressure differences that pull in moisture from elsewhere. Although this has been shown using methods common to the atmospheric science literature, these results are not well known or understood by economists, partly because of a large methodological gap. Economists, focused on causality within social science data, have little experience with evaluating and understanding atmospheric science models. The paper that comes closest to an economist's preferred analysis uses real world data and finds that, indeed, air that has passed over forest cover releases more precipitation than air that has passed over areas with little forest (Spracklen et al., 2012). However, the work done by Spracklen et al. (2012) is far from a causally identified econometric analysis. In this paper, I build on the contributions of Spracklen et al. (2012) and other atmospheric scientists to estimate the effect of deforestation on rainfall in a way that is comprehensible and convincing to economists.

I use an instrumental variables analysis to examine the effects of exogenous variation in forest cover. Many papers within the atmospheric science literature, including Spracklen et al. (2012), do not take into account potentially important factors such as the impact that rainfall can have on forest cover, as shown in Saatchi et al. (2013) and Samanta et al. (2012b). It is clear that there is correlation between precipitation and upwind forest cover, but one should not make a causal conclusion. In addition to seasonal changes in forest greenness, forests can be affected by human activity in ways that are correlated with rainfall. For example, areas with an ideal combination of rainfall and soil type for a given crop should experience relatively more conversion to agriculture. If the profitability of different crops changes over time, spatial and seasonal fixed effects will not capture the effect of precipitation on forest cover. I measure the correlation between precipitation at weather stations and upwind vegetative cover, using spatial and temporal fixed effects to remove some endogeneity concerns. Additionally, IMF commodity price data is used to instrument for changes in vegetative cover. This is done to correct for issues of reverse causality (rainfall causing changes in forest cover), as well as control for any other factors such as land or ocean temperature which may be affecting both forest cover and precipitation. As Brazil is large enough and has enough tropical forest cover to affect its own rainfall (Oliveira et al., 2013; Stickler et al., 2013; Makarieva et al., 2014), I limit my analysis to Brazil. This paper tests whether there is correlation between rainfall at a weather station and instrumented upwind forest cover, given that there are fixed effects to account for spatial and temporal variation in rainfall.

When not instrumenting for forest cover, I find results that one unit of "Leaf Area Index" is associated with approximately 0.16-0.17 additional mm/day of precipitation, which is similar to the 0.25 - 0.40 mm/day found in Spracklen et al. (2012). However, when I use instrumented changes in forest cover, my results increase by an order of magnitude. This suggests that deforestation may have a much greater effect on precipitation than what has previously been believed by scientists. If human activity can indeed have such an impact on rainfall, that would contradict many assumptions economists make when including rainfall as a variable in econometric research. Though we know that humans are able to affect large-scale climate patterns, the exclusion restriction is satisfied as long as short-term and local deviation in precipitation is not correlated with the outcome of interest except via the first stage. However, if my results and those of Spracklen et al. (2012) are correct, rainfall would be an invalid instrument for some regressions in countries with enough tropical forest cover to affect their own weather.

Furthermore, beyond its use in instrumental variables, these findings suggest that forests have additional importance not typically accounted for in non-market valuation studies. There are a number of ecosystem services studies that attempt to value of the Amazon rainforest, starting with Peters et al. (1989) on the value of the intact Amazon rainforest for non-wood products, and Balick and Mendelsohn (1992) on its value for traditional medicine. Others have estimated the value of the rainforest based on its contribution to biodiversity (Fearnside, 1999), carbon storage (Boerner et al., 2007), and ecotourism (Kirkby et al., 2010). Protection of the Amazon is crucial for climate change mitigation (Nepstad et al., 2008; Soares-Filho et al., 2010) as well as for biodiversity. This paper shows that there is an additional marginal benefit of forest cover that has not been previously accounted for. I do not contribute a precisely estimated value of the Amazon via its effects on precipitation, but provide a rough estimate. More research is needed to better estimate the dollar value of the services that trees provide in terms of added precipitation.

In section 2.2, I describe the data used in this analysis, and section 2.3 covers the empirical strategy used. Section 2.4 includes the results, including robustness checks, and Section 2.5 is the conclusion.

2.2 Data

Forest Cover: MODIS Leaf Area Index

In order to measure vegetative cover, I use the MODBUV5 version of the Moderate-resolution Imaging Spectroradiometer (MODIS) Leaf Area Index dataset, available through the Land Processes Distributed Archive Center. Leaf Area Index (LAI) is a measure of leaf area per unit of ground area in a given grid cell, estimated via satellite. For example, a measure of "3.2" for Leaf Area Index indicates that, on average, there are 3.2 leaves between a point on the ground and the sky above. Because this data is already reported as a fraction, I do not transform it by taking its log. LAI is estimated by measuring the amount of light reflectance in the blue and green range, reflected by chlorophyll, and in the near-infrared, from internal leaf structures. Figure 2.1 shows the mean LAI across Brazil by month, and Figure 2.3 shows the variation in LAI across Brazil in December 2000 and December 2013. Clearly, there is variation in LAI within a year, across space, and over time. Figure 2.1 and Figure 2.2 show the mean and standard deviation of monthly Leaf Area Index and monthly precipitation across Brazil, with a curve plotted to match the average of each. When comparing the levels of LAI in Figure 2.1 and Figure 2.2, LAI does not appear to be perfectly correlated with the wet and dry seasons in Brazil. However, Saatchi et al. (2013) and Samanta et al. (2012b) have found a strong relationship between rainfall and LAI in Brazil, so differences in my figures may be due to aggregation across Brazil, which has a broad variety of climates. Comparing Figures 2.3 and 2.5, which include spatial variation, we see that the most forested regions of Brazil also appear to be the rainiest regions. This is not necessarily evidence that forests generate rainfall; it may be evidence that rainfall generates forests.

Using LAI has the disadvantage that agriculture is not excluded. Agricultural fields typically reflect less light in the blue, green, and near-infrared spectra, and will thus show up in the dataset with lower LAI than forests, but will not return "0" values. It would be valuable to replicate this analysis using the University of Maryland Landsat Forest Cover Change dataset, which reports changes from forest to non-forest cover. Deforestation should be more strongly affected by my instruments than LAI, and is a more relevant variable to economists. However, I chose to use LAI rather than change in forest cover because LAI has the advantage of providing a comparable analysis to the Spracklen et al. (2012) paper, and because leaf area is related to evapotranspiration, providing a more direct causal link between the independent and dependent variables.

The data are available starting February 18, 2000, which is when I begin my analysis. The MOD15A2 data set contains estimated Leaf Area Index at the 1 kilometer x 1 kilometer level, with one observation for every 8 days. I chose to use the same dataset as Spracklen et al. (2012), MODBUV5, which is a refined version of MOD15A2, developed by Ranga Myneni and others at Boston University.[1] MODBUV5 is aggregated to the monthly, 0.25° x 0.25° level, and only includes the highest-quality satellite measurements. Using this aggregated dataset has the potential to increase measurement error in my analysis, and I do so under the assumption that Leaf Area Index at an 8-day level is very similar to Leaf Area Index aggregated to the month level. Furthermore, the dataset loses observations by keeping only the highest quality observations. Choosing between these two datasets involves a tradeoff, and by selecting MODBUV5 I am using the assumption that the remaining observations are representative of the missing observations. This assumption is potentially problematic, because often observations are missing due to cloud and aerosol cover (Samanta et al., 2012a). If the areas covered by clouds and aerosols in a given month differ systematically from the freely visible areas, there may be measurement error that is biased in a certain direction. To mitigate this issue, I replace the missing data points with the mean of LAI within a 30 kilometer radius. Using this strategy, I am able to interpolate approximately 50% of the missing data, and the coefficients of my results increase.

[1]The MOD15A2 data is available at `https://lpdaac.usgs.gov/products/modis_products_table/mod15a2` and the MODBUV5 version is available by contacting Ranga Myneni (ranga.myneni@gmail.com).

Wind Data: NOAA ISD

I combine Leaf Area Index with Integrated Surface Data (ISD) from the Automated Weather Observing System, provided by the National Oceanic and Atmospheric Administration (NOAA).[2] I extract the data for 143 automated weather stations in Brazil, from 2000-2013. The NOAA ISD for Brazil include wind speed, wind direction, temperature, and dew point, as well as station elevation and observational quality. Table 1 provides summary statistics of this data. Each station reports once every 8 hours, so I transform these variables into the mean wind speed, wind direction, temperature and dew point per day. This introduces attenuation bias. An additional concern is that approximately 44% of the station-day observations do not include atmospheric pressure, and none include precipitation. The most important element of this data set for my analysis is the wind direction, which I interact with Leaf Area Index to create an index of upwind leaf area.

My analysis depends on the level of upwind forest cover in a given region. To determine this, I set my unit of observation to be a NOAA ISD station in Brazil. I therefore have 143 "individuals" in my analysis. For each station-month, I find the mean Leaf Area Index in each of four surrounding quadrants: northeast, southeast, southwest, and northwest. In my main analysis, these quadrants are 10° x 10°, or approximately 1111 km wide by 1111 km long (the width in particular will change throughout Brazil as the distance between points of longitude changes). The size of the quadrants was chosen because Spracklen et al. (2012) find that rainfall can be affected by forest cover as far as 1000 km away. Figure 2.7 provides an example of the relative size of a 10° x 10° quadrant. Clearly, stations near the coast will have some quadrants that contain few LAI data points. In Section 2.4, I repeat my analysis using smaller quadrants.

I combine the mean LAI per quadrant with dummy variables for quadrant of daily wind direction to create an index of upwind leaf cover which varies on a daily level. Because the wind directions 90°, 180°, 270° and 360° appear disproportionately often in the data, and my estimated upwind forest cover is particularly inaccurate for wind directions on the edge of a quadrant, I calculate LAI differently for due north, due south, due east, and due west. The relevant LAI is calculated to be the mean of the LAI in the two neighboring quadrants; for example, upwind LAI on a day when the wind is blowing due north is the mean of LAI in the northwest quadrant and LAI in the northeast quadrant.

Precipitation: TRMM Data

Precipitation is the dependent variable in my analysis. This variable is not available at the station level. I use Tropical Rainfall Measuring Mission (TRMM)-based precipitation estimates, which are satellite-interpolated rainfall estimates at the 0.25° x 0.25° level.[3] I use the daily version of these data, which are reported in mm/day of precipitation. I transform

[2]NOAA AWOS ISD data available at `http://www.ncdc.noaa.gov/data-access/land-based-station-data/land-based-datasets/automated-weather-observing-system-awos`.

[3]TRMM-based precipitation estimates available at `http://trmm.gsfc.nasa.gov/data_dir/data.html`.

the gridded data to an estimate of station-level precipitation using the mean of precipitation at the four grid points within a 0.25° radius of a given station. This introduces some measurement error, because the estimate of precipitation will not exactly match the true precipitation level at a given station, but has the advantage of eliminating station-level precipitation measurement error. See Figure 2.2 for the mean precipitation throughout the year (transformed to mm/hr) and Figure 2.5 for examples of spatial variation in rainfall across Brazil. The variation of rainfall within a year, across years, and across space will be crucial for my analysis.

Prices: IMF Commodity Price Indices

I use the International Monetary Fund (IMF) Commodity Prices and Price Indices for the years 1999 - 2013.[4] The data combine monthly average prices of goods at a specific location, in 2005 U.S. dollars. For example, the IMF soybean price is based on the $/metric ton price of a soybean futures contract for No. 2 yellow soybeans via the Chicago Board of Trade. I use the following price indices based on the more important agricultural goods in Brazil: hardwood, softwood, maize, beef, and soybeans. I lag the prices of these goods by one year. The IMF data vary monthly, and include one global price per commodity. Table 2 provides summary statistics for the IMF price data.

2.3 Empirical Strategy

To overcome the endogeneity of vegetative cover, I employ the following two stage least squares approach:

$$y_{it} = \beta UW\hat{L}AI_{it} + x_{it}'\theta + \delta_m + \delta_i + \delta_w + \varepsilon_{it}$$
$$UW\hat{L}AI_{it} = log(z_m)'\hat{\lambda} + x_{it}'\hat{\theta} + \hat{\delta_m} + \hat{\delta_i} + \hat{\delta_w},$$

where y_{it} represents rainfall measured in mm/day, UWLAI is upwind leaf area index, x_{it} is a series of covariates that change based upon the specification, δ_m is a series of month fixed effects, δ_i is a series of fixed effects for each station in Brazil, and δ_w are wind quadrant fixed effects. Some variables vary monthly and some vary daily, hence the notation m for month and t for day. UWLAI is estimated using instruments z_t, as well as all fixed effects and exogenous covariates.

I also run a naive regression:

$$y_{it} = \beta UWLAI_{it} + x_{it}'\theta + \delta_m + \delta_i + \delta_w + \epsilon_{it},$$

where $UWLAI_{it}$ is not instrumented, to explore the extent to which endogeneity is important in this model.

[4] IMF Commodity Price Data available at http://www.imf.org/external/np/res/commod/index.aspx.

Choice of Instruments

Commodity prices, as reported by the IMF, are used as instruments. There are 180 commodities or indices reported in the IMF data, and in my main specifications I use the 12-month-lagged price of hardwoods (wood from non-coniferous trees). The price of hardwood was chosen as an instrument because the price of timber has a less complex relationship with LAI than the price of a leafy crop, such as maize. In the IMF dataset, the price of hardwood is calculated using the price of Malaysian logs imported to Japan, and the price of Malaysian sawnwood in the UK. I assume that these prices are correlated with the price of Brazilian hardwoods, and so I expect an increased price of hardwood to be correlated with decreased leaf area index 12 months afterward.

Because Brazil has a great deal of heterogeneity in terms of land types, I interact the lagged price of hardwoods with state fixed effects to allow the response to IMF hardwood prices to vary by state. Including station-level fixed effects creates station-level spatial variation in my estimates of upwind leaf area index. Furthermore, I use daily wind direction fixed effects to create variation on the daily, rather than monthly, level. Because of the number of variables to report in the first stage, I only use the lagged price of hardwoods in my main specification. As a robustness check, I conduct a regression that includes all relevant prices that show a strong first stage.

Exclusion Restriction

To use IMF prices as instruments for LAI, I need to assume that $cov(z_m, UWLAI_{it}) \neq 0$ and $cov(z_m, \varepsilon_{it}) = 0$. The first condition is easily satisfied by choosing only prices where $cov(z_m, UWLAI_{it}) \neq 0$. However, I cannot mathematically prove that the exclusion restriction holds. The second stage dependent variable is rainfall, which luckily is exogenous to many factors. But, as was stated earlier in this paper, rainfall can be affected by human activity such as pollution and heat from urban centers (Dettwiller and Changnon, 1976; Rosenfeld, 2000; Shepherd et al., 2002; Shepherd, 2005). According to the World Bank, the urban population of Brazil steadily increased from 141.7 million in 2000 to 170.7 million in 2013. This would create increased rainfall around cities, which coincides with the increasing prices of hardwoods over time. However, any violation of the exclusion restriction will disappear when including station-time-interacted fixed effects. I test this in Robustness Checks section.

Even if rainfall is not affected by anything that correlates with hardwood prices, there may be correlations between hardwood prices and the second-stage error term. For example, if rainfall in Brazil is sufficiently predictable, or has sufficient serial correlation, and if global hardwood prices are affected by the prices of Brazilian hardwood, lagged hardwood prices may correlate with current rainfall. I use a full 12-month lag to reduce predictability of rain. Because this analysis include month-by-year fixed effects, day-to-day variation in rainfall would need to be predictable 12 months prior, and its prediction would need to affect global prices, in order for the exclusion restriction to fail. I assume that this is not the case.

As I will discuss in the Robustness Checks section, omitted variables may be the underlying cause of the relationship between prices and LAI. As long as these omitted variables do not affect Brazilian rainfall, the exclusion restriction still holds. Because rainfall is determined by temperature, available moisture, and air pressure, it seems unlikely that many omitted variables could affect local rainfall on such a short time scale.

Interpretation of Treatment Effect

As is the case with any instrumental variables strategy, the coefficient of interest in my analysis will represent a local average treatment effect. The $\hat{\beta}$ that I estimate will represent the effect of vegetative cover on rainfall for vegetation that is affected by IMF-reported prices. This will be the effect of vegetative cover on rainfall for marginal land; that is, land whose use will change given reasonable changes in the profitability of agriculture or logging. When I use the price of hardwood as an instrument, I should be measuring the treatment effect of changes in forest cover on rainfall, for land on which forest cover changes in response to exogenous world prices. Though this treatment effect cannot be applied to other types of land, it is useful in that forest policy prescriptions should be based on the impact of human-induced changes in use of marginal land.

It is important to determine whether I have measured a true change in total precipitation, or merely the relocation of precipitation from one area to another. Forests affect rainfall by increasing moisture availability, but also by changing air pressure, which draws in moisture from elsewhere. To evaluate the importance of forests, we must know whether this moisture come from nearby, less forested regions, or from areas such as the ocean. Makarieva et al (2014) find that changes in forest cover in Brazil affect large-scale, ocean-to-land air transport systems; that is, the total amount of precipitation in Brazil increases when there is more forest cover. Thus the changes in rainfall that we see do not merely reflect a different distribution pattern for rain.

2.4 Results

First I present the naive regressions, in Table 2.1. Column (1) contains the results of a regression that does not account for any of the endogeneity between LAI and precipitation. We see that precipitation is strongly correlated with upwind vegetative cover; a 1-unit increase in upwind on a given day is correlated with a 0.993 mm/day increase in precipitation. However, this regression does not account for spatial or temporal factors that may affect both leaf area index and precipitation. Regressions (2) - (4) include wind direction and weather station fixed effects, clustered standard errors, and ways of accounting for changes that correlate with time. Column (2) reports a regression with fixed effects for month of the year (Jan, Feb, etc) and a quadratic year time trend, and column (3) replaces the time trend with year fixed effects. We see that when accounting for these factors, upwind leaf area is not significantly correlated with changes in precipitation. Finally, column (4) includes temperature and dew point, which are part of the mechanism for how forest cover affects rainfall. I would expect that including these variables should reduce the coefficient on UWLAI, thus demonstrating that I have specified the correct mechanism. Here we see that the coefficient on UWLAI is reduced enough to become negative and significant at the 0.1% level. I do not have an explanation for why we might see negative correlations between UWLAI and precipitation when the ways that forest cover affects precipitation are taken into account.

Tables 2.2 and 2.3 show the coefficients of the first stage of the two stage least squares analysis. In Table 2.2, we see that a 1% increase in the lagged price of hardwoods is correlated with 0.4 - 0.5 unit decrease in LAI, or 0.4-0.5 fewer leaves above a given point on the ground. This is a very large effect, suggesting a strong instrument. In addition, almost all of the coefficients on price-state interaction variables are significant. It appears that I do not have a weak instruments problem. It may be the case that the size of these coefficients is due to the true relationship between farmgate prices (which I do not use), deforestation choices, and LAI. However, I cannot rule out the option that omitted variables drive the significant relationship seen in Tables 2.2 and 2.3.

Table 2.4 shows the second-stage results of the two stage least squares approach. Column (1) contains the same fixed effects and clustering as column (2) of Table 2.1 and column (1) of Tables 2.2 and 2.3. The regression in column (1) includes station, month, and wind direction fixed effects, as well as a year time trend. Column (2) includes year fixed effects. We see that the results are not sensitive to how I control for time. There is a less than 0.1% chance that we would have the data that we have if the null hypothesis of no correlation between rainfall and UWLAI is correct. As is to be expected, the coefficient of interest is no longer significant when including the mechanism, as is the case in column (3).

When I do not instrument for UWLAI, my results are similar to those of Spracklen et al. (2012); I find a coefficient of 0.16-0.17 whereas they find 0.25 - 0.40 mm/day for an additional unit of OWLAI. However, when I instrument for changes in forest cover, these results increase substantially. If the causal interpretation of these results is to be believed, a 1 unit increase in upwind LAI on a given day leads to a 2.613- 2.674 mm increase in rain on that day. This is a large change in precipitation; the mean daily rainfall in my sample is 4.12 mm, and the mean daily rainfall on days when it rains is 9.074. However, a 1 unit increase in LAI is also a substantial change. The mean day-to-day change in UWLAI seen by a weather station in my sample is 0.474 units, and much of that change is due to switching quadrants from day to day, and from seasonal changes in forest cover. The mean change in LAI for a given 0.25° x 0.25° square in Brazil, from one year to the next, is 0.308 units. I cannot calculate the mean change in LAI for a given square from day to day, as the LAI data is reported on a monthly level. It is clear that we would expect to see somewhat less change in LAI than 1 unit per day, but the coefficients in Table 2.4 suggest that if a every spot in 10° x 10° quadrant were to be covered by one more leaf, the area downwind of that quadrant would receive approximately 2.6 mm more rain per day. Of course, we must also take into account the fact that Table 2.4 presents a local average treatment effect, and the previous sentence holds only if treatment of all units would affect rainfall in the same way. Makarieva et al. (2014) suggests that this is not the case, and that small patches of deforestation in otherwise forested areas have larger effects on rainfall than deforestation in mostly deforested areas. Depending on whether the marginal lands affected by changes in prices are in heavily-forested inland Brazil, or less-forested coastal areas, the average treatment could be larger or smaller than the coefficients in Table 2.4.

Robustness Checks

Though the significance levels of my results are not sensitive to the specification, the coefficient of interest does change slightly. I conduct the following robustness checks to test the sensitivity of my results.

First, I change the size of the quadrants that are used to calculate upwind leaf area. The quadrants used for Table 2.4 are 10° x 10° , or approximately 1110 km wide x 1110 km long, a size chosen because Spracklen et al. found that the effect of LAI on precipitation holds for up to 1000 km. Using such large quadrants has the advantage of taking a mean over more MODIS data points, which is useful for stations with sparse data points nearby. The large quadrants are also useful because they take into account all of the LAI that could reasonably affect station-level precipitation, if wind direction did not change over space. The 1110 km x 1110 km have the disadvantage of introducing large measurement error if wind direction changes as distance from a station increases, or if the effect of LAI on precipitation increases as distance from a station decreases. I run the same regression as in column (2) of Table 2.4, but for quadrants that are 5° x 5° and 2° x 2° in size. For the 5° x 5° quadrants, the coefficient on upwind leaf area index is 2.307, significant at the 0.001 level. When constraining the areas of interest to 2° x 2° quadrants around each station, the coefficient decreases to 1.726, significant at the 0.01 level. This may be because more distant regions do have a strong effect on precipitation, and they also have different LAI. At the very least, it appears that the effect holds over the range of distances determined by Spracklen et al.

Another potential concern is whether my findings are specific to the instruments chosen. To test this, I choose different price variables from the IMF data set. We may see different local average treatment effects when using different instruments, but should expect each coefficient to be significant and positive. For this robustness check, I once again use 10° x 10° quadrants. The coefficient on upwind LAI does not change in a meaningful way when I use the 12-month lagged prices of soy and maize, the 12-month lagged price of beef, or lagged soy, maize, beef and hardwood prices simultaneously, as can be seen in Table 2.5. I also instrument UWLAI using the log of 12-month lagged prices of softwoods, and the coefficient in this regression is not statistically significant. This is surprising because Brazil does produce softwoods and the first-stage coefficients (not shown) are statistically significant. Softwoods come from coniferous trees, and it is possible that these trees have different effects on moisture availability or atmospheric pressure, or that these trees are in areas that are not moist enough to generate their own rainfall. It is also possible that there are other problems with my analysis, such as omitted variables that are correlated with rainfall and with the prices of some goods. To test for spurious correlations between prices and rainfall, I run a placebo test using the log of the current price of metal as my instrument for UWLAI. This coefficient is not statistically significant, and so it appears to be unlikely that omitted variables drive the results we see elsewhere.

Next, I test whether wind direction drives much of the effects we see in Table 2.4. In previous regressions, I did not include interaction terms between station fixed effects and wind direction fixed effects, which create station-quadrant fixed effects. This does not account for any time-invariant differences between quadrants, or for wind direction to have different effects on rainfall in different parts of Brazil. Table 2.7 includes quadrant fixed effects. Column (4) reports a regression with softwood as the instrument for UWLAI. The non-significance of the coefficient of interest is not surprising given the previous results when using softwood as an instrument. Column (2) reports results from a regression using lagged soy and maize prices, and column (3) reports results from a regression using lagged beef prices. When accounting for quadrant fixed effects, the coefficients have not changed by much, but the standard errors increase. Including quadrant fixed effects greatly reduces the overall variation in my data, given that I already include month-by-year fixed effects and wind direction fixed effects, and that my LAI data only varies at the month level. It could be the case that my specification is overparameterized when including this additional level of fixed effects beyond station fixed effects, or it could be that differences in the effect of wind direction across Brazil drives most of the variation that generates my results.

2.5 Conclusion

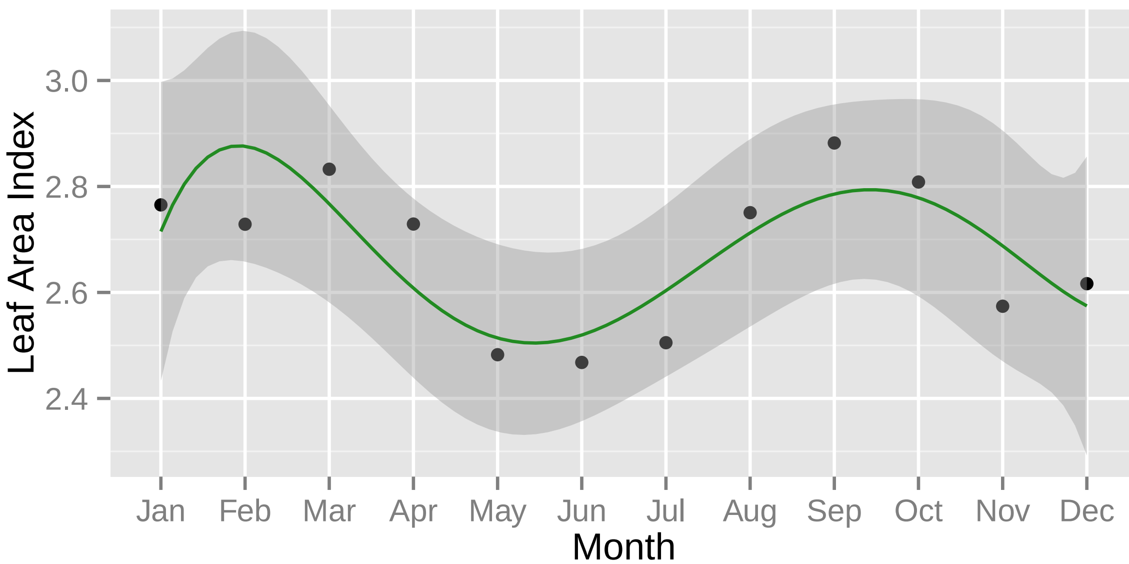
This paper presents suggestive evidence that vegetative cover does indeed affect precipitation in Brazil. To account for issues of reverse causality, I instrument upwind leaf cover using lags of IMF-reported prices. Using some of these instruments, I estimate that a 1 unit increase in the fraction of upwind leaf cover yields a 2.577-2.636 mm/day increase in rainfall (see Table 2.4). Similar, albeit smaller, results are obtained when investigating LAI in smaller areas around a station. However, these results do not hold when I interact station fixed effects with wind direction or year fixed effects.

If forest cover does affect precipitation, this would be an important finding. Many attempts have been made to estimate the value of forests, but none that I am aware of incorporate the effects that forests can have on precipitation. The focus of this paper is on whether there is indeed an effect on rainfall, but it is simple to calculate a rough estimate of the additional value that forests have, given this effect. First, I need a value of water; I use an estimate from Schlenker et al. (2005), who find that the value of water in the United States is \$324-\$656 per acre-foot. Looking at one specific month, the year-over-year mean LAI in Brazil in December 2013 increased by 0.130 units. Assuming that 10% of the rainfall in Brazil can be captured and used for agriculture, the extra rain that fell in Brazil in December 2013 alone due to increased LAI is worth \$2.3 -\$4.7 billion. Clearly, the effects of forest cover on precipitation are economically significant.

These results also have methodological implications within applied econometrics. Many economists assume that rainfall is an exogenous variable that will satisfy the exclusion restriction for most regressions. If forest cover can affect precipitation, many instrumental variables analyses would need to be reassessed for possible exclusion restriction violations. These violations are thus far only known to possibly occur in areas with enough moisture and unbroken forest cover for forests to generate rainfall, but this includes vast areas that are popular for development economics research.

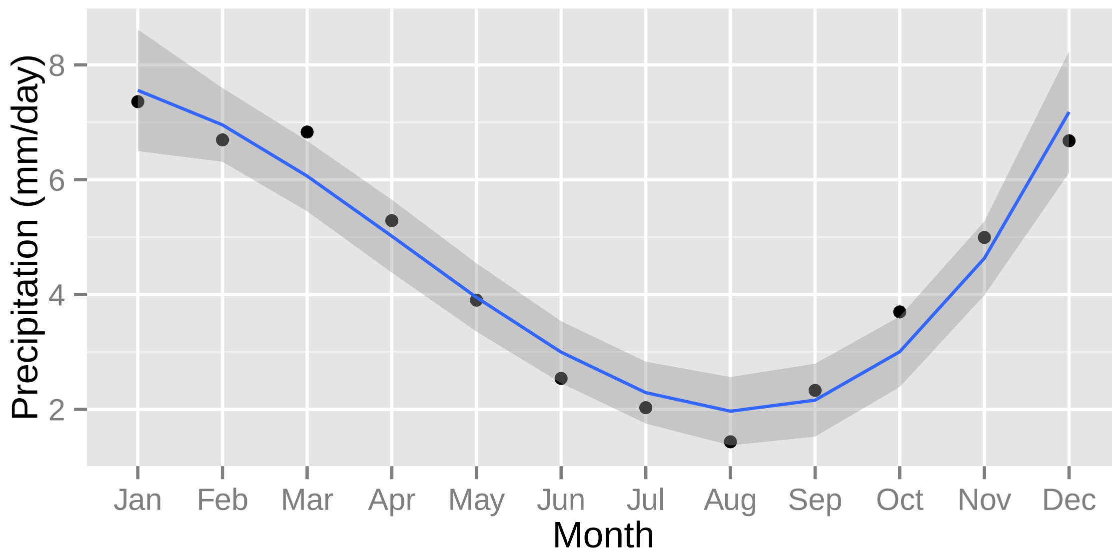
Further research is certainly needed in this area. It would be valuable to conduct an analysis that uses deforestation data, rather than Leaf Area Index data. It would also be valuable to conduct an analysis that better accounts for the passage of air over land, and includes additional important variables such as distance from the ocean. This paper has shown that there is most likely an effect of forest cover on rainfall; the next step is to have a more precise and accurate measure of the effect, and an estimation of the value of forests due to their effect on rainfall.

Figure 2.1: Mean Leaf Area Index in Brazil by Month, 2000-2013



Notes: The figure above depicts the mean Leaf Area Index value across Brazil for each month from 2000-2013. Leaf Area Index (LAI) is a measure of leaf area per unit of ground area in a given grid cell. These data come from the MODBUV5 version of the Moderate-resolution Imaging Spectroradiometer (MODIS) data packages available through the Land Processes Distributed Archive Center. A best-fit curve is depicted by the green line, with a 95% confidence interval depicted in gray.

Figure 2.2: Mean Precipitation in Brazil by Month, 2000-2013



Notes: The figure above depicts the mean precipitation, measured in mm/day, across Brazil for each month from 2000-2013. These data come from the Tropical Rainfall Measuring Mission (TRMM), and are satellite-interpolated at the 0.25° x 0.25° level. A best-fit curve is depicted by the blue line, with a 95% confidence interval depicted in gray.

(a) December 2000 (b) December 2013

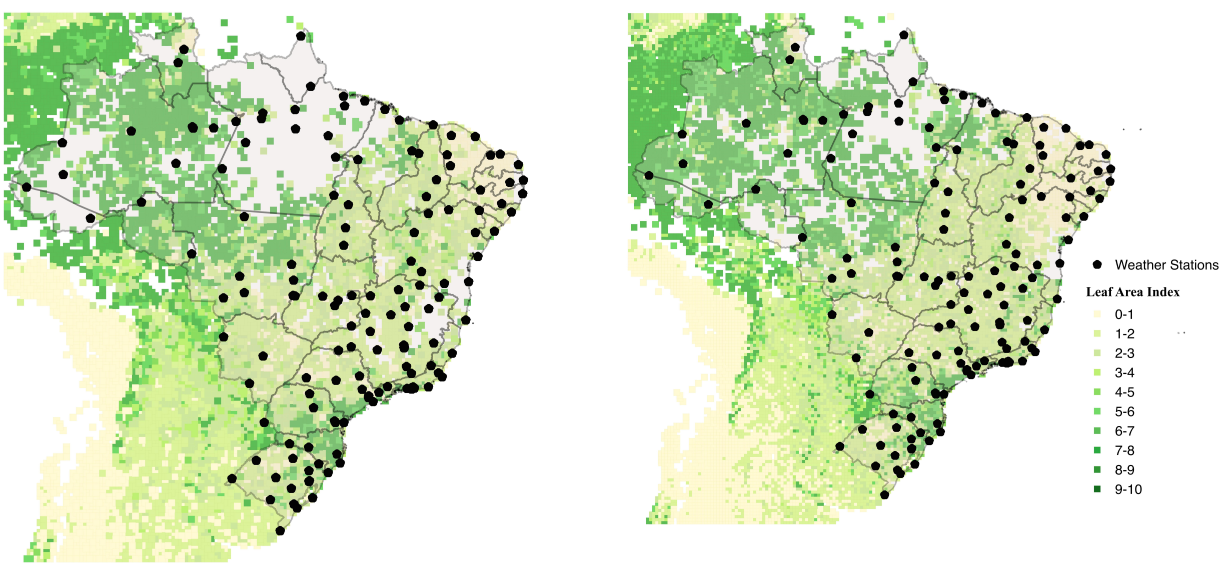


Figure 2.3: LAI across Brazil and Surrounding Regions

Notes: The figures above depict the Leaf Area Index (LAI) in each 0.25°x 0.25°grid cell, as reported by the MODBUV5 data product. LAI is a measure of leaf area per unit of ground area in a given grid cell, estimated via satellite. Areas without green or yellow grid cells have missing data, most likely due to cloud cover. The black circles are locations of the automated weather stations that I use in my analysis. The states of Brazil are outlined in light gray. The figure on the left reports LAI in December 2000, and the figure on the right reports LAI in December 2013.

(a) December 2000 (b) December 2013

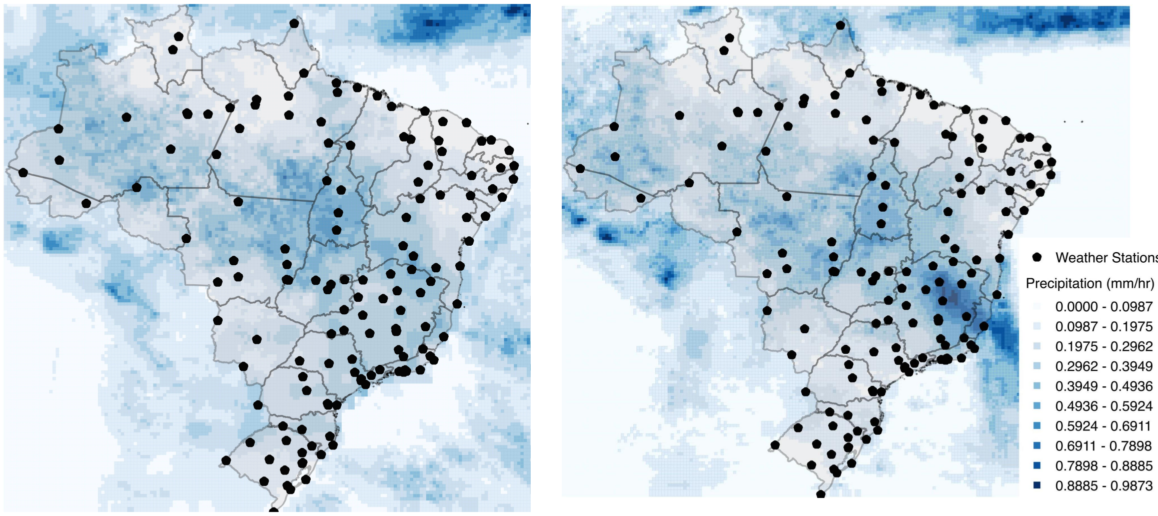


Figure 2.5: Mean Precipitation across Brazil and Surrounding Regions

Notes: The figures above depict mean monthly precipitation across Brazil and the surrounding regions, measured in millimeters per hour, in December 2000 (left) and December 2013 (right). Precipitation estimates come from the Tropical Rainfall Measuring Mission (TRMM) and are satellite-interpolated. The black circles are locations of the automated weather stations that I use in my analysis. The states of Brazil are outlined in light gray.

Table 2.1: Coefficients of Covariates on Precipitation, Non-Instrumented LAI

	Dependent variable:			
	Precipitation (mm/day)			
	(1)	(2)	(3)	(4)
Upwind Leaf Area Index	0.993***	0.174	0.162	−0.372***
	(0.011)	(0.159)	(0.156)	(0.107)
Elevation	0.0003***			
	(0.00004)			
Year		−0.002		
		(0.009)		
Temperature				−0.896***
				(0.095)
Dew Point				0.964***
				(0.140)
Constant	1.521***			
	(0.038)			
Station Fixed Effects	no	yes	yes	yes
Month (Jan, Feb, Mar, etc) Fixed Effects	no	yes	yes	yes
Year Fixed Effects	no	no	yes	yes
Year Time Trend	no	yes	no	no
Wind Direction Fixed Effects	no	yes	yes	yes
Month-by-Year Clustering	no	yes	yes	yes
State-Level Clustering	no	yes	yes	yes
Observations	544,450	544,450	544,450	544,450
R^2	0.016	0.065	0.066	0.147
Adjusted R^2	0.016	0.064	0.065	0.147

Notes: This table shows the results of a regression of precipitation, measured in mm/day, on upwind leaf area index (UWLAI). The regression in column (1) does not include any fixed effects, but it does control for the elevation of the weather station. The regression in column (2) includes fixed effects for weather station, month, and wind direction (north, northeast, east, southeast, south, southwest, west, or northwest), and a time trend. The standard errors for this regression are clustered at the month of sample and state level. The regression in column (3) is similar to that of column (2), but with year fixed effects in place of a year time trend. The regression in column (4) is identical to that of column (3) except for added controls for temperature and dew point at the weather station. Standard errors are in parentheses. Significance: $^{*}p < 0.1;^{**} p < 0.05;^{***} p < 0.01$.

Table 2.2: First Stage Coefficients of Covariates on Upwind Leaf Area Index, Part 1

	Dependent variable:		
	Upwind Leaf Area Index		
	(1)	(2)	(3)
$log(P_{Hardwood,m-12})$	0.458^{*}	0.541^{**}	0.646^{***}
	(0.186)	(0.199)	(0.194)
Year Time Trend	0.006		
	(0.003)		
Temperature			-0.004^{***}
			(0.001)
Dew Point			0.055^{***}
			(0.002)
$log(P_{Hardwood,m-12})$∗Alagoas	-0.680^{**}	-0.676^{**}	-0.671^{**}
	(0.217)	(0.216)	(0.206)
$log(P_{Hardwood,m-12})$∗Amapa	-0.687^{**}	-0.679^{**}	-0.854^{***}
	(0.243)	(0.243)	(0.228)
$log(P_{Hardwood,m-12})$∗Amazonas	-0.427^{**}	-0.431^{**}	-0.533^{***}
	(0.164)	(0.164)	(0.159)
$log(P_{Hardwood,m-12})$∗Bahia	-0.600^{**}	-0.605^{**}	-0.733^{***}
	(0.203)	(0.203)	(0.191)
$log(P_{Hardwood,m-12})$∗Ceara	-0.824^{***}	-0.823^{***}	-0.949^{***}
	(0.236)	(0.236)	(0.219)
$log(P_{Hardwood,m-12})$∗Distrito Federal	-0.561^{**}	-0.560^{**}	-0.599^{**}
	(0.210)	(0.210)	(0.196)
$log(P_{Hardwood,m-12})$∗Espirito Santo	-0.408^{*}	-0.405^{*}	-0.451^{*}
	(0.203)	(0.203)	(0.198)
$log(P_{Hardwood,m-12})$∗Goias	-0.517^{*}	-0.510^{*}	-0.601^{**}
	(0.205)	(0.206)	(0.189)
$log(P_{Hardwood,m-12})$∗Maranhao	-1.131^{***}	-1.124^{***}	-1.181^{***}
	(0.249)	(0.249)	(0.239)
$log(P_{Hardwood,m-12})$∗Mato Grosso	-0.501^{*}	-0.505^{*}	-0.608^{**}
	(0.201)	(0.201)	(0.188)
$log(P_{Hardwood,m-12})$∗Mato Grosso do Sul	-0.610^{**}	-0.610^{**}	-0.640^{***}
	(0.200)	(0.200)	(0.190)
$log(P_{Hardwood,m-12})$∗Minas Gerais	-0.537^{**}	-0.534^{**}	-0.631^{***}
	(0.201)	(0.201)	(0.188)
$log(P_{Hardwood,m-12})$∗Para	-0.807^{***}	-0.807^{***}	-0.928^{***}
	(0.163)	(0.163)	(0.158)

Note: $^{*}p < 0.05;^{**} p < 0.01;^{***} p < 0.001$

Table 2.3: First Stage Coefficients of Covariates on Upwind Leaf Area Index

	Dependent variable:		
	Upwind Leaf Area Index		
	(1)	(2)	(3)
$log(P_{Hardwood,m-12})$*Paraiba	-0.519^{*}	-0.518^{*}	-0.673^{**}
	(0.248)	(0.249)	(0.232)
$log(P_{Hardwood,m-12})$*Parana	-0.569^{**}	-0.570^{**}	-0.663^{***}
	(0.202)	(0.202)	(0.190)
$log(P_{Hardwood,m-12})$*Pernambuco	-0.720^{***}	-0.717^{***}	-0.777^{***}
	(0.214)	(0.214)	(0.196)
$log(P_{Hardwood,m-12})$*Piaui	-0.836^{***}	-0.833^{***}	-0.921^{***}
	(0.217)	(0.217)	(0.200)
$log(P_{Hardwood,m-12})$*Rio de Janeiro	-0.426^{*}	-0.416^{*}	-0.500^{*}
	(0.205)	(0.205)	(0.198)
$log(P_{Hardwood,m-12})$*Rio Grande do Norte	-0.484^{*}	-0.477^{*}	-0.543^{*}
	(0.225)	(0.225)	(0.211)
$log(P_{Hardwood,m-12})$*Rio Grande do Sul	-0.601^{**}	-0.597^{**}	-0.697^{***}
	(0.196)	(0.196)	(0.183)
$log(P_{Hardwood,m-12})$*Rondonia	-0.712^{***}	-0.712^{***}	-0.779^{***}
	(0.151)	(0.151)	(0.148)
$log(P_{Hardwood,m-12})$*Roraima	-0.295	-0.297	-0.529^{**}
	(0.209)	(0.209)	(0.194)
$log(P_{Hardwood,m-12})$*Santa Catarina	-0.445^{*}	-0.438^{*}	-0.565^{**}
	(0.193)	(0.194)	(0.183)
$log(P_{Hardwood,m-12})$*Sao Paulo	-0.251	-0.241	-0.360
	(0.212)	(0.212)	(0.196)
$log(P_{Hardwood,m-12})$*Sergipe	-0.668^{*}	-0.667^{*}	-0.685^{*}
	(0.296)	(0.291)	(0.273)
$log(P_{Hardwood,m-12})$*Tocantins	-0.424^{*}	-0.422^{*}	-0.501^{**}
	(0.213)	(0.213)	(0.194)
Station Fixed Effects	yes	yes	yes
Month (Jan, Feb, Mar, etc) Fixed Effects	yes	yes	yes
Year Fixed Effects	no	yes	yes
Year Time Trend	yes	no	no
Wind Direction Fixed Effects	yes	yes	yes
Month-by-Year Clustering	yes	yes	yes
State-Level Clustering	yes	yes	yes
Observations	544,450	544,450	544,450
R^2	0.731	0.731	0.741
Adjusted R^2	0.731	0.731	0.741

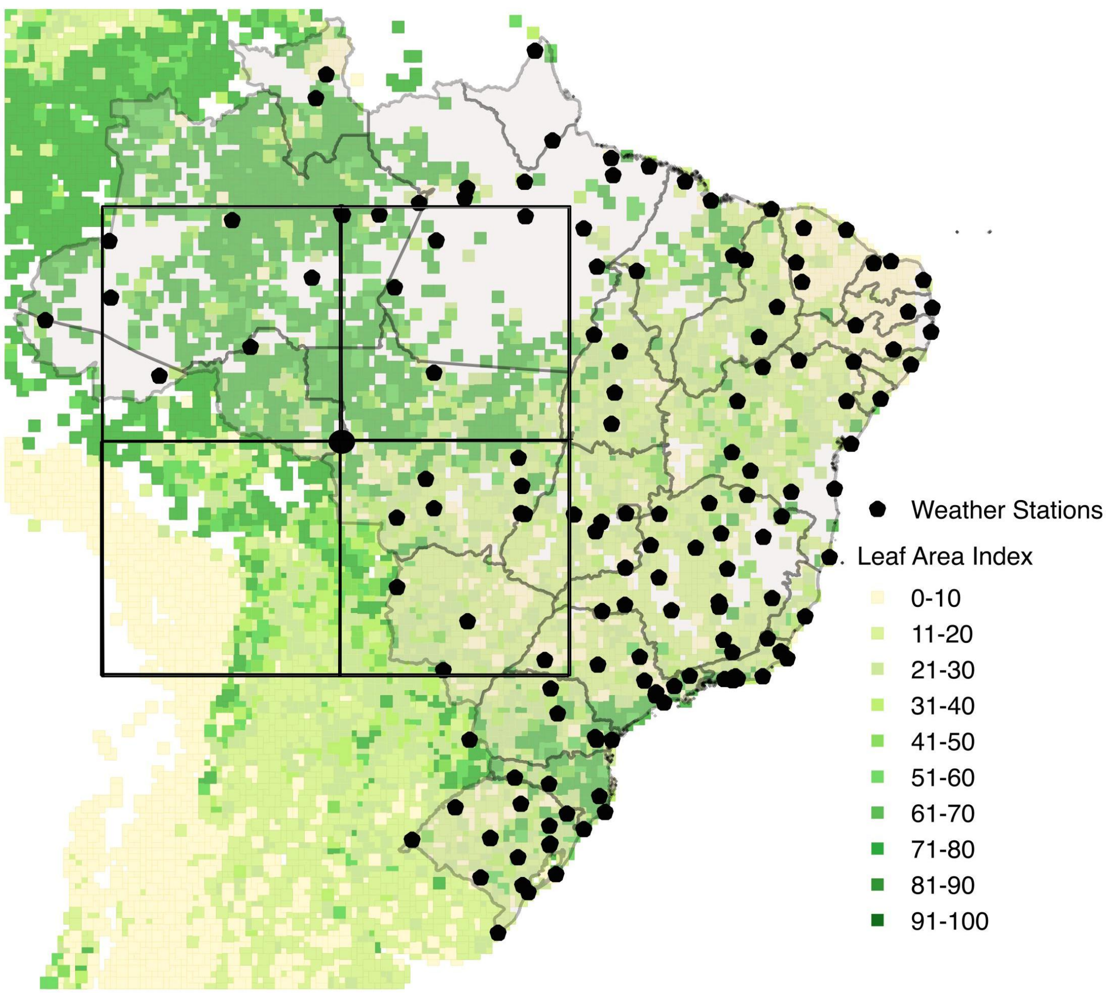
Notes: This table shows the results of a regression of UWLAI on the IMF-reported 12 month lagged price of hardwood, and interactions between this price and state fixed effects. The omitted state is Acre. Standard errors are in parentheses, and are two-way clustered by state and month-of-sample.

Table 2.4: Coefficients of Covariates on Precipitation, Instrumented LAI

	Dependent variable:		
	Precipitation (mm/day)		
	(1)	(2)	(3)
$\widehat{\text{UWLAI}}$	2.613***	2.674***	1.624
	(0.749)	(0.717)	(1.063)
Year Time Trend	−0.001		
	(0.008)		
Temperature			−0.887***
			(0.097)
Dew Point			0.854***
			(0.159)
State-Price Interaction Variables	yes	yes	yes
Station Fixed Effects	yes	yes	yes
Month (Jan, Feb, Mar, etc) Fixed Effects	yes	yes	yes
Year Fixed Effects	no	yes	yes
Year Time Trend	yes	no	no
Wind Direction Fixed Effects	yes	yes	yes
Month-by-Year Clustering	yes	yes	yes
State-Level Clustering	yes	yes	yes
Observations	544,450	544,450	544,450
R^2	0.038	0.037	0.130
Adjusted R^2	0.037	0.037	0.130

Notes: This table shows the results of a regression of precipitation, measured in mm/day, on instrumented upwind leaf area index (UWLAI). UWLAI is instrumented using the IMF-reported 12 month lagged price of hardwood, and interactions between this price and state fixed effects (see Tables 2 and 3). All regressions in this table have weather station, wind direction, and month fixed effects. Column (1) has a year time trend, and columns (2) and (3) use year fixed effects. Column (3) includes controls for temperature and dew point. Standard errors are in parentheses, and are two-way clustered by state and month-of-sample. Significance: $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$.

Figure 2.7: Sample 10°x 10° Quadrants around Vilhena Station



Notes: This figure depicts the size and shape of quadrants used in my analysis, around a single weather station. In most regressions, I measure LAI in 10°x 10°quadrants to the northeast, northwest, southeast, and southwest of each station. The size of the quadrants was chosen because Spracklen et al. (2012) find that rainfall can be affected by forest cover as far as 1000 km away.

Table 2.5: Coefficients on Instrumented LAI, Varying Instruments

	Dependent variable:			
	Precipitation (mm/day)			
	Soy, Maize	Beef	Soy, Maize, Beef, Hardwood	Softwood
	(1)	(2)	(3)	(4)
Upwind Leaf Area Index	1.896**	2.193***	1.226	0.871
	(0.677)	(0.476)	(1.923)	(1.029)
Station Fixed Effects	yes	yes	yes	yes
Month-by-Year Fixed Effects	yes	yes	yes	yes
Wind Direction Fixed Effects	yes	yes	yes	yes
State-Level Clustering	yes	yes	yes	yes
Month-by-Year Clustering	yes	yes	yes	yes
Observations	538,290	538,290	538,290	538,290
R^2	0.054	0.049	0.062	0.064
Adjusted R^2	0.053	0.048	0.061	0.064

Notes: This table shows the results of a regression of precipitation, measured in mm/day, on instrumented upwind leaf area index (UWLAI). Each column uses different instruments for UWLAI. In column (1), I use the 12-month lagged prices of soy and maize. Column (2) uses the 12-month lagged price of beef, and column (3) uses lagged soy, maize, beef and hardwood prices simultaneously. Column (4) uses 12-month lagged prices of softwoods. All regressions in this table have weather station, wind direction, and month fixed effects. Standard errors are in parentheses, and are two-way clustered by state and month-of-sample. Significance: $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$.

Table 2.6: Coefficients of Covariates on Precipitation, Varying Size of Quadrants

	Dependent variable:					
	Precipitation					
	10 x 10 Quadrant		5 x 5 Quadrant		2 x 2 Quadrant	
	(1)	(2)	(3)	(4)	(5)	(6)
Upwind Leaf Area Index	0.162	2.674***	0.307	2.307***	0.254	1.726**
	(0.156)	(0.717)	(0.209)	(0.516)	(0.242)	(0.533)
Instrumented UWLAI	no	yes	no	yes	no	yes
Station Fixed Effects	yes	yes	yes	yes	yes	yes
Wind Direction Fixed Effects	yes	yes	yes	yes	yes	yes
State-Level Clustering	yes	yes	yes	yes	yes	yes
Month-by-Year Clustering	yes	yes	yes	yes	yes	yes
Observations	544,450	544,450	538,290	538,290	538,290	538,290
R^2	0.067	0.045	0.067	0.051	0.067	0.058
Adjusted R^2	0.066	0.044	0.066	0.050	0.066	0.057

Notes: This table shows the results of a regression of precipitation, measured in mm/day, on instrumented upwind leaf area index (UWLAI). Columns (1), (3), and (5) use non-instrumented UWLAI. In columns (2), (4), and (6), UWLAI is instrumented using the IMF-reported 12 month lagged price of hardwood, and interactions between this price and state fixed effects. Columns (1) and (2) use UWLAI within 10°x 10°quandrants, columns (3) and (4) use UWLAI within quadrants that are 5°x 5°in size, and columns (5) and (6) use UWLAI within quandrants 2°x 2°in size. All regressions in this table have weather station, wind direction, and month fixed effects. Standard errors are in parentheses, and are two-way clustered by state and month-of-sample. Significance: $^{*}p < 0.1$;$^{**}p < 0.05$;$^{***}p < 0.01$.

Table 2.7: Coefficients of Covariates on Precipitation, Station-Wind Direction Interaction Fixed Effects

	Dependent variable:			
	Precipitation (mm/day)			
	(1)	(2)	(3)	(4)
Upwind Leaf Area Index	2.293*	1.580	2.354**	3.426
	(0.989)	(0.817)	(0.769)	(2.395)
Station-Wind Direction Fixed Effects	yes	yes	yes	yes
Month-by-Year Fixed Effects	yes	yes	yes	yes
Month-by-Year Clustering	yes	yes	yes	yes
State-Level Clustering	yes	yes	yes	yes
Observations	556,568	556,568	556,568	556,568
R^2	0.076	0.079	0.075	0.066
Adjusted R^2	0.072	0.076	0.071	0.062

Notes: This table shows the results of a regression of precipitation, measured in mm/day, on instrumented upwind leaf area index (UWLAI). All regressions in this table have weather station by wind direction fixed effects as well as month of sample fixed effects. The regression for column (1) uses lagged hardwood prices. Column (2) reports results from a regression using lagged soy and maize prices, and column (3) reports results from a regression using lagged beef prices. Column (4) reports a regression with softwood as the instrument for UWLAI. Standard errors are in parentheses, and are two-way clustered by state and month-of-sample. Significance: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Chapter 3

Risk updating versus renovation: A study of post-flood property values

3.1 Introduction

Estimating the economic costs of flooding is crucial for policymakers, investors, and insurers. Understanding these costs helps us know how much to invest in flood defenses and climate change mitigation, as well as to predict future expenses for insurers or disaster relief agencies. Floods impose direct costs by damaging physical property, and occasionally by injuring or killing local residents. Floods also impose indirect costs: they can cause a drop in economic productivity due to infrastructure damage, and they can also affect perception about future flood risk. It is difficult to capture all of these costs simultaneously, so many papers in the field focus on changes in property sale prices as a proxy for some of the costs of flooding. The are multiple advantages to this strategy: data on post-flood property sale prices are widely available, and will capture both changes to the value of physical structures and changes in the perception of flood risk.

However, using changes in property sale prices to estimate the cost of flooding has its drawbacks. First, post-flood property sale prices will reflect changes in the physical state of buildings (physical damage net repairs) as well as changes to perceived risk. These effects can work in opposite directions, as property owners may receive insurance payouts and renovate their homes before putting them on the market. Secondly, it is difficult to precisely estimate which buildings are flooded; most data on flood events provide a coarse estimate of flood boundaries. An imprecise estimate of flood boundaries will lead to bias when estimating the impact of a flood on property values. If many non-flooded homes are marked as flooded (as often occurs when measuring floods at the municipality or county level, for example), the measured effect will be a weighted average of the effects of the flood on flooded homes and the effects of a flood on non-flooded homes, which may be very different. This paper demonstrates that it is crucial to precisely measure flooding and to decompose changes in property sale price due to risk perception from changes due to the physical state of a property.

In this paper, I am able to tackle both the issue of imprecision and the issue of multiple forces affecting house prices. I have extremely precise flood outline data, which I combine with residential property sale data to measure flooding at the house level. Thus I greatly reduce classical measurement error, and I am able to compare homes very close to one another that have different flood status. Secondly, I restrict my sample to homes that should not have differences in post-flood risk perception, and so any price differential will be due to flood damages and post-flood repairs. Immediately after a flood, buildings should clearly be less valuable due to mold damage and damage to the structure and electrical system of a property. By the time a property is sold, however, it may have been renovated to a level above its pre-flood quality. When keeping post-flood risk perception constant, I find that homes in fact increase in value after a flood. This suggests that the average flooded home in my sample is renovated beyond its pre-flood quality before being sold. Since the sign of damages net repairs is positive, estimates that do not parse out willingness to pay from the physical value of the home are underestimating the impact of a flood on property values.

In order to control for the effect of post-flood risk updating on the value of flooded and non-flooded homes, I limit my sample to homes that are very close to a flood boundary (within 100-1000 meters from a flood, depending on the specification). If flood boundaries are random, and homebuyers are well-informed about flood boundaries, then homes just outside of a flood boundary should face the same changes in homebuyer risk perception as homes just inside a flood boundary. I then conduct a regression discontinuity analysis around the flood boundary. I find that, after a flooded, homes just inside a flood boundary are actually worth approximately 2% - 7% *more* than homes just outside a flood boundary. This suggests that the value of flood damages net repairs is positive, and the average flooded home is updated and renovated before being sold.

Of course, this regression discontinuity analysis assumes that flood boundaries are random. Flood boundaries are not entirely random; the boundaries clearly depend on elevation and distance to a river or coast, which are correlated with amenity values. I test whether it is random for a home to be just inside or just outside a flood boundary by looking at whether other property characteristics are correlated with flood status, and discuss these results in section 1.6. I find that, after a flood, there is discontinuity across flood boundaries in the proportion of apartments versus detached homes and in the proportion of new properties. However, there is little to no discontinuity in these variables before a flood occurs. I argue that flood boundaries are sufficiently random to interpret my results as the true value of damages net repairs for a flooded home.

In section 3.2 of this paper, I present a simple hedonic model to explain the multiple factors that can affect post-flood property sale prices. Section 3.3 provides context for how this model is helpful for interpreting results from previous literature in the field of hedonic valuation of the costs of natural disasters. Sections 3.4 and 3.5 outline the data and methods I use, respectively. Section 3.6 presents my results, and section 3.7 concludes the paper.

3.2 Background

In the Rosen model, goods can be described using a vector of their characteristics $z = (z_1, z_2, ...z_n)$, where z_i represents the amount of the ith characteristic in the good. The various prices of versions of the good with different vectors z implicitly reveal the value of each characteristic using the function $p(z) = p(z_1, z_2, ...z_n)$. The first derivative $\partial p/\partial z_i$ represents the implicit market value for an additional unit of characteristic z_i.

After a natural disaster event such as a flood, the change in value of a flooded home will be a function of the characteristics z that change after a flood. These include flood damages or other direct effects on the physical structure (d), post-flood housing repairs and renovations (h), and the market's updated risk beliefs about the home (r). The total change in market value can be represented by the following equation:

$$\begin{aligned}
\Delta p_z &= p_{t=1}(z_{t=1}) - p_{t=0}(z_{t=0}) \\
&= p_{z,t=1}(d_{t=1}, h_{t=1}, r_{t=1}, z_1, ...z_n) - p_{z,t=0}(d_{t=0}, h_{t=0}, r_{t=0}, z_1, ...z_n) \\
&= \Delta p_z(\Delta d, \Delta h, \Delta r) \\
&= \frac{\partial p_z}{\partial d} \cdot \Delta d + \frac{\partial p_z}{\partial h} \cdot \Delta h + \frac{\partial p_z}{\partial r} \cdot \Delta r \qquad (3.1)
\end{aligned}$$

where $t = 0$ is the period before the flood and $t = 1$ is the period after the flood. The model above assumes that there are no interaction effects with other variables in z when d, h, and/or r change. The new value of the home will depend on the sum of $\frac{\partial p_z}{\partial d} \cdot \Delta d$, $\frac{\partial p_z}{\partial h} \cdot \Delta h$, and $\frac{\partial p_z}{\partial r} \cdot \Delta r$ when moving from $t = 0$ to $t = 1$. We should expect $\frac{\partial p_z}{\partial d} \cdot \Delta d$, to be negative, as floods and other natural disasters tend to cause damage that lowers the value of a home. Based on previous literature about risk updating, we should also expect $\frac{\partial p_z}{\partial r} \cdot \Delta r$ to be negative, as risk tends to become more salient to consumers after a disaster event (Bin and Polasky (2004); Bin and Landry (2013b); Kousky (2010); Rambaldi et al. (2013); Atreya et al. (2013)). We expect $\frac{\partial p_z}{\partial h} \cdot \Delta h$ to be zero or positive, and the size of $\frac{\partial p_z}{\partial h} \cdot \Delta h$ will depend on the contextual factors such as the local insurance regime, speed of post-disaster payouts, and the time since the disaster occurred. Thus the sign of Δp_z is undetermined.

Disaster events can also impact the price of similar homes that were not directly hit. Although $\Delta d = 0$ and $\Delta h = 0$ for homes that were not flooded, there could be a change in r via risk updating. Thus, for non-flooded homes:

$$\begin{aligned}
\Delta p_z &= p_{t=1}(z_{t=1}) - p_{t=0}(z_{t=0}) \\
&= p_{t=1}(r_{t=1}, d, h, z_1, ...z_n) - p_{t=0}(r_{t=0}, d, h, z_1, ...z_n) \\
&= \Delta p_z(\Delta r) \\
&= \frac{\partial p_z}{\partial r} \cdot \Delta r. \qquad (3.2)
\end{aligned}$$

We should expect $\frac{\partial p_z}{\partial r} \cdot \Delta r$ to be negative, as previous literature has shown that high-risk homes that were not hit also tend to decrease in value after a natural disaster (Bin and Polasky (2004); Bin and Landry (2013b); Naoi et al. (2009); Kousky (2010); Rambaldi et al. (2013); Atreya et al. (2013)).

The framework of this model is useful for outlining a few issues with using property sale prices as a measure of the costs of flooding or other natural disasters. First, the sale price of a flooded home will reflect damages (d) and risk updating (r), but also any post-flood renovations (h). Thus changes in the sale price of a home will underestimate the true costs of flooding ($\frac{\partial p_z}{\partial d} \cdot \Delta d + \frac{\partial p_z}{\partial r} \cdot \Delta r$) . Secondly, most analyses are not able to decompose the effect of risk updating ($\frac{\partial p_z}{\partial r} \cdot \Delta r$) from the effect of damages ($\frac{\partial p_z}{\partial d} \cdot \Delta d$). These two effects have very different policy implications; damages d reflect the direct impacts of the flood on the state of the housing stock, and thus are informative for estimates of the insurance payouts or future flooding costs. Changes in r and $\frac{\partial p_z}{\partial r} \cdot \Delta r$ reflect changes in beliefs about risk and willingness to pay to live in a high risk area, which is informative for understanding homebuyer behavior and flood salience. Finally, most analyses have fairly coarse estimates of whether a home was actually flooded. If, for example, $\frac{\partial p_z}{\partial d} \cdot \Delta d + \frac{\partial p_z}{\partial h} \cdot \Delta h > 0$ and $\frac{\partial p_z}{\partial r} \cdot \Delta r < 0$, and many homes that are not flooded are labeled as flooded, the estimated treatment effect for of a flood will be downwardly biased.

In this paper, I am not able to decompose $\frac{\partial p_z}{\partial d} \cdot \Delta d$ from $\frac{\partial p_z}{\partial h} \cdot \Delta h$, but I am able to isolate the effect of $\frac{\partial p_z}{\partial d} \cdot \Delta d + \frac{\partial p_z}{\partial h} \cdot \Delta h$ from $\frac{\partial p_z}{\partial r} \cdot \Delta r$. By looking only at homes just inside or outside a flood boundary, I study homes that should have the same or similar $\frac{\partial p_z}{\partial r} \cdot \Delta r$. Any treatment effect I find will be the sum of $\frac{\partial p_z}{\partial d} \cdot \Delta d$ and $\frac{\partial p_z}{\partial h} \cdot \Delta h$. In order to conduct this analysis, I need to be able to correctly measure whether homes close to a flood boundary are flooded or not. Fortunately, the flood boundary data provided by the UK Environment Agency allow me to estimate flood status at the property level. Thus I should have very little measurement error in my treatment variable, and little bias in my estimate of a treatment effect.

3.3 Literature Review

This paper contributes to the literature on using property prices to estimate the costs of flooding and other natural disasters. There are a number of papers that find surprisingly small or positive effects of disasters on property values. Skantz and Strickland (1987) find that homes in Houston that were flooded do not decrease in price immediately after the flood. Beron et al. (1997) find that Bay Area homes with high earthquake risk have lower prices, but this effect is *smaller* after the massive Loma Prieta earthquake. Smith et al. (2006) finds negative effects of being in an area at high risk of hurricanes, but no impact of actual hurricane damage on house prices. It is possible that these results are due to changes in risk perception for non-damaged homes, or imprecision in the measure of which homes are directly affected by natural disasters.

Of course, there are many papers that find a negative impact of disasters on property values. Many researchers have found that house prices drop after flood events (Bin and

Polasky (2004); Bin and Landry (2013b); Kousky (2010); Rambaldi et al. (2013); Atreya et al. (2013)). Similar results exist for homes after fracking-related earthquakes (Koster and Van Ommeren, 2015) and nuclear contamination (Yamane et al., 2013). I argue that, by not separating the effects of risk updating from the effects of post-disaster repairs, these papers underestimate the true costs of natural disasters.

There are some papers that do not conflate physical damage and risk perception, typically by studying the effect of disasters on homes that are not directly damaged. Naoi et al. (2009) examine post-earthquake home values for "near-miss" homes; that is, homes that are at high earthquake risk but are not in the area hit by a quake. They find that "near-miss" homes decrease in value, suggesting that risk updating can decrease nearby home values after a natural disaster. Other papers looking at "near-miss" homes near wildfires find increases (Mueller et al., 2009) or decreases (Hansen and Naughton, 2013) in sale prices, although both of these results may be due to changes in amenity values. There are also papers that examine the effects of risk updating though a treatment of pure information: earthquake and volcano hazard notices tend to decrease property values (Bernknopf et al., 1990), as does the introduction of wildfire risk maps (Donovan et al., 2007). These papers inform us that the housing market is subject to risk updating when disaster events become more salient. If both treated and non-treated homes experience a change in estimated risk ($\frac{\partial p_z}{\partial r} \cdot \Delta r \neq 0$), and $\frac{\partial p_z}{\partial r} \cdot \Delta r$ is not the same across treated and control units, the estimated treatment effect of a natural disaster will be biased.

Although the papers listed above are effective in decomposing damages net repairs ($\frac{\partial p_z}{\partial d} \cdot \Delta d + \frac{\partial p_z}{\partial h} \cdot \Delta h$) from risk updating ($\frac{\partial p_z}{\partial r} \cdot \Delta r$), they do this via examining price changes for homes that are *not* hit by natural disasters. There are very few papers that control for risk updating and look only at the effect of disasters on damaged homes. Daniel et al. (2009) do this by studying home prices along the Meuse River in the Netherlands. The authors use aerial photos to gain a relatively accurate estimate of which homes were flooded in 1993 and in 1995, and find that inundated homes are worth 7.4% less than other homes after a flood. However, they compare values of flooded homes to all other homes within riverbank municipalities. The flood should bring down the value of some but not all non-flooded homes, and not at the same level as within the flood zone, thus biasing this estimate of flood damages. There is one other paper that tries to deal with this issue: Atreya and Ferreira (2015) look at property sales of 2,685 homes in Albany, Georgia before and after a 1994 flood. The authors use USGS flood inundation maps to measure which homes are flooded, and compare value of these homes to other homes within Albany, all of which are within two miles of the inundation zone. They find that flooded homes are worth 45% - 64% less than comparable homes in the year after a flood, although this effect decreases over time. The difference between these results and ours may be due to the fact that their study region is a low-income area, which may have contextual differences about post-flood insurance and disaster relief payouts. My paper follows a similar methodology, but in a broader region with many more observations, and contributes an additional (and contrary) estimate to this literature.

3.4 Data

The data on actual flood events are published by the UK Environment Agency. Since 1946, the Environment Agency has been tracking flood events within England using aerial photography and in-person surveyors. These flood records include shapefiles with precise outlines of each flood, and so are at a much higher geographic resolution than is available in most countries. In the United States, for example, flood history is usually measured at the county level. The flood outline data I use include the start and end date of flooding, the source of the flood (from a river, groundwater, etc), and the geographical extent of flooding. I use flood outlines through November 2017. There are 27,924 separate flood outline shapes in the data, but many of these are associated with the same flood event. I combine flood shapes by event code to identify 1,966 separate flood events. These flood events are mapped in Figure 3.1. One can see that floods are common in some coastal areas, but there are also many floods along rivers and in low-lying floodplains. Thus there is a good deal of geographic diversity among the homes used in my analysis.

Data on home sale prices and house characteristics are provided by Her Majesty's Land Registry. The UK government records all residential property sales for full market value since January 1, 1995. (Homes are not sold for full market value if they are transferred during an inheritance or a divorce, for example.) The Price Paid Data include the property address, date of sale, sale price (in contemporaneous British £), type of property (detached, semi-detached, flat, etc), whether the sale is freehold or leasehold[1], and whether the property is newly built or an established residential building. These data are periodically updated, and we include transactions through February 2018.

I use the addresses in the Price Paid Data to geocode the location of each property. I then use the geographic coordinates of each property to estimate whether it is inside or outside each flood boundary. I limit my sample spatially to homes that are within 1000 meters of a buffer around the flood boundary. For each unit within the flood boundary buffer, I include all sales of that property that happen during or after the flood. (That is, property sales before the flood occurred are not considered). The resulting dataset includes 8,700,200 unique residential property sales for 2,910,893 unique homes.

Table 3.1 presents the mean and standard deviation of variables of interest. The main variable of interest is the property sale price, which is converted to 2017 £. The data included in this table are the same subset of data used in my analysis; that is, sales for properties that are within one kilometer of a flood that happened prior to the sale. The timing of the sales is fairly evenly distributed; the mean sale year is approximately the median year of data (Jan 1995- Feb 2018), and, surprisingly, sales are fairly stable across seasons. Most mortgages are for freehold properties, and the proportion of leasehold mortgages is close to the proportion of properties that are flats, which are the typical source of leasehold sales in England. If we look only at homes that are actually flooded, there are far fewer (approximately 3.6% of

[1]A "leasehold" sale gives the buyer ownership of a building or unit itself for a fixed amount of time, but not the land underneath. The leases are typically for 99 years. A "freehold" sale does not have this stipulation, and is analogous to a typical sale in the United States.

the data), and the average flooded home is inundated 1.26 times during the sample period (1946-present). Most home flooding in this dataset occurs during winter, which is the rainy season in England. We see that the mean year of flooding in the data is 1962, which is far less than the median year of flood recording (1946-present). This is likely because floods only appear in the data if they occur before a home sale, and all floods that occur before home sales appear in this subset of the data. When conducting my analysis, I also try restricting the dataset to only include floods that occur 10, 5, or 1 years before a sale.

3.5 Methods

First, I create plots to observe whether there is a discontinuity in my data at the threshold point, where distance to a flood boundary is zero. The main plot is created using the following model:

$$ln(p_i) = \beta_0 + \beta_1 \cdot T_{if} + \beta_2 \cdot dist_{if} + \beta_3 \cdot T_i \cdot dist_{if} + \varepsilon_i, \tag{3.3}$$

where i represents an individual sale, p_{it} is the sale price of the property, in 2017 £, T_{it} is an indicator variable for whether the home was inside or outside the flood boundary, and $dist_{if}$ is the distance to the flood boundary. This plot only include homes sold between 0 and 10 years after a flood occurs. To test whether the threshold is random, I also create plots using the model:

$$z_i = \beta_0 + \beta_1 \cdot T_{if} + \beta_2 \cdot dist_{if} + \beta_3 \cdot T_i \cdot dist_{if} + \varepsilon_i, \tag{3.4}$$

where z_i is a property characteristic during an individual sale, such as the type of home (detached, semi-detached, apartment, terraced, or other) and whether the property is newly built. Finally, I test my results by creating a plot using equation 3.3 but only using home sales that occur *before* a flood. We should not expect there to be a discontinuity in prices before a flood occurs.

Along with creating regression discontinuity plots, I, run regressions using the following model:

$$ln(p_{if}) = \beta_0 + \beta_1 \cdot T_{if} + \beta_2 \cdot dist_{if} + \beta_3 \cdot T_i \cdot dist_{if} + \gamma_i + \lambda_f + \kappa_s + \rho_t + \varepsilon_{if}, \tag{3.5}$$

where i represents an individual sale, and f represents a single flood. Thus the unit of observation is at the sale-flood level, and property sales appear multiple times in the data if the property was flooded multiple times. p_{it} is the sale price of the property, in 2017 £, T_{it} is an indicator variable for whether the home was inside or outside the flood boundary, and $dist_{if}$ is the distance to the flood boundary. The baseline regression includes only homes within a 1000 meter buffer around each flood, and only includes property sales that occur between 0 and 10 years after the flood. γ_i, λ_f, κ_s,and ρ_t are sale-, flood-, spatial-, and time-specific fixed effects, respectively. For most regressions, these are a district fixed effect, a flood-specific fixed effect, month of sample fixed effects, and controls for the property type (detached, semi-detached, flat, etc), whether the property is new construction, and whether the sale is freehold or leasehold.

3.6 Results

Figure 3.2 depicts logged home sale price for homes within 1000 meters inside or outside of a flood boundary. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary. Only property sales that occur between 0 and 10 years after a flood are included in this plot. We see that, although home prices increase slightly as distance to the flood boundary decreases, there is a sharp jump for homes just inside versus just outside the boundary. This is also true when I plot home price in levels (2017 £) for homes just inside and outside a flood boundary, as can be seen in Figure 3.3.

Flood risk is correlated with geographic features such as elevation and distance to coast, so it is certainly possible that this jump is due to confounding factors. To test for this, first I plot property characteristics by distance to a flood boundary for homes sold after a flood. In Figures 3.4, 3.5, and 3.6 we see that homes just inside a flood boundary are more likely to be apartments versus detached homes, and more likely to be brand new construction. This is concerning, as it suggests that the flood boundary is not random but is instead highly correlated with property characteristics that can affect market value. However, because these plots are for home sales after a flood, it is possible that the flood itself affects property characteristics by damaging buildings and incentivizing different types of post-flood construction. To test for this, I plot percent apartments, percent detached homes, and percent new homes *before* a flood occurs in Figures 3.7, 3.8, and 3.9. We see that, before floods occur, properties just inside a flood boundary are only slightly more likely to be apartments, and are not less likely to be detached homes or more likely to be new homes. I also examine logged home price by distance to flood boundary for homes sold before a flood, in Figure 3.10. The price of homes inside a flood boundary is not statistically significantly different from the price of homes outside a flood boundary. However, these property sales before a flood occurs happen earlier in time, and so the differences between Figures 3.2, 3.4, 3.5, and 3.6 versus Figures 3.7, 3.8, 3.9, and 3.10 could be due to compositional changes of flood-prone areas over time.

Table 3.2 presents results from my main regression specification. The regressions in this table include all property sales that within 1000 meters of a flood, and happen less than 10 years after a flood occurs. Column (1) includes all sales that occur within 10 years after the flood, column (2) includes all sales that occur within 5 years after the flood, and column (3) includes all sales that occur within 1 year after the flood. All regressions also have district fixed effects, month of sample fixed effects, flood event fixed effects, and controls for the property type (detached, semi-detached, flat, terraced, or other), mortgage type (freehold versus leasehold), and whether the property is newly built. Standard errors are in parentheses, and are two-way clustered by postcode and month of sample. In this table, we see that homes that are flooded and then sold are worth 3.1% - 6.8% more than similar properties just outside the flood boundary. This is possible if the effect of renovations after a flood outweighs the effect of post-flood risk updating.

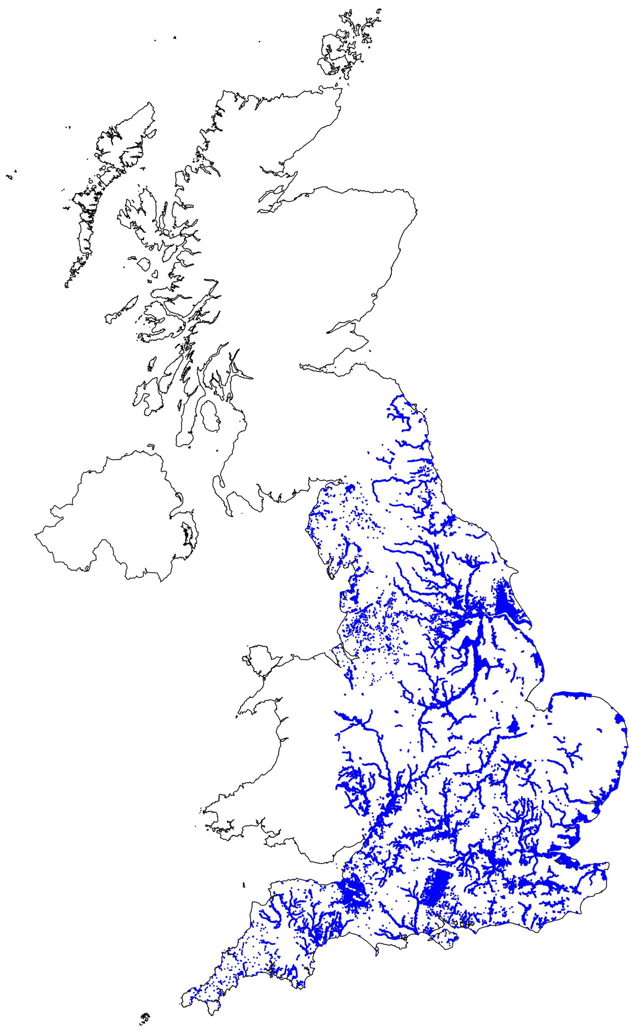
In Tables 3.3, 3.4, and 3.5 I check whether these results are sensitive to the spatial bandwidth or time period that I choose. Table 3.3 includes all homes sold within 10 years after a flood, Table 3.4 includes all homes sold within 5 years after a flood, and Table 3.5 includes all homes sold within 1 year after a flood. The columns in each table represent decreasing spatial bandwidth around the flood boundary. We see in all of these tables that, after a flood, homes that are flooded and then sold are worth approximately 2-8% more than homes that are sold after a flood but are just outside of a flood boundary. The statistical significance of these results does drop as I lose observations due to an increasingly restricted sample. Overall, it appears that the net value of post-flood damages and post-flood repairs is positive, which suggests that, on average, homes that are sold after a flood are repaired and renovated before being put on the market. Thus using property sale prices to estimate flood damages will lead to an underestimation of the true impacts of flooding.

3.7 Conclusion

There is a rich literature on estimating the costs of and reactions to natural disasters using property sale prices. In this paper, I demonstrate that property sale prices can be problematic, as after a disaster there are forces that push house prices in different directions. Property damages should decrease the value of affected homes, post-disaster renovations should increase the value of the subset of affected homes that are renovated before being sold, and risk updating should decrease home values for both affected homes and homes that are not affected but face similar risk profiles. By using unusually high resolution flood data, and isolating the effect of damages net repairs from the effect of risk updating, I find that, on average, homes that are sold after a flood actually *increase* in value. From these results I conclude that the effect of post-flood insurance payouts and subsequent renovation outweighs the effect of direct flood damages, at least for the subset of homes that are sold. Because the net effect of damages and repairs is positive, many other papers that use post-flood property values as an estimate of damages or consumer preferences likely underestimate the true cost of flooding.

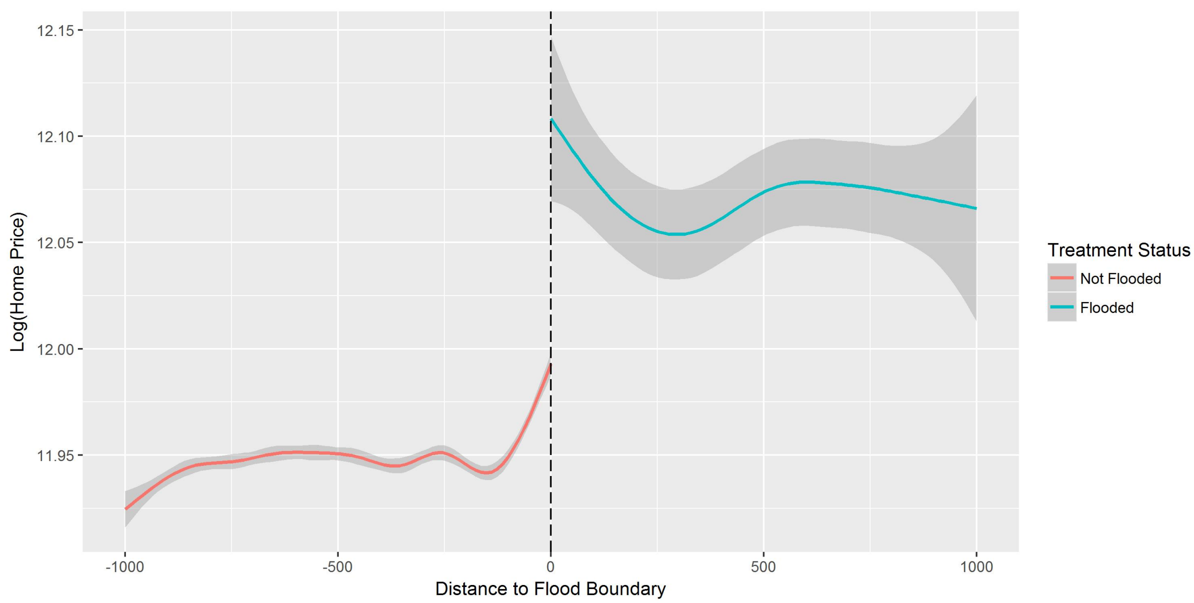
Further research is needed to understand whether this is true under different insurance or disaster relief regimes. In the UK, all homeowners must have flood insurance, which facilitates post-flood renovation by providing homeowners with an injection of capital. Nations such as the United States with different insurance laws or other barriers to renovation might not see a net positive increase in home values after a flood. For example, Atreya and Ferreira (2015) conduct a similar analysis and find a net negative impact of a flood on property values in the relatively low-income town of Albany, Georgia. Additionally, further work is needed to understand whether my results are due to flood boundaries being non-random, and correlated with other property characteristics. Property characteristics do not have a discontinuity at the flood boundary *before* homes are flooded, but this may be due to changes in the composition of home values over time, and not due to floods themselves. Overall, my work suggests that researchers must think carefully about the multiple competing forces that affect property values after a natural disaster.

Figure 3.1: Historical Extent of Flooding in England



Notes: The figure above depicts in blue the areas that have ever been flooded in England since 1946, as reported by the UK Environment Agency. Wales, Scotland, and Northern Ireland are not included in this dataset. The data I use have shapefiles for all recorded floods from January 1946 through November 2017. There are 27,924 separate flood outline shapes in the data, but many of these are associated with the same flood event. I combine flood shapes by event code to identify 1,966 separate flood events.

Figure 3.2: Regression Discontinuity Plot: Log Property Sale Price by Flood Status

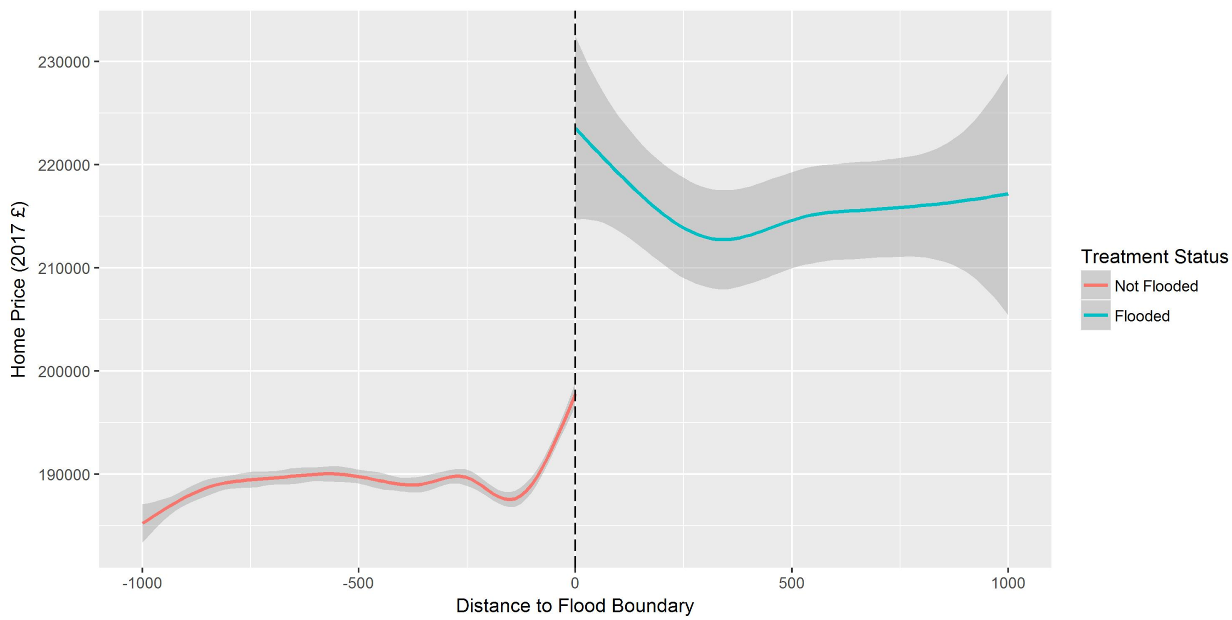


Notes: The figure above depicts logged home sale price (in 2017 £) for homes just inside and outside a flood boundary. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary. There are relatively few homes inside of a flood but 1000 meters away from a flood boundary, so the standard errors for flooded homes increase as distance from the flood boundary increases.

Table 3.1: Summary Statistics of Properties Sold

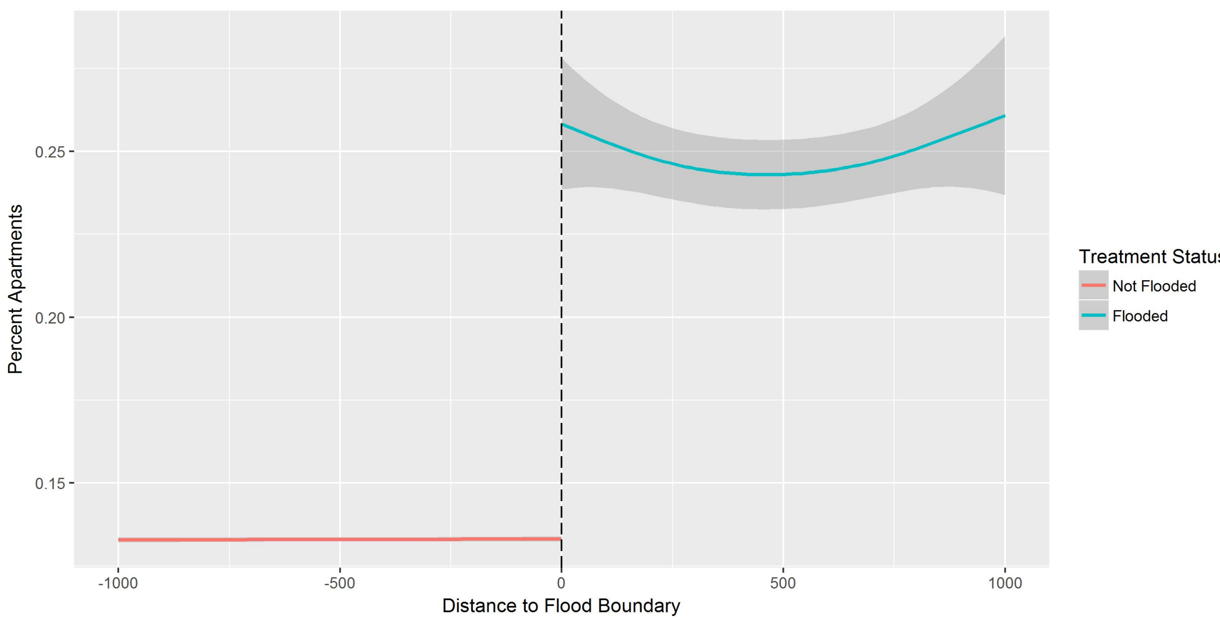
Property Characteristics	Mean
Sale Characteristics	
Number of Observations	8,700,200
Price (2017 £)	236,681.18
	(483,793.76)
New Sale	0.10
	(0.00)
Year of Sale	2005.42
	(6.56)
Sold in Winter	0.22
	(0.00)
Sold in Spring	0.24
	(0.00)
Sold in Summer	0.28
	(0.00)
Sold in Autumn	0.26
	(0.00)
Mortgage Type	
Freehold	0.73
	(0.00)
Leasehold	0.25
	(0.00)
Property Type	
Detached	0.20
	(0.00)
Semi-Detached	0.26
	(0.00)
Terraced	0.33
	(0.00)
Flat	0.19
	(0.00)
Other	0.01
	(0.00)
Flooded Homes Only	
Number of Observations	309,963
Number of Floods per Home	1.26
	(0.65)
Flood Year	1962.18
	(15.70)
Flooded in Winter	0.77
	(0.00)
Flooded in Spring	0.10
	(0.00)
Flooded in Summer	0.01
	(0.00)
Flooded in Autumn	0.12
	(0.00)

Figure 3.3: Regression Discontinuity Plot: Property Sale Price by Flood Status



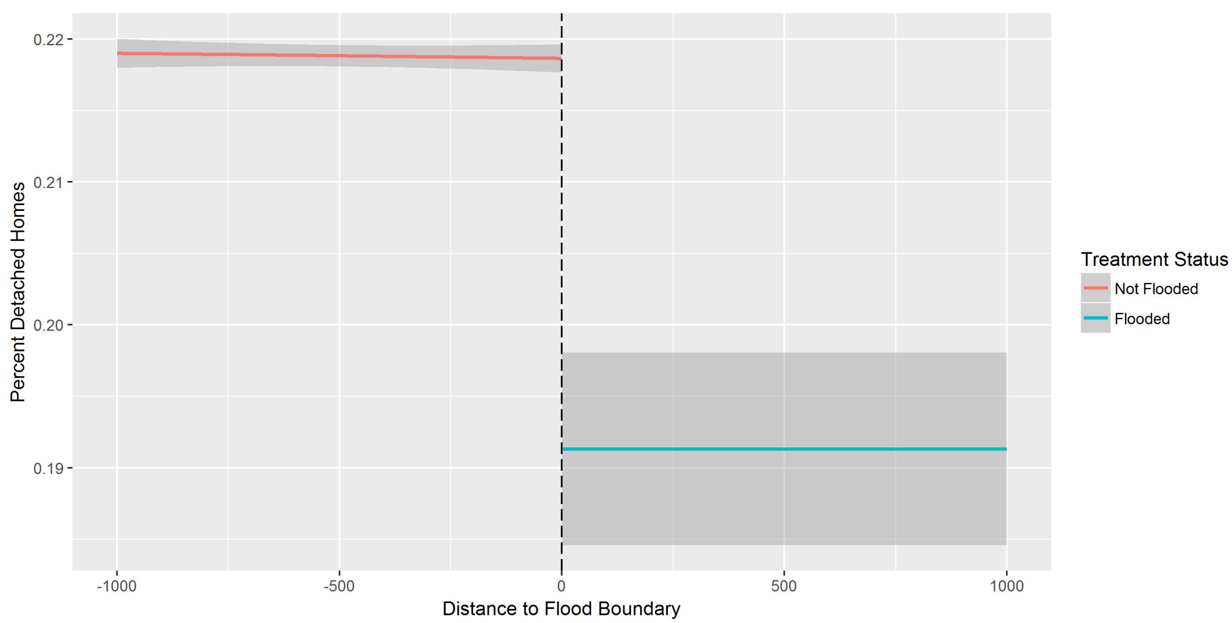
Notes: The figure above depicts home sale price (in 2017 £) for homes just inside and outside a flood boundary. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary. There are relatively few homes inside of a flood but 1000 meters away from a flood boundary, so the standard errors for flooded homes increase as distance from the flood boundary increases.

Figure 3.4: Proportion of Homes that are Apartments, by Distance to a Flood Boundary



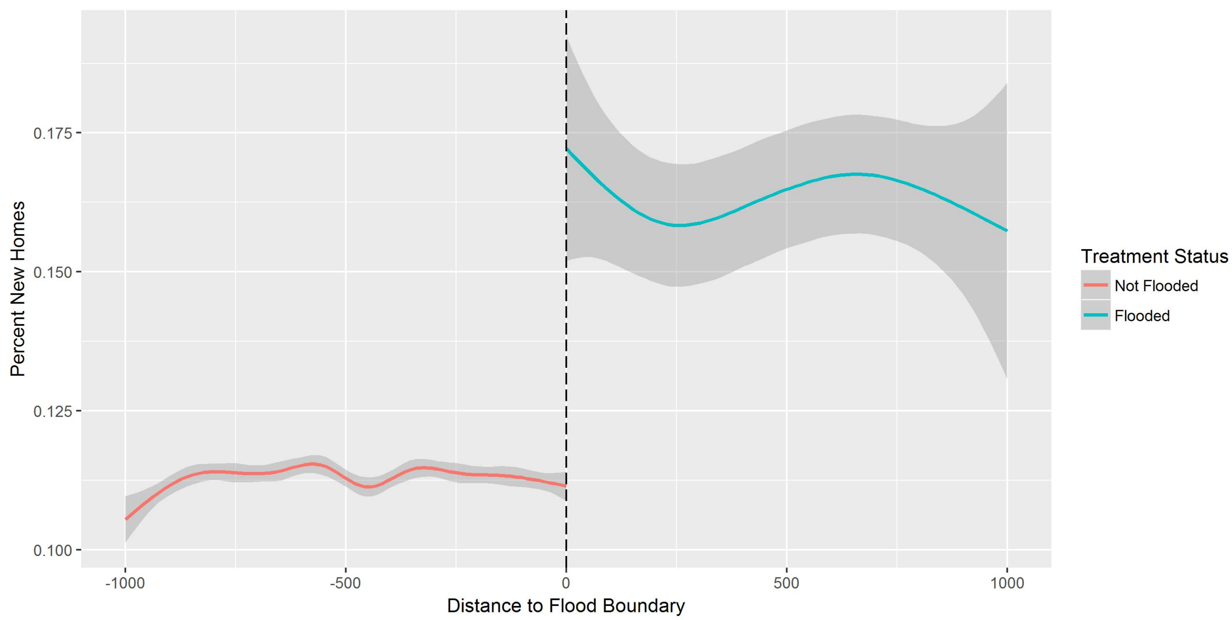
Notes: The figure above depicts the portion of homes that are apartments, for homes just inside and outside a flood boundary. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary.

Figure 3.5: Proportion of Detached Homes, by Distance to a Flood Boundary



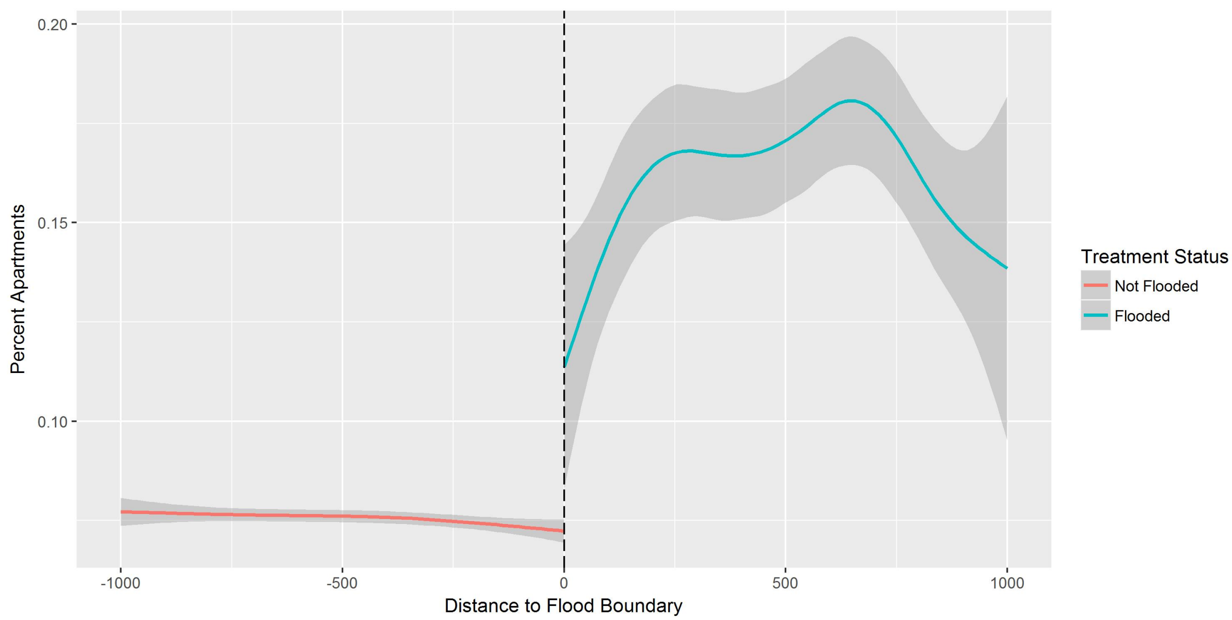
Notes: The figure above depicts the portion of homes that are detached homes, for homes just inside and outside a flood boundary. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary.

Figure 3.6: Proportion of New Homes, by Distance to a Flood Boundary



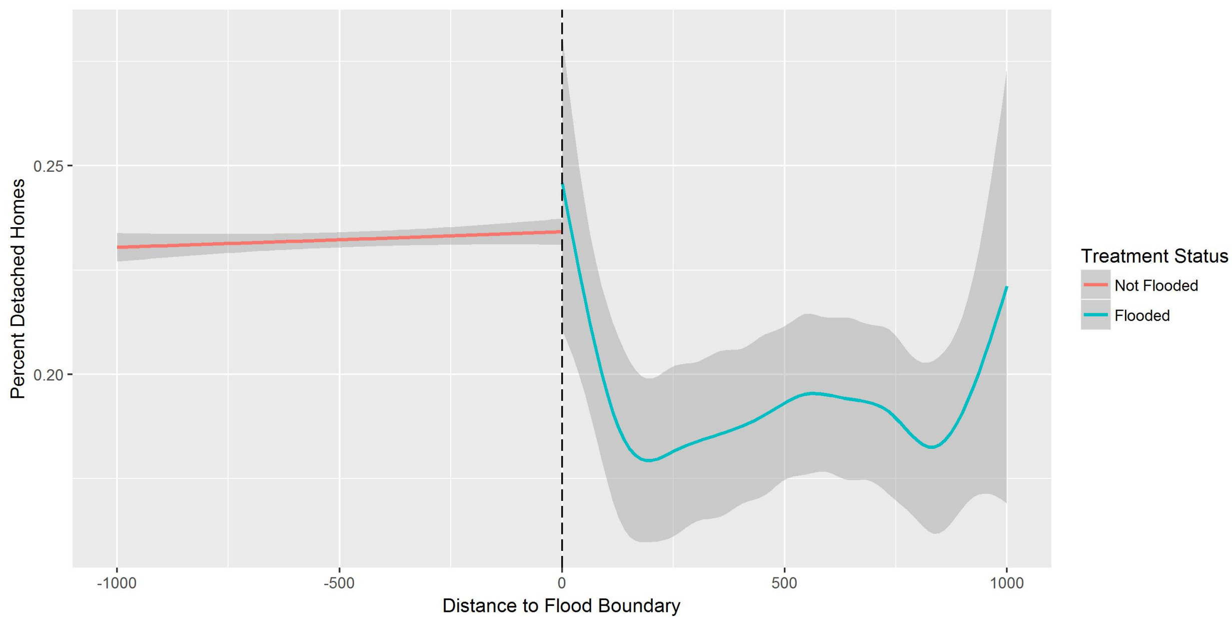
Notes: The figure above depicts the portion of homes that are brand new construction, for homes just inside and outside a flood boundary. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary.

Figure 3.7: Proportion of Homes that are Apartments before a Flood, by Distance to a Flood Boundary



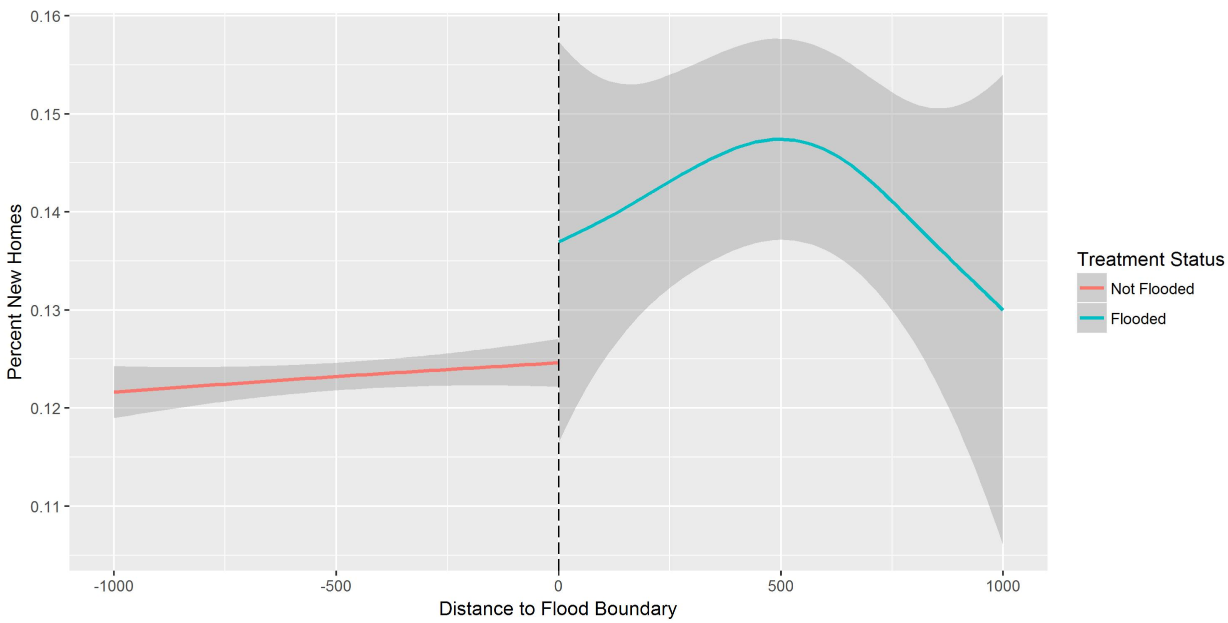
Notes: The figure above depicts the portion of homes that are apartments, for homes just inside and outside the boundary that have *not* yet been flooded. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary.

Figure 3.8: Proportion of Detached Homes before a Flood, by Distance to a Flood Boundary



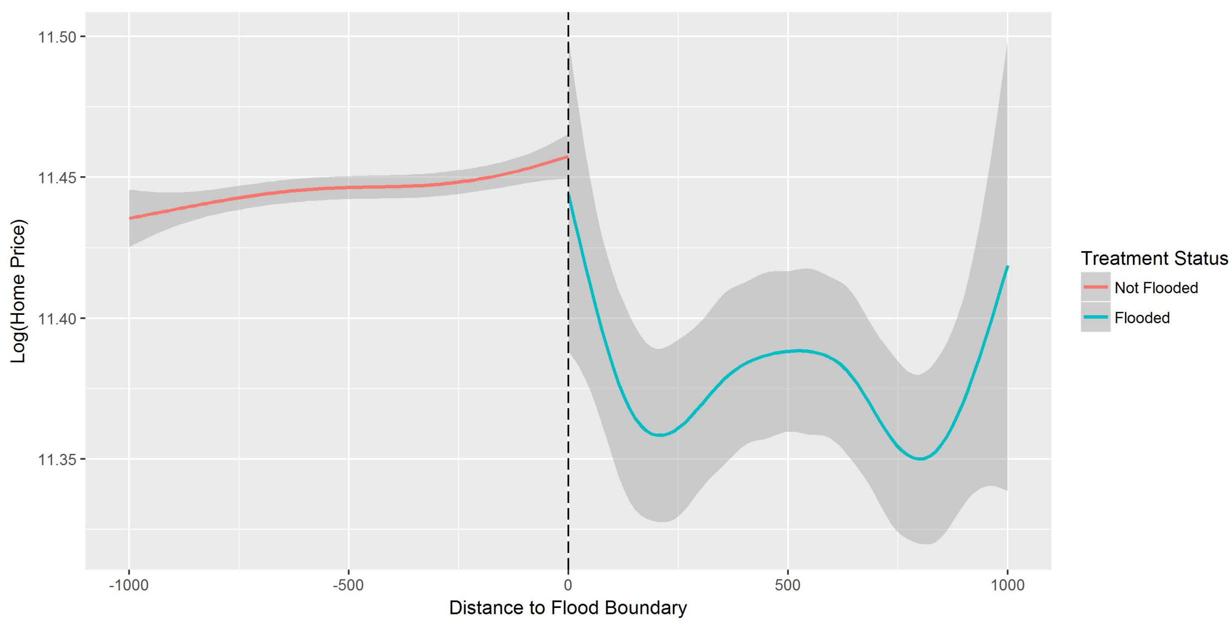
Notes: The figure above depicts the portion of homes that are detached, for homes just inside and outside the boundary that have *not* yet been flooded. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary.

Figure 3.9: Proportion of New Homes before a Flood, by Distance to a Flood Boundary



Notes: The figure above depicts the portion of homes that are brand new construction, for homes just inside and outside the boundary that have *not* yet been flooded. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary.

Figure 3.10: Log Property Sale Price by Distance to a Flood Boundary, Homes Not Yet Flooded



Notes: The figure above depicts logged home sale price (in 2017 £) for homes just inside and outside a flood boundarythat have *not* yet been flooded. The 95% confidence interval for these estimates is depicted in gray. Properties to the left of the dashed line are outside the flood boundary, and properties to the right of the dashed line are inside the flood boundary.

Table 3.2: Regression Discontinuity Results, All Sales within 1000 Meters

	Dependent variable: Log (Price)		
Maximum Years Since Flood	10	5	1
	(1)	(2)	(3)
Flooded	0.045***	0.031*	0.068**
	(0.013)	(0.018)	(0.035)
Distance to Flood Boundary	−0.00000	0.00000	0.00001
	(0.00000)	(0.00000)	(0.00000)
Flooded × Distance to Flood Boundary	0.00001	0.00003	0.00003
	(0.00002)	(0.00003)	(0.0001)
Observations	1,233,207	631,901	112,653
District FE	Y	Y	Y
Month of Sample FE	Y	Y	Y
Flood FE	Y	Y	Y
Sale Controls	Y	Y	Y
R^2	0.658	0.659	0.638
Adjusted R^2	0.658	0.658	0.635

Notes: This table shows the results of a regression of residential property sale price, in 2017 £, on the traditional regression discontinuity variables of treatment status (flooded versus non-flooded), distance to the flood boundary, and distance × treatment status. All sales within 1000 meters of a flood are included. Column (1) includes all sales that occur within 10 years after the flood, column (2) includes all sales that occur within 5 years after the flood, and column (3) includes all sales that occur within 1 year after the flood. All regressions also have district fixed effects, month of sample fixed effects, flood event fixed effects, and controls for the property type (detached, semi-detached, flat, terraced, or other), mortgage type (freehold versus leasehold), and whether the property is newly built. Standard errors are in parentheses, and are two-way clustered by postcode and month of sample. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table 3.3: Regression Discontinuity Results, Sales within 10 Years After Flood

	Dependent variable: Log (Price)			
Bandwidth (m)	750	500	250	100
Flooded	0.043***	0.035**	0.036**	0.031
	(0.013)	(0.015)	(0.016)	(0.020)
Distance to Flood Boundary	−0.00000	0.00000	−0.00000	−0.00000
	(0.00000)	(0.00000)	(0.00001)	(0.00003)
Flooded × Distance to Flood Boundary	0.00002	0.00005	0.0001	0.0001
	(0.00002)	(0.00004)	(0.0001)	(0.0004)
Observations	952,780	614,498	292,976	118,476
District FE	Y	Y	Y	Y
Month of Sample FE	Y	Y	Y	Y
Flood FE	Y	Y	Y	Y
Sale Controls	Y	Y	Y	Y
R^2	0.659	0.659	0.658	0.663
Adjusted R^2	0.658	0.659	0.656	0.659

Notes: This table shows the results of a regression of residential property sale price, in 2017 £, on the traditional regression discontinuity variables of treatment status (flooded versus non-flooded), distance to the flood boundary, and distance × treatment status. All sales within 10 years after a flood are included. Columns (1) through (4) include different bandwidths of data; column (1) includes all sales within 750 meters from the flood, column (2) includes all sales within 500 meters from the flood, column (3) includes all sales within 250 meters from the flood, and column (4) includes all sales within 100 meters from the flood. All regressions also have district fixed effects, month of sample fixed effects, flood event fixed effects, and controls for the property type (detached, semi-detached, flat, terraced, or other), mortgage type (freehold versus leasehold), and whether the property is newly built. Standard errors are in parentheses, and are two-way clustered by postcode and month of sample. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table 3.4: Regression Discontinuity Results, Sales within 5 Years After Flood

	Dependent variable: Log (Price)			
Bandwidth (m)	750	500	250	100
Flooded	0.028	0.024	0.045**	0.036
	(0.020)	(0.022)	(0.021)	(0.026)
Distance to Flood Boundary	−0.00000	0.00000	0.00001	0.0001**
	(0.00000)	(0.00000)	(0.00001)	(0.00004)
Flooded × Distance to Flood Boundary	0.00004	0.00005	−0.0002	−0.00003
	(0.00004)	(0.0001)	(0.0002)	(0.001)
Observations	488,575	317,106	152,127	61,071
District FE	Y	Y	Y	Y
Month of Sample FE	Y	Y	Y	Y
Flood FE	Y	Y	Y	Y
Sale Controls	Y	Y	Y	Y
R^2	0.659	0.660	0.655	0.664
Adjusted R^2	0.658	0.659	0.653	0.658

Notes: This table shows the results of a regression of residential property sale price, in 2017 £, on the traditional regression discontinuity variables of treatment status (flooded versus non-flooded), distance to the flood boundary, and distance × treatment status. All sales within 5 years after a flood are included. Columns (1) through (4) include different bandwidths of data; column (1) includes all sales within 750 meters from the flood, column (2) includes all sales within 500 meters from the flood, column (3) includes all sales within 250 meters from the flood, and column (4) includes all sales within 100 meters from the flood. All regressions also have district fixed effects, month of sample fixed effects, flood event fixed effects, and controls for the property type (detached, semi-detached, flat, terraced, or other), mortgage type (freehold versus leasehold), and whether the property is newly built. Standard errors are in parentheses, and are two-way clustered by postcode and month of sample. Significance: *p<0.1; **p<0.05; ***p<0.01.

Table 3.5: Regression Discontinuity Results, Sales within 1 Year After Flood

	Dependent variable: Log (Price)			
Bandwidth (m)	750	500	250	100
Flooded	0.064*	0.070*	0.031	0.079
	(0.036)	(0.036)	(0.020)	(0.064)
Distance to Flood Boundary	−0.00000	−0.00001	−0.00000	0.0001
	(0.00001)	(0.00001)	(0.00003)	(0.0001)
Flooded × Distance to Flood Boundary	0.0001	0.0001	0.0001	−0.0003
	(0.0001)	(0.0001)	(0.0004)	(0.001)
Observations	87,440	57,886	118,476	11,821
District FE	Y	Y	Y	Y
Month of Sample FE	Y	Y	Y	Y
Flood FE	Y	Y	Y	Y
Sale Controls	Y	Y	Y	Y
R^2	0.640	0.643	0.663	0.651
Adjusted R^2	0.636	0.638	0.659	0.627

Notes: This table shows the results of a regression of residential property sale price, in 2017 £, on the traditional regression discontinuity variables of treatment status (flooded versus non-flooded), distance to the flood boundary, and distance × treatment status. All sales within 1 year after a flood are included. Columns (1) through (4) include different bandwidths of data; column (1) includes all sales within 750 meters from the flood, column (2) includes all sales within 500 meters from the flood, column (3) includes all sales within 250 meters from the flood, and column (4) includes all sales within 100 meters from the flood. All regressions also have district fixed effects, month of sample fixed effects, flood event fixed effects, and controls for the property type (detached, semi-detached, flat, terraced, or other), mortgage type (freehold versus leasehold), and whether the property is newly built. Standard errors are in parentheses, and are two-way clustered by postcode and month of sample. Significance: *p<0.1; **p<0.05; ***p<0.01.

Bibliography

Adger, W Neil, Suraje Dessai, Marisa Goulden, Mike Hulme, Irene Lorenzoni, Donald R Nelson, Lars Otto Naess, Johanna Wolf, and Anita Wreford, "Are there social limits to adaptation to climate change?," *Climatic change*, 2009, *93* (3-4), 335–354.

Agarwal, Sumit, Itzhak Ben-David, and Vincent Yao, "Systematic mistakes in the mortgage market and lack of financial sophistication," *Journal of Financial Economics*, 2017, *123* (1), 42–58.

_ , Richard J Rosen, and Vincent Yao, "Why do borrowers make mortgage refinancing mistakes?," *Management Science*, 2015, *62* (12), 3494–3509.

Allcott, Hunt and Nathan Wozny, "Gasoline prices, fuel economy, and the energy paradox," *Review of Economics and Statistics*, 2014, *96* (5), 779–795.

Arnell, Nigel W, MJ Clark, and AM Gurnell, "Flood insurance and extreme events: the role of crisis in prompting changes in British institutional response to flood hazard," *Applied Geography*, 1984, *4* (2), 167–181.

Association of British Insurers, "ABI statement of principles on the provision of flood insurance," 2002.

_ , "UK Insurance and Long-Term Savings: The state of the market," 2017.

Atreya, Ajita and Susana Ferreira, "Seeing is believing? Evidence from property prices in inundated areas," *Risk Analysis*, 2015, *35* (5), 828–848.

_ , _ , and Warren Kriesel, "Forgetting the flood? An analysis of the flood risk discount over time," *Land Economics*, 2013, *89* (4), 577–596.

Auffhammer, Maximilian and Wolfram Schlenker, "Empirical studies on agricultural impacts and adaptation," *Energy Economics*, 2014, *46*, 555–561.

Balick, Michael J and Robert Mendelsohn, "Assessing the economic value of traditional medicines from tropical rain forests," *Conservation Biology*, 1992, pp. 128–130.

Barreca, Alan, Karen Clay, Olivier Deschenes, Michael Greenstone, and Joseph S Shapiro, "Adapting to climate change: The remarkable decline in the US temperature-mortality relationship over the twentieth century," *Journal of Political Economy*, 2016, *124* (1), 105–159.

Barrios, Salvador, Luisito Bertinelli, and Eric Strobl, "Trends in rainfall and economic growth in Africa: A neglected cause of the African growth tragedy," *The Review of Economics and Statistics*, 2010, *92* (2), 350–366.

Bernknopf, Richard L, David S Brookshire, and Mark A Thayer, "Earthquake and volcano hazard notices: An economic evaluation of changes in risk perceptions," *Journal of Environmental Economics and Management*, 1990, *18* (1), 35–49.

Beron, Kurt J, James C Murdoch, Mark A Thayer, and Wim PM Vijverberg, "An analysis of the housing market before and after the 1989 Loma Prieta earthquake," *Land Economics*, 1997, pp. 101–113.

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor, "Do individuals make sensible health insurance decisions? Evidence from a menu with dominated options," Technical Report, National Bureau of Economic Research 2015.

Bin, Okmyung and Craig E Landry, "Changes in implicit flood risk premiums: Empirical evidence from the housing market," *Journal of Environmental Economics and management*, 2013, *65* (3), 361–376.

_ **and** _ , "Changes in implicit flood risk premiums: Empirical evidence from the housing market," *Journal of Environmental Economics and management*, 2013, *65* (3), 361–376.

_ **and Stephen Polasky**, "Effects of flood hazards on property values: evidence before and after Hurricane Floyd," *Land Economics*, 2004, *80* (4), 490–500.

Blackmore, Nicole, "Flood insurance: 'How has my premium leapt 550pc when I'm nowhere near water?'," *The Telegraph*, 2015.

Boerner, Jan, Arisbe Mendoza, and Stephen A Vosti, "Ecosystem services, agriculture, and rural poverty in the Eastern Brazilian Amazon: Interrelationships and policy prescriptions," *Ecological Economics*, 2007, *64* (2), 356–373.

Bohlken, Anjali Thomas and Ernest John Sergenti, "Economic growth and ethnic violence: An empirical investigation of HinduMuslim riots in India," *Journal of Peace research*, 2010, *47* (5), 589–600.

Bucks, Brian and Karen Pence, "Do borrowers know their mortgage terms?," *Journal of urban Economics*, 2008, *64* (2), 218–233.

Burke, Marshall and Kyle Emerick, "Adaptation to climate change: Evidence from US agriculture," *American Economic Journal: Economic Policy*, 2016, *8* (3), 106–40.

Busse, Meghan R, Christopher R Knittel, and Florian Zettelmeyer, "Are consumers myopic? Evidence from new and used car purchases," *American Economic Review*, 2013, *103* (1), 220–56.

_ , Devin G Pope, Jaren C Pope, and Jorge Silva-Risso, "The psychological effect of weather on car purchases," *The Quarterly Journal of Economics*, 2015, *130* (1), 371–414.

Camerer, Colin F and Howard Kunreuther, "Decision processes for low probability events: Policy implications," *Journal of Policy Analysis and Management*, 1989, *8* (4), 565–592.

Chan, Sewin and Ann Huff Stevens, "What you don't know can't help you: Pension knowledge and retirement decision-making," *The Review of Economics and Statistics*, 2008, *90* (2), 253–266.

Chaney, Eric, "Revolt on the Nile: Economic shocks, religion, and political power," *Econometrica*, 2013, *81* (5), 2033–2053.

Chivers, James and Nicholas E Flores, "Market failure in information: the national flood insurance program," *Land Economics*, 2002, *78* (4), 515–521.

Choi, James J, David Laibson, Brigitte C Madrian, and Andrew Metrick, "Defined contribution pensions: Plan rules, participant choices, and the path of least resistance," *Tax policy and the economy*, 2002, *16*, 67–113.

da Silva, Renato Ramos, David Werth, and Roni Avissar, "Regional impacts of future land-cover changes on the Amazon basin wet-season climate," *Journal of climate*, 2008, *21* (6), 1153–1170.

Daniel, Vanessa E, Raymond JGM Florax, and Piet Rietveld, "Floods and residential property values: a hedonic price analysis for the Netherlands," *Built environment*, 2009, *35* (4), 563–576.

Davis, Lucas W and Lutz Kilian, "Estimating the effect of a gasoline tax on carbon emissions," *Journal of Applied Econometrics*, 2011, *26* (7), 1187–1214.

Department for Environment Food & Rural Affairs, "Securing the future availability and affordability of home insurance in areas of flood risk," Technical Report June 2013.

_ , "Impact Assessment: Managing the future financial risk of flooding," Technical Report Oct 2014.

Dettwiller, J. and S. A. Changnon, "Possible Urban Effects on Maximum Daily Rainfall at Paris, St. Louis and Chicago," *Journal of Applied Meteorology*, May 1976, *15* (5), 517–519.

Donovan, Geoffrey H, Patricia A Champ, and David T Butry, "Wildfire risk and housing prices: a case study from Colorado Springs," *Land Economics*, 2007, *83* (2), 217–233.

Edmonds, Timothy, "Household flood insurance," Technical Report, House of Commons Library 2017.

Eltahir, Elfatih AB and Rafael L Bras, "Precipitation recycling in the Amazon basin," *Quarterly Journal of the Royal Meteorological Society*, 1994, *120* (518), 861–880.

Environment Agency, "Floods cost UK economy nearly 600 million," 2013.

_ , "The costs and impacts of the winter 2013 to 2014 floods," Technical Report 2016.

Fankhauser, Samuel, Joel B Smith, and Richard SJ Tol, "Weathering climate change: some simple rules to guide adaptation decisions," *Ecological economics*, 1999, *30* (1), 67–78.

Fearnside, Philip M, "Biodiversity as an environmental service in Brazil's Amazonian forests: risks, value and conservation," *Environmental conservation*, 1999, *26* (04), 305–321.

Figueiredo, Miguel FP De, F Daniel Hidalgo, and Yuri Kasahara, "When do voters punish corrupt politicians? Experimental evidence from Brazil," *Yale University*, 2010.

Fjelde, Hanne and Nina von Uexkull, "Climate triggers: Rainfall anomalies, vulnerability and communal conflict in sub-Saharan Africa," *Political Geography*, 2012, *31* (7), 444–453.

Flood Re, "Transitioning to an affordable market for household flood insurance: The first Flood Re transition plan," Technical Report 2016.

Ford, James D, Lea Berrang-Ford, and Jaclyn Paterson, "A systematic review of observed climate change adaptation in developed nations," *Climatic change*, 2011, *106* (2), 327–336.

Fullerton, Don and Sarah E West, "Can taxes on cars and on gasoline mimic an unavailable tax on emissions?," *Journal of Environmental Economics and Management*, 2002, *43* (1), 135–157.

Gallagher, Justin, "Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States," *American Economic Journal: Applied Economics*, July 2014, *6* (3), 206–33.

Garcia-Carreras, L and DJ Parker, “How does local tropical deforestation affect rainfall?,” *Geophysical Research Letters*, 2011, *38* (19).

Gibson, Matthew, Jamie T Mullins, and Alison Hill, “Climate change, flood risk, and property values: Evidence from New York City,” Technical Report, Working Paper 2017.

Grigolon, Laura, Mathias Reynaert, and Frank Verboven, “Consumer valuation of fuel costs and tax policy: Evidence from the European car market,” *American Economic Journal: Economic Policy*, 2018, *10* (3), 193–225.

Grothmann, Torsten and Anthony Patt, “Adaptive capacity and human cognition: the process of individual adaptation to climate change,” *Global Environmental Change*, 2005, *15* (3), 199–213.

Gustman, Alan L and Thomas L Steinmeier, “Imperfect knowledge, retirement and saving,” Technical Report, National Bureau of Economic Research 2001.

Hansen, Winslow D and Helen T Naughton, “The effects of a spruce bark beetle outbreak and wildfires on property values in the wildland–urban interface of south-central Alaska, USA,” *Ecological Economics*, 2013, *96*, 141–154.

Henderson-Sellers, Ann and V Gornitz, “Possible climatic impacts of land cover transformations, with particular emphasis on tropical deforestation,” *Climatic Change*, 1984, *6* (3), 231–257.

HM Treasury, “The UK insurance growth action plan,” Technical Report 2013.

Ho, Kate, Joseph Hogan, and Fiona Scott Morton, “The impact of consumer inattention on insurer pricing in the Medicare Part D program,” *The RAND Journal of Economics*, 2017, *48* (4), 877–905.

Hsiang, Solomon M and Daiju Narita, “Adaptation to cyclone risk: Evidence from the global cross-section,” *Climate Change Economics*, 2012, *3* (02), 1250011.

Keller, Carmen, Michael Siegrist, and Heinz Gutscher, “The role of the affect and availability heuristics in risk communication,” *Risk analysis*, 2006, *26* (3), 631–639.

Keys, Benjamin J, Devin G Pope, and Jaren C Pope, “Failure to refinance,” *Journal of Financial Economics*, 2016, *122* (3), 482–499.

Kirkby, Christopher A, Renzo Giudice-Granados, Brett Day, Kerry Turner, Luz Marina Velarde-Andrade, Agusto Dueñas-Dueñas, Juan Carlos Lara-Rivas, and W Yu Douglas, “The market triumph of ecotourism: an economic investigation of the private and social benefits of competing land uses in the Peruvian Amazon,” *PloS one*, 2010, *5* (9), e13015.

Koster, Hans RA and Jos Van Ommeren, "A shaky business: Natural gas extraction, earthquakes and house prices," *European Economic Review*, 2015, *80*, 120–139.

Kousky, Carolyn, "Learning from extreme events: Risk perceptions after the flood," *Land Economics*, 2010, *86* (3), 395–422.

Kunreuther, Howard, Nathan Novemsky, and Daniel Kahneman, "Making low probabilities useful," *Journal of risk and uncertainty*, 2001, *23* (2), 103–120.

Lacetera, Nicola, Devin G Pope, and Justin R Sydnor, "Heuristic thinking and limited attention in the car market," *American Economic Review*, 2012, *102* (5), 2206–36.

Lamond, Jessica E, DG Proverbs, and FN Hammond, "Accessibility of flood risk insurance in the UK: confusion, competition and complacency," *Journal of Risk Research*, 2009, *12* (6), 825–841.

Lean, Judith and David A Warrilow, "Simulation of the regional climatic impact of Amazon deforestation," *Nature*, 1989, *342* (6248), 411.

Maccini, Sharon and Dean Yang, "Under the weather: Health, schooling, and economic consequences of early-life rainfall," *American Economic Review*, 2009, *99* (3), 1006–26.

Makarieva, A. M., V. G. Gorshkov, D. Sheil, A. D. Nobre, P. Bunyard, and B.-L. Li, "Why Does Air Passage over Forest Yield More Rain? Examining the Coupling between Rainfall, Pressure, and Atmospheric Moisture Content*," *Journal of Hydrometeorology*, February 2014, *15* (1), 411–426.

Makarieva, AnastassiaM., VictorG. Gorshkov, and Bai-Lian Li, "Conservation of water cycle on land via restoration of natural closed-canopy forests: implications for regional landscape planning," *Ecological Research*, 2006, *21* (6), 897–906.

Martin, Ralf, Laure B De Preux, and Ulrich J Wagner, "The impact of a carbon tax on manufacturing: Evidence from microdata," *Journal of Public Economics*, 2014, *117*, 1–14.

Maystadt, Jean-Francois, Olivier Ecker, and Athur Mabiso, *Extreme weather and civil war in Somalia: Does drought fuel conflict through livestock price shocks?*, Intl Food Policy Res Inst, 2013.

McCoy, Shawn J and Randall P Walsh, "Wildfire risk, salience & housing demand," *Journal of Environmental Economics and Management*, 2018, *91*, 203–228.

Miguel, Edward, Shanker Satyanath, and Ernest Sergenti, "Economic Shocks and Civil Conflict: An Instrumental Variables Approach," *Journal of Political Economy*, August 2004, *112* (4), 725–753.

Mueller, Julie, John Loomis, and Armando González-Cabán, "Do repeated wildfires change homebuyers demand for homes in high-risk areas? A hedonic analysis of the short and long-term effects of repeated wildfires on house prices in Southern California," *The Journal of Real Estate Finance and Economics*, 2009, *38* (2), 155–172.

Myers, Erica, "Are Home Buyers Inattentive? Evidence From Capitalization of Energy Costs," *American Economic Journal: Economic Policy*, 2018.

Naoi, Michio, Miki Seko, and Kazuto Sumita, "Earthquake risk and housing prices in Japan: Evidence before and after massive earthquakes," *Regional Science and Urban Economics*, 2009, *39* (6), 658–669.

Negri, Andrew J, Robert F Adler, Liming Xu, and Jason Surratt, "The impact of Amazonian deforestation on dry season rainfall," *Journal of Climate*, 2004, *17* (6), 1306–1319.

Nepstad, Daniel C, Claudia M Stickler, Britaldo Soares-Filho, and Frank Merry, "Interactions among Amazon land use, forests and climate: prospects for a near-term forest tipping point," *Philosophical Transactions of the Royal Society B: Biological Sciences*, 2008, *363* (1498), 1737–1746.

Oliveira, Leydimere JC, Marcos H Costa, Britaldo S Soares-Filho, and Michael T Coe, "Large-scale expansion of agriculture in Amazonia may be a no-win scenario," *Environmental Research Letters*, 2013, *8* (2), 024021.

Penning-Rowsell, Edmund C, Sally Priest, and Clare Johnson, "The evolution of UK flood insurance: incremental change over six decades," *International Journal of Water Resources Development*, 2014, *30* (4), 694–713.

Peters, Charles M, Alwyn H Gentry, and Robert O Mendelsohn, "Valuation of an Amazonian rainforest," *Nature*, 1989, (339), 655–656.

Pizer, William A, "Combining price and quantity controls to mitigate global climate change," *Journal of public economics*, 2002, *85* (3), 409–434.

Rambaldi, Alicia N, Cameron S Fletcher, Kerry Collins, and Ryan RJ McAllister, "Housing shadow prices in an inundation-prone suburb," *Urban Studies*, 2013, *50* (9), 1889–1905.

Rivers, Nicholas and Brandon Schaufele, "Salience of carbon taxes in the gasoline market," *Journal of Environmental Economics and Management*, 2015, *74*, 23–36.

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

Rosenfeld, Daniel, "Suppression of Rain and Snow by Urban and Industrial Air Pollution," *Science*, 2000, *287* (5459), 1793–1796.

Rudgard, Olivia, "Is Flood Re insurance scheme working? 'I was quoted £3,500 a year - now it's only £425'," *The Telegraph*, 2016.

Saatchi, Sassan, Salvi Asefi-Najafabady, Yadvinder Malhi, Luiz E. O. C. Arago, Liana O. Anderson, Ranga B. Myneni, and Ramakrishna Nemani, "Persistent effects of a severe drought on Amazonian forest canopy," *Proceedings of the National Academy of Sciences*, 2013, *110* (2), 565–570.

Sallee, James M, Sarah E West, and Wei Fan, "Do consumers recognize the value of fuel economy? Evidence from used car prices and gasoline price fluctuations," *Journal of Public Economics*, 2016, *135*, 61–73.

Samanta, Arindam, Sangram Ganguly, Eric Vermote, Ramakrishna R. Nemani, and Ranga B. Myneni, "Why Is Remote Sensing of Amazon Forest Greenness So Challenging?," *Earth Interactions*, June 2012, *16* (7), 1–14.

_ , Yuri Knyazikhin, Liang Xu, Robert E. Dickinson, Rong Fu, Marcos H. Costa, Sassan S. Saatchi, Ramakrishna R. Nemani, and Ranga B. Myneni, "Seasonal changes in leaf area of Amazon forests from leaf flushing and abscission," *Journal of Geophysical Research: Biogeosciences*, 2012, *117* (G1), n/a–n/a.

Schlenker, Wolfram, W Michael Hanemann, and Anthony C Fisher, "Will US agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach," *American Economic Review*, 2005, pp. 395–406.

Sheil, Douglas and Daniel Murdiyarso, "How Forests Attract Rain: An Examination of a New Hypothesis," *BioScience*, 2009, *59* (4), pp. 341–347.

Shepherd, J. Marshall, "A Review of Current Investigations of Urban-Induced Rainfall and Recommendations for the Future," *Earth Interactions*, July 2005, *9* (12), 1–27.

_ , Harold Pierce, and Andrew J. Negri, "Rainfall Modification by Major Urban Areas: Observations from Spaceborne Rain Radar on the TRMM Satellite," *Journal of Applied Meteorology*, July 2002, *41* (7), 689–701.

Shukla, J and Yale Mintz, "Influence of land-surface evapotranspiration on the earth's climate," *Science*, 1982, *215* (4539), 1498–1501.

Skantz, Terrance and Thomas Strickland, "House prices and a flood event: an empirical investigation of market efficiency," *Journal of Real Estate Research*, 1987, *2* (2), 75–83.

Smith, V Kerry, Jared C Carbone, Jaren C Pope, Daniel G Hallstrom, and Michael E Darden, "Adjusting to natural disasters," *Journal of Risk and Uncertainty*, 2006, *33* (1-2), 37–54.

Soares-Filho, Britaldo, Paulo Moutinho, Daniel Nepstad, Anthony Anderson, Hermann Rodrigues, Ricardo Garcia, Laura Dietzsch, Frank Merry, Maria Bowman, Letícia Hissa et al., "Role of Brazilian Amazon protected areas in climate change mitigation," *Proceedings of the National Academy of Sciences*, 2010, *107* (24), 10821–10826.

Spracklen, D. V., S. R. Arnold, and C. M. Taylor, "Observations of increased tropical rainfall preceded by air passage over forests," *Nature*, September 2012, *489* (7415), 282–285.

Stickler, Claudia M., Michael T. Coe, Marcos H. Costa, Daniel C. Nepstad, David G. McGrath, Livia C. P. Dias, Hermann O. Rodrigues, and Britaldo S. Soares-Filho, "Dependence of hydropower energy generation on forests in the Amazon Basin at local and regional scales," *Proceedings of the National Academy of Sciences*, 2013, *110* (23), 9601–9606.

Taylor, Christopher M, Eric F Lambin, Nathalie Stephenne, Richard J Harding, and Richard LH Essery, "The influence of land use change on climate in the Sahel," *Journal of Climate*, 2002, *15* (24), 3615–3629.

Thomalla, Frank, Tom Downing, Erika Spanger-Siegfried, Guoyi Han, and Johan Rockström, "Reducing hazard vulnerability: towards a common approach between disaster risk reduction and climate adaptation," *Disasters*, 2006, *30* (1), 39–48.

Tompkins, Emma L, W Neil Adger, Emily Boyd, Sophie Nicholson-Cole, Keith Weatherhead, and Nigel Arnell, "Observed adaptation to climate change: UK evidence of transition to a well-adapting society," *Global environmental change*, 2010, *20* (4), 627–635.

Wang, Jingfeng, Rafael L Bras, and Elfatih AB Eltahir, "The impact of observed deforestation on the mesoscale distribution of rainfall and clouds in Amazonia," *Journal of Hydrometeorology*, 2000, *1* (3), 267–286.

Xue, Yongkang and Jagadish Shukla, "The influence of land surface properties on Sahel climate. Part II. Afforestation," *Journal of Climate*, 1996, *9* (12), 3260–3275.

Yamane, Fumihiro, Hideaki Ohgaki, and Kota Asano, "The immediate impact of the Fukushima Daiichi accident on local property values," *Risk Analysis*, 2013, *33* (11), 2023–2040.

Yang, Dean and HwaJung Choi, "Are remittances insurance? Evidence from rainfall shocks in the Philippines," *The World Bank Economic Review*, 2007, *21* (2), 219–248.