American Economic Journal: Applied Economics 2014, 6(3): 206–233
http://dx.doi.org/10.1257/app.6.3.206

Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States[†]

By Justin Gallagher*

I examine the learning process that economic agents use to update their expectation of an uncertain and infrequently observed event. I use a new nation-wide panel dataset of large regional floods and flood insurance policies to show that insurance take-up spikes the year after a flood and then steadily declines to baseline. Residents in nonflooded communities in the same television media market increase take-up at one-third the rate of flooded communities. I find that insurance take-up is most consistent with a Bayesian learning model that allows for forgetting or incomplete information about past floods. (*JEL* D12, D83, D84, G22, Q54)

Economists have long been interested in understanding how individuals form beliefs over the likelihood of random events such as natural disasters. One reason why natural disasters have garnered attention is the finding that economic agents appear to overreact to the occurrence of a new disaster (e.g., Slovic, Kunreuther, and White 1974; Kunreuther 1976; Kunreuther et al. 1978).[1] Kahneman (2011) points to the research on natural disasters as among the earliest evidence of the judgment heuristic known as availability bias.[2] Nevertheless, a large and immediate change in beliefs after a disaster could be consistent with the common Bayesian learning model (DeGroot 1970; Viscusi 1991; Davis 2004).

Flooding is an example of a type of rare stochastic event where detailed information regarding the likelihood of the event is accessible, but personal experience is infrequent. In most communities in the United States, decades of historical flood

*Department of Economics, Weatherhead School of Management, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106-7235 (e-mail: justin.p.gallagher@case.edu). I thank David Card, Mariana Carrera, David Clingingsmith, Stefano DellaVigna, Michael Greenstone, Brad Howells, Patrick Kline, Vikram Maheshri, Enrico Moretti, Owen Ozier, Mark Votruba, and Philippe Wingender, four anonymous referees, and seminar participants at Case Western Reserve University, Kent State University, Loyola Marymount University, Massachusetts Institute of Technology, University of California-Berkeley, and Wharton for their many helpful comments on this project. Andrew Loucky, Di Tang, Yuhe Wang, and especially Trevor Allen and Anthony Gatti provided outstanding research assistance. I am grateful for funding provided by the EPA's Science to Achieve Results (STAR) fellowship, the Institute of Business and Economic Research, the NBER working group on Household Finance, and Weatherhead School of Management. All errors are my own.

[†] Go to http://dx.doi.org/10.1257/app.6.3.206 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] This finding is sometimes described as an underreaction in terms of preparedness and expectations before a disaster rather than an overreaction afterward.

[2] Availability bias is described as "situations in which people assess the frequency of a class or the probability of an event by the ease with which instances or occurrences can be brought to mind" (Tversky and Kahneman 1982, 11).

records exist. Detailed parcel-level flood maps indicating the precise location of each property vis-a-vis the flood plain are also available to residents.[3] Other settings that share similar characteristics to flooding include certain types of crime (e.g., home robberies) and health risks (e.g., work-place injuries).

This paper examines how flood risk beliefs change after floods using a new panel dataset on flooding and the purchase of flood insurance.[4] The dataset includes information on all flood insurance policies in the United States for each calendar year and whether a community is hit by a Presidential Disaster Declaration (PDD) flood that year. The 18-year community-level flood panel includes data on approximately 27 million annual flood insurance policies, 11,025 county floods, and 643 distinct PDD floods. Virtually the entire country (92 percent of the counties in the sample) was hit by at least 1 of these floods.

I use the change in the number of insurance policies per capita as a measure of changing homeowner beliefs over the expectation of a future flood. A simple homeowner flood insurance model implies that the demand for flood insurance increases as the expected probability of a future flood increases. Homeowner insurance policies explicitly exempt coverage for damage due to flooding and homeowners must decide each year whether to purchase a separate flood insurance policy. Importantly, the price of flood insurance is not experience rated. The federal government sets the rates for flood insurance and insurance is available to homeowners before and after each flood at nearly identical rates.

An assumption of this paper is that community-level flood probabilities are constant from 1958–2007. Overall this is consistent with the view of the National Flood Insurance Program, which sets the insurance rates, and the Army Corps of Engineers, which creates the flood maps. Further, there is no evidence of annual serial correlation in PDD floods.[5]

I use a flexible event study framework to nonparametrically estimate the causal effect of large regional floods on insurance take-up for hit and neighboring homeowners. The identifying assumption is that, conditional on a community's geography and calendar time trends, whether or not a community is flooded in a particular year is random. I find strong evidence of an immediate rise in the fraction of homeowners covered by flood insurance in flooded communities. The effect peaks at 9 percent and then begins to steadily decline. After nine years, the effect of a flood is no longer statistically distinguishable from zero. The same spike and decay pattern in insurance take-up repeats if a community is hit by multiple floods during the panel. Take-up is the same after high and low per capita cost floods, suggesting that homeowners do not use the new floods to learn about flood costs.

[3] New homeowners are required by law to receive a copy of the flood map at the time the property is purchased. Also, community flood maps are required to be displayed publicly (e.g., at the Town Hall), and, more recently, are available online. Bin, Kruse, and Landry (2008) show that there is a price differential between similar homes inside and outside the 100-year floodplain. Since the housing market reflects market-level knowledge of the flood map boundaries, it is likely that most potential home buyers receive this information.

[4] All property owners can purchase insurance, but for the ease of exposition, I refer to flood insurance policy holders as homeowners. A community is defined by the National Flood Insurance Program as a local political entity (e.g., village, town, city). This definition is similar to a US Census place.

[5] I test the assumption of independence in PDD floods using a Wald-Wolfowitz Runs Test (Swed and Eisenhart 1943). Section I and online Appendix Sections B.6 and C provide more details on this fixed probability assumption.

The large jump in insurance take-up implies that homeowners do not make a one time decision on whether to purchase flood insurance based, for example, on the risk-based flood maps. The size of the jump is also striking given the long history of past floods in most communities. A new flood provides very little new statistical information given the history of past floods. The jump combined with the quick decline to baseline levels suggests that homeowners are not incorporating all available information. This could occur if current homeowners forget about past floods, or by migration if homeowners only use flood information from the years spent living in the community. In both cases, the amount of flood information is limited and the relative importance of a new flood in forming flood beliefs is large. In the years after a new flood, the effect of the recent flood on the expectations of a future flood will quickly lessen (implying a quick return to baseline) as residents begin to forget or because of the entry of new residents.

The event study framework is also used to examine whether homeowners in communities close to a flood learn about flood risks from the experience of their neighbors. The goal is to provide evidence on whether homeowners incorporate the experience of others when updating beliefs over the risk of a flood (Camerer and Ho 1999 and Ho and Chong 2003). I am able to separately measure how direct and indirect experience affect perceptions of a future flood and to compare the relative importance of each.

I consider two different measures for proximity to a flood: geographic distance and the sharing of TV media exposure (Snyder and Strömberg 2010). We are able to separately identify TV media market and geographic neighbor effects by taking advantage of the exogenously determined media markets and the random timing and location of the floods. We might expect homeowners in geographically neighboring communities to increase insurance if there is minor flooding outside the highly impacted areas. Also, if geographic areas share similar flood risks, then homeowners could use nearby flooding to learn about their own flood risk. I find that insurance take-up in communities not hit by a flood, but located either within or just outside a flooded county, increases by about 3 percent in the years after a nearby flood.

Local TV news is a potential source of general flood risk information and a means to learn about nearby floods. The content of TV news broadcasts vary by media market. I use closed captioning information on local TV news broadcasts to show that there are three times as many flood news stories in media markets when there is a PDD flood. The number of news stories increases with the proportion of the media market that is flooded.

Insurance take-up after a flood for nonflooded communities that share a TV media market is one-third as large as in flooded communities and persists for six years. Take-up for nonflooded media neighbors increases with the proportion of the media market that is flooded. The geographic neighbor take-up effect mostly disappears after accounting for whether nonflooded homeowners are in the same media market as a flood. Take-up within a media market does not vary by distance from the flood.

There is no evidence that nonflooded homeowners distinguish between the relevancy of the new flood information from media market floods. Homeowners respond to media market floods the same regardless of whether the flooded community shares a very similar flood history. This is surprising if we believe that the difference between flooding in two communities with very similar flood histories is due to randomness and not differences in community flood risk characteristics.

In Section IV, I test how well a full information Bayesian learning model fits the observed changes in insurance take-up. I simulate changes in conditional flood probabilities under the assumption that homeowners update their beliefs using the 50-year history of PDD floods (1958–2007). Changes in conditional flood probabilities cannot match the pattern of insurance take-up. The event study and simulation evidence points toward a learning model that allows homeowners to weigh recent floods more heavily than earlier floods (Camerer and Ho 1999; Malmendier and Nagel 2011). The data are also consistent with Availablity Bias, a nonlearning model interpretation (Tversky and Kahneman 1982).

There are several possible underlying learning model explanations including a mistaken understanding of the flooding process, forgetting by current residents, and migration. One challenge in distinguishing between the forgetting and migration explanations is that flood insurance data are aggregated at the community-level and I am unable to observe which policies are dropped because a homeowner moved and sold the property. Nevertheless, there is suggestive evidence for the role of migration. Insurance take-up returns to baseline levels after a flood faster in population-increasing communities than in population-decreasing communities. I also show that a learning model calibrated using county migration rates could match observed insurance take-up.

A number of previous studies examine the immediate change in flood expectations after a flood using stated preferences (e.g., Kunreuther 1976; 1995) and land prices (e.g., Bin, Kruse and Landry 2008; Kousky 2010; Bin and Landry 2013). A more recent literature uses panel datasets on flood insurance policies to evaluate factors that affect demand for insurance (Browne and Hoyt 2000; Kriesel and Landry 2004), characteristics of policyholders (Michel-Kerjan and Kousky 2008), and policy tenure (Michel-Kerjan, de Forges, and Kunreuther 2012). This paper differs from the previous literature in that it is the first (to my knowledge) to document the dynamic multi-year effect of new floods on insurance take-up, and to use this pattern to evaluate possible risk learning models. This paper is also the first to show how neighboring floods, including floods in the same TV media market, affect take-up.[6]

Studies that document spikes in revised beliefs after nonflooding environmental events include Palm (1995) (earthquakes), Davis (2004) (cancer clusters), and Deryugina (2013) (weather). Malmendier and Nagel (2011) and Davis (2004) both study learning environments similar to flooding and document the persistence of beliefs over time. Malmendier and Nagel (2011) examine how past stock market returns affect investment portfolio purchasing decisions. Davis (2004) studies how the public disclosure of new cancer cases affects beliefs over environmental cancer risk. Davis (2004) finds that the standard (full information) Bayesian model can fit the data, while Malmendier and Nagel (2011) find support for a discounting model.

[6] I am not aware of another paper that studies how TV media affect beliefs about the environment. A related economic literature examines the effect of media coverage on voting behavior and political outcomes: e.g., Ansolabehere, Snowberg, and Snyder (2006) and DellaVigna and Kaplan (2007) (television), Ferraz and Finan (2008) (radio), Snyder and Strömberg (2010), Gentzkow, Shapiro, and Sinkinson (2011), and Gerber, Karlan, and Bergan (2009) (newspaper).

This paper differs from Davis (2004) in that there are low-frequency signals over a relatively long time horizon.

The prevailing view is that the overall level of flood insurance take-up is too low relative to the social optimum (e.g., Kunreuther 1996; Kriesel and Landry 2004; Kunreuther, Meyer, and Michel-Kerjan 2009). The learning model interpretation– that homeowners discount past floods—underscores this conclusion. Discounting past floods (for whatever reason) is likely to lead homeowners to underestimate their true risk—and thus underinsure.[7] If homeowners are underinsured, then a temporary increase in flood insurance could be welfare improving from the perspective of the homeowner. A policy that seeks to lock in insurance purchase at the higher level immediately after a flood, for example, through either multiyear or automatic renewal insurance contracts, would likely improve homeowner welfare (Jaffee, Kunreuther, and Michel-Kerjan 2008). However, this conclusion must be tempered by the fact that most homeowners are charged a price for flood insurance that is 30–40 percent above NFIP determined actuarial rates. Also, the finding that nonflooded homeowners increase insurance purchase by the same amount regardless of the underlying information content of the flood increases the likelihood that some homeowners may overreact and initially overinsure.[8] A complete welfare calculation would take into account government expenditures and how flood damage impacts banks and financial companies (Michel-Kerjan, de Forges, and Kunreuther 2012).

I. Flooding and Flood Insurance in the United States

A. *The National Flood Insurance Program*

Flood insurance was not available to home or business owners in the United States for most of the twentieth century.[9] The federal government created the National Flood Insurance Program (NFIP) in 1968.[10] The NFIP sets flood insurance premiums at "actuarial" rates based on historical flood data, hydrological modeling, and detailed community flood maps created by the Army Corps of Engineers. Engineering data and historical observations are used to determine expected damage. The expected damage-based rates are then increased by 30–40 percent to cover the expenses of running the program.[11]

[7] Mechanically this can be seen by comparing equations (3) and (4) in Section IV.

[8] Data restrictions prevent a more precise homeowner welfare calculation. A rigorous (homeowner) welfare calculation would, at a minimum, require knowledge of the precise geographic location of each homeowner policy, the level of flood insurance purchased by each homeowner, flood insurance policy premium rates at each location, and NFIP expected damages at each location.

[9] The reasons stated for why no private flood insurance market existed include the lack of accurate flood risk information that could prevent adverse selection, and the view that many homeowners are unwilling to pay actuarially fair prices (American Insurance Association 1956; Anderson 1974).

[10] This section provides a short overview of the NFIP. Online Appendix Section B has a more detailed discussion of several important aspects of the NFIP. FEMA(2002a, b), and Michel-Kerjan (2010) provide good descriptions of the NFIP and its history.

[11] The exception to this rate setting process are grandfathered structures built before 1975 (or the introduction of NFIP in each community). The rates for these structures are lower and approximately equal to expected flood damage (US Government Accountability Office 2008).

To simplify the rate-setting process, the NFIP specifies a limited number of nationally designated flood zones. The Army Corps of Engineers flood maps divide each part of each community as falling into one of approximately ten flood zones. The zones with the highest flood risk correspond to the 100-year flood plain. Different premium base rates are offered for each zone and adjusted within each zone according to a number of factors.[12]

Homeowners decide whether to purchase flood insurance each calendar year. Flood insurance polices are sold by private insurance companies at the rates specified by the NFIP. A homeowner's policy will be dropped if the homeowner doesn't pay the premium for the subsequent year.[13] Flood insurance and risk information is transmitted to home and business owners in a number of ways. First, each community offering NFIP insurance posts detailed publicly accessible copies of the Army Corps of Engineers flood maps. These maps allow each homeowner to precisely identify the location of his home and its corresponding flood zone. Second, flood zone documents are required at the time of purchase of a new home if the home is within the 100-year flood plain.[14] Finally, private insurance companies are compensated by the NFIP for each flood insurance policy transaction. Thus, insurance companies have an incentive to directly market flood insurance to homeowners.

One important implication of the NFIP rate setting process is that premium rates are unaffected by whether your home is flooded. The base premium rates (and adjustments) for the ten nationally designated flood zones are set for the entire country. A second implication of the rate setting process is that the base flood rates for the various zones remain virtually unchanged in real dollars for the years included in the panel analysis. For example, during the 10 years from 1996–2005, the average annual real rate increase was 0.61 percent for those properties built after 1975 and 1.49 percent for those properties grandfathered into the program (see Appendix Table 2).[15] Nevertheless, all econometric models in this paper will include flexible nonparametric controls for calender time.

All flood insurance policies in the United States are sold through the NFIP. Through a Freedom of Information Act request, I received NFIP data on all flood insurance policies from 1980–2007. The number of annual flood insurance policies has increased steadily from about 2 million in 1980 to 5.5 million in 2007. This paper focuses on the decision to purchase flood insurance after a large, regional flood and does not attempt to explain the overall trend in flood insurance take-up.

[12] The 100-year flood plain is defined by FEMA as the area of land that will be "inundated by the flood event having a 1 percent chance of being equaled or exceeded in any given year." See FEMA (2008) for more details regarding the rate setting process.

[13] Homeowners receive renewal notices from the insurance company handling the policy. Flood insurance can only be purchased in communities that officially participate in the NFIP. Approximately 90 percent of communities participate. Homeowners living in these communities can also purchase insurance at the same rates directly from the NFIP. Online Appendix Section B.1 provides more details.

[14] There are often building restrictions on new structures within the 100-year flood plain. In addition, all new structures that have a bank loan underwritten by the federal government are ostensibly required to have flood insurance. However, this law is not widely enforced (Dixon et al. 2006; FEMA 2007). Online Appendix Table 1 calculates, using GAO data, that 97 percent of homeowners purchase flood insurance by choice and not due to existing Mandatory Purchase Laws.

[15] These ten years are the only years for which I was able to receive a breakdown for annual premium price changes. NFIP personnel have assured me that this period is representative of the program's history.

The insurance data are aggregated at the community level for each calendar year by the NFIP. There are several limitations of using the aggregated flood insurance policy count data. For example, I am not able to distinguish between new and continuing flood policies. A second limitation is that the NFIP does not currently track which policies are for properties located in the 100-year flood plain.

B. *Presidential Disaster Declaration Floods*

The Disaster Relief Act of 1950 established the Presidential Disaster Declaration (PDD) system. The PDD system is a formalized process to request and receive federal assistance following large natural disasters. The declaration process has several steps. The governor of a state must write an official letter to the President requesting that a PDD be declared for specific counties in the state. In the letter, the governor outlines the scope of the disaster, including weather and damage information collected by local agencies. The letter must specify the list of counties in the state that would be part of a PDD. Historically, three-quarters of flooding PDD requests have been granted.[16]

A Presidential Disaster Declaration opens the door to two major types of disaster assistance. The largest component of disaster assistance is Public Assistance. Public Assistance is available to local and state governments, as well as nonprofit organizations located in a PDD county. These groups can access grant money to remove debris, repair infrastructure, and to aid in reconstruction of public buildings. The damage must have been caused by the natural disaster. The second type of disaster assistance is Individual Assistance. Individual Assistance is available to residents in PDD counties. Home and business owners can access low interest disaster loans to rebuild. Direct cash assistance is also available for temporary and emergency expenses, such as interim housing.

This paper uses PDD events as a data source of large regional floods. The data collected include the date of the PDD, the type of disaster, location information (county), and an estimate of disaster cost.[17] All communities participating in the NFIP that have nonmissing population data for the 1990–2007 panel are included in the event study analysis. There are 2,704 such counties (or county equivalents). This includes approximately 86 percent of all US counties and covers 93 percent of the US population.[18] Nearly every county in the sample (92 percent) is hit by at least 1 PDD flood during the 18 years from 1990–2007. The median number of PDD floods for a county is three.

Figure 1 shows a county delineated map of the continental United States. The map is color coded based on the number of Presidential Disaster Declarations from 1990–2007. The darker the shade of grey, the greater the number of floods. Black

[16] In 1986, FEMA established criteria to use when evaluating whether to grant a request. These criteria include estimated damage costs (Downton and Pielke 2001; Sylves and Búzás 2007).

[17] The paper uses all flooding-related Disaster Declarations. PDD data were downloaded from the Public Risk Institute Website. I also downloaded county flood cost data from SHELDUS, but opted not to use these data. Please see online Appendix Section D for details on the PDD data, and a cautionary note regarding the use of the SHELDUS data.

[18] The population data are from the US Census. This population calculation uses US Census 2000 data. Please refer to online Appendix Section D for details on the census data.

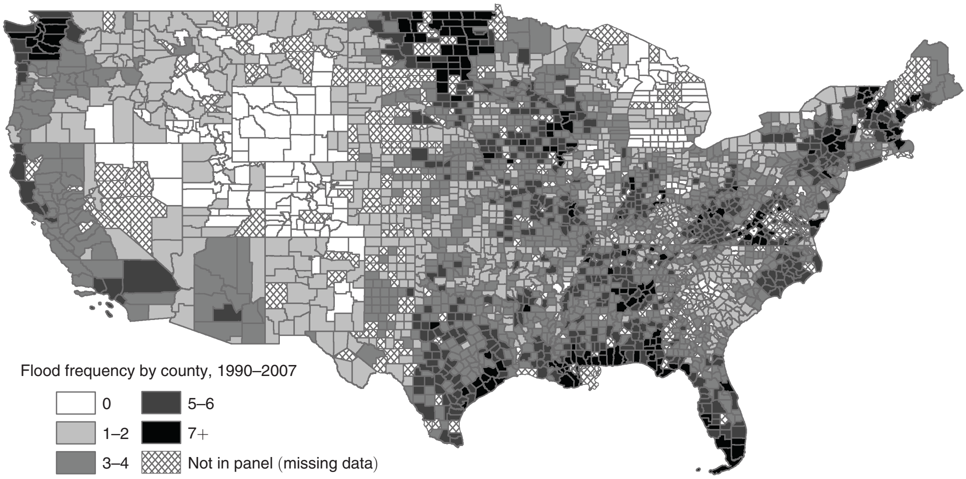

FIGURE 1. PRESIDENTIAL DISASTER DECLARATION FLOOD INTENSITY BY COUNTY 1990–2007

Notes: Map of the continental United States delineated by county. The map shows the Presidential Disaster Declaration (PDD) flood intensity by county from 1990–2007 for the counties included in the 1990–2007 community panel. Counties with no coloring have zero PDD floods. The darker the shade of grey, the greater the number of PDD floods. Counties with hash marks are not included in the panel. A county is dropped from the panel if no community in the county is included in the 1990–2007 community panel. The online text and data Appendix provide more details on the Presidential Disaster Declaration flood data and the 1990–2007 community panel.

corresponds to counties with seven or more PDD floods, while counties with zero floods are colored white. The counties with diagonal black lines are those excluded from the analysis.

PDD floods are determined at the county level. However, not all communities within a county may be affected by the flood. I construct a variable to identify which communities in PDD counties are "hit" by each flood. As described above, state and local governments, as well as nonprofits, are entitled to grant money to repair infrastructure and rebuild damaged structures.

Through a Freedom of Information Act request, I received a data file that lists the location of every Public Assistance damage claim paid out from 1990–2007. There are more than 800,000 unique observations. Using these data, I create an indicator variable for whether a community within a PDD county is hit by a particular flood. I consider a community to be hit if there is at least one Public Assistance claim with a damage location within the community.[19] Thirty-two percent of communities in PDD counties are hit by a PDD county-level flood in the year of a flood.

An assumption of this paper is that community-level flood probabilities are constant from 1958–2007. Overall, this is consistent with the view of the NFIP and the Army Corps of Engineers. Very few of the community flood maps have been modified since the maps were first created in the 1970's and early 1980's.[20] A second assumption is that there is no annual serial correlation in PDD floods. I test

[19] We are able to match 98.6 percent of the Public Assistance claims to a NFIP community. Please refer to online Appendix Section E for matching details.

[20] Online Appendix Section B.6 provides a more detailed discussion of this assumption.

this assumption of independence using a Wald-Wolfowitz Runs Test (Swed and Eisenhart 1943). Importantly, this test does not assume that the probability of a flood in each county is the same. I fail to reject the null hypothesis of the independence of annual floods at all conventional significance levels.[21]

II. Econometric Model

We use a flexible event study framework that nonparametrically estimates the causal effect that large regional floods have on the take-up of flood insurance. Equation (1) is the main estimating equation.

$$\ln(takeup_{ct}) = \sum_{\tau=-T}^{T} \beta_\tau W_{c\tau} + \alpha_c + \gamma_{st} + \epsilon_{ct}. \tag{1}$$

The unit of observation is a community calendar year. A community is defined by FEMA and roughly equal to the US Census Place definition (i.e., village, town, city, etc.). The dependent variable in equation (1), $\ln(takeup_{ct})$, is log Flood Policies Per Person for community c in year t.[22] The independent variables of interest are the event time indicator variables, $W_{c\tau}$. These variables track the year of a PDD flood and the years immediately preceding and following a flood. The indicator variable W_{c0} equals 1 if community c is hit by a flood in that calendar year.[23] The indicator variable $W_{c\tau}$ equals 1 if a community is hit by a flood in $-\tau$ years. Many communities are hit by more than one PDD flood during the event study. For these communities, each flood is coded with its own set of indicator variables.[24] The event time indicator variable W_{c-1} is normalized to zero when I estimate equation (1). In practice, this is done by excluding W_{c-1} from the regression. The estimated coefficients are interpreted as the percent change in the take-up of flood insurance in community c relative to the year before a flood.

In most of the specifications of equation (1), I bin the $W_{c\tau}$ by creating a single indicator variable for the end periods. The bin indicator variables serve a practical purpose. I am most interested in the years shortly before and after a flood. The event time indicator variables, $W_{c\tau}$, near the tails of the event study, are identified off of many fewer observations, and therefore have large standard errors. Binned indicator variables pool the effect on take-up over multiple event years to increase statistical power.[25]

[21] A Runs Test on the sample of 1990–2007 panel counties with at least two PDDs results in a p-value of 0.30. Online Appendix Section C provides a more detailed discussion.

[22] The number of policies in force is an extensive margin measure of insurance demand. An alternative is to use the quantity of insurance purchased (intensive measure). Using the number of policies in force avoids several theoretical and empirical challenges that are involved with using the quantity of insurance purchased. Please see online Appendix Section B.5 for a discussion.

[23] Occasionally, a community is hit by more than one PDD flood in the same calendar year. I don't distinguish between communities hit by one or more than one PDD flood in a particular year when estimating equation (1). The reason for this is that the flood insurance policy count data are aggregated by year.

[24] For example, Hazlehurst, Georgia is hit by a PDD flood in 1991 and 2004. Thus, in year 2000, $W_{c9} = 1$, since it has been 9 years since the 1991 PDD and $W_{c-4} = 1$, since it is 4 years before the 2004 PDD.

[25] For example, in the 1990–2007 panel event study $W_{c,17} = 1$ only if there is a flood in 1990. I create $W_{c,early} = 1$ if $\tau \in [-17, -11]$, and $W_{c,late} = 1$ if $\tau \in [11, 17]$. Equation (1) is then estimated with these two bin indicator

Equation (1) also includes community fixed effects (α_c), state by year fixed effects (γ_{st}), and a stochastic error term (ϵ_{ct}). The fixed effects nonparametrically control for unobserved (and unchanging) community characteristics and state-specific yearly factors. Community geography is important in predicting the likelihood of a flood. The underlying community geography includes surface characteristics, such as the percent of a community located in the flood plain, as well as location specific factors, such as average rainfall. State-by-year fixed effects account for state-specific yearly trends that may affect take-up, such as state-level responses to flooding, state economic conditions, and changes in NFIP institutional factors. Standard errors from the estimation of equation (1) are clustered at the state level.[26] Finally, the causal interpretation of equation (1) comes from the assumption that whether a community is hit by a flood in a particular year is random conditional on community and state-by-year fixed effects.

We are also interested in estimating the take-up of flood insurance for communities not directly hit by a flood.

$$\ln(takeup_{ct}) = \sum_{\tau=-T}^{T} \beta_\tau W_{c\tau} + \sum_{\tau=-T}^{T} \lambda_\tau N_{c\tau} + \alpha_c + \gamma_{st} + \epsilon_{ct}. \tag{2}$$

We estimate equation (2) when we consider "neighboring" communities that were not directly hit by a flood. Equation (2) is identical to equation (1), except that it also includes event time indicator variables for neighboring communities, $N_{c\tau}$.

We estimate equation (1) and equation (2) on a panel of communities over two different time periods: (i) 1980–2007 and (ii) 1990–2007. A community is included in each of these panels only if there is nonmissing data for each year.[27] These time periods are selected based on data availability. Community-level flood insurance policy data are available beginning in 1978, but the community-level population data are not widely available until 1980. Thus, the 28-year period from 1980–2007 is the longest panel for which we can estimate flood insurance take-up for a large sample of communities. In all of these regressions, the definition of a flood is whether a homeowner resides in a community that is in a Presidential Disaster Declaration county. For the period 1990–2007, we can use a more detailed definition of a flood hit. Beginning in 1990, we confirm whether a PDD flood declared at the county level damaged infrastructure or public buildings in each community in the county.

The approach of this paper is to use the more geographically precise 1990–2007 flood panel to establish the basic empirical result. Next, we confirm that the 1980–2007 panel reproduces the same pattern of flood insurance take-up. We then switch the remainder of the analysis to the 1980–2007 panel.

In addition to having a longer estimating panel, the 1980–2007 panel has at least four important advantages. First, I am able to exactly control for the lagged effect

variables rather than including the individual variables $W_{c,-11}, \ldots, W_{c,-17}$ and $W_{c,11}, \ldots, W_{c,17}$.

[26] Online Appendix Table 5, column 5 considers how the standard errors change when we account for a general form of spatial correlation as proposed by Driscoll and Kraay (1998).

[27] These two panels are balanced in calender time. Please see online Appendix Table 6 for estimates from models that are balanced in event time. The point estimates from a panel balanced in event time are remarkably similar. I focus on the balanced calendar time panel because the sample is much larger.

of PDD floods that occur before the start of the panel. The identification concern is that the 1990–2007 panel may incorrectly attribute the lagged take-up effect of a flood that occurs before the beginning of the panel to flood events during the panel.[28] Second, the county-level flood definition is consistent with the geographic precision of the data used to define a community flood "neighbor."[29] Third, national panel data on migration and income are available at the county-level. These data are used to test how sensitive the estimated insurance take-up results are to differing levels of household migration and income. Fourth, Section IV considers whether a learning model that incorporates past flood information when forming expectations about a future flood can explain the observed pattern of flood insurance take-up. The county-level definition of a flood allows for consistency between the historical flood data and the flood data in the 1980–2007 panel.

III. Estimation Results

A. *Communities Hit by a Flood*

Figure 2 plots the event time indicator coefficients, β_τ, from the estimation of equation (1) on the 1990–2007 panel. Event time is plotted on the x-axis. Year 0 corresponds to a year a community is hit by a PDD flood, while years $-1, \ldots, -10$ and $1, \ldots, 10$ are the years before and after a flood, respectively. The leftmost (rightmost) point on the graph is a pooled coefficient for the years -11 to -17 (11 to 17). The results are normalized to the year before a flood hit. The plotted event time coefficients can be interpreted as the percent change in the take-up of per capita flood insurance policies in the community relative to the year before a flood. The bands represent the 95 percent confidence interval and show whether each point estimate is statistically different from 0.

There is no discernable trend in take-up in the years before a flood. The effect of a future flood is economically small and not statistically different from zero for all time periods before the flood. In the year of a flood, there is an 8 percent increase in the take-up of flood insurance relative to the year before a flood. Take-up peaks at 9 percent the year after a flood. Take-up after the flood remains positive and statistically significant for nine years. After nine years, take-up is not statistically different relative to the year before a flood.[30]

Figure 3 plots the point estimates of hit and nonhit communities within a flooded county. The point estimates are from the estimation of equation (2) that specifically controls for the impulse response function of nonhit communities in PDD counties.

[28] County-level PDD flood data are available beginning in 1958. I use these earlier floods to precisely control for the lagged effect of floods that occur before 1980. The 1990–2007 panel only considers leads and lags for a flood if the PDD occurred within the time frame of the event study. There is no way to determine whether a community within a PDD county was "hit" by the county-level flood before 1990. The $W_{c\tau}$ indicator variables all equal zero for a community for any PDD flood outside the event study window.

[29] An important definition of a flood neighbor will be whether a community is in the same television media market. Media markets are defined at the county-level by Nielson Media Research.

[30] The point estimates and standard errors for specifications of equation (1) with year fixed effects are larger than those with state-specific time trends. Take-up in the year of a flood is about 2 percentage points larger, while the effect of a flood persists for 1 fewer year.

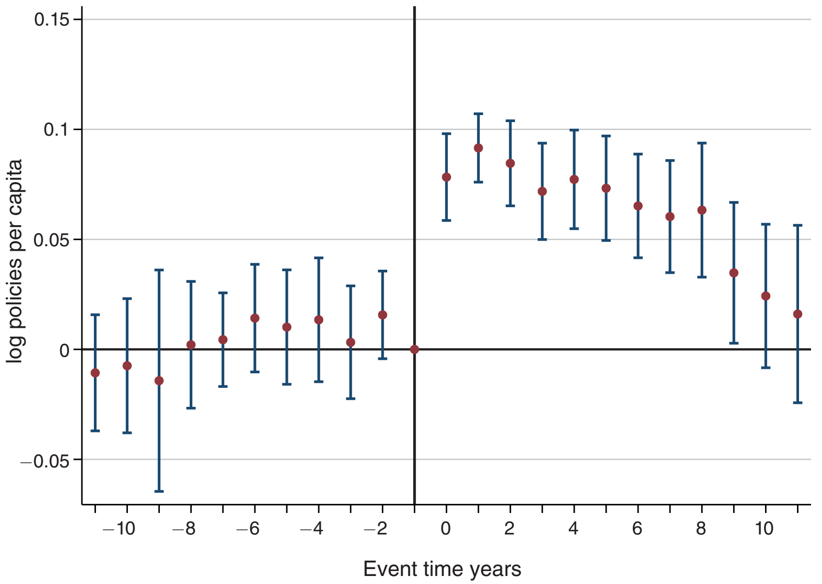

Figure 2. Flood Insurance Take-Up for Communities Hit by a Presidential Disaster Declaration Flood 1990–2007

Notes: The figure plots event time insurance take-up coefficients from estimation of equation (1) on the 1990–2007 panel. All estimated coefficients can be interpreted as the percent increase in flood insurance policies per capita for a hit community relative to the year before a flood (−1 on the *x*-axis). The end points on the graph are binned so that −11 (+11) is a bin for years −11 to −17 (+11 to +17). The vertical axis measures log per capita flood insurance take-up. The coefficient for the year before a flood is normalized to zero. The bars show the 95 percent confidence interval. Standard errors are clustered by state. There are 10,841 communities in the event study. A community is defined by the National Flood Insurance Program (NFIP) and corresponds to political jurisdictions: city, town, village, etc. A community is defined as hit if there is a Public Assistance damage claim submitted to the Federal Emergency Management Agency (FEMA) for damage from a Presidential Disaster Declaration flood.

There is a 2–3 percent increase in insurance take-up in nonhit communities within flooded counties. This effect persists for five years after a flood. The magnitude of the increase in take-up for nonhit communities is about one-third as large as that of hit communities.[31]

Figure 4 plots estimates of insurance take-up using equation (1) and the 1980–2007 panel. Recall that the definition of a flood for the 1980–2007 panel is whether the community is located in a PDD county. One advantage of this panel is the ability to precisely control for the lag effect of PDD floods that occur before the beginning of the panel. Flood insurance take-up peaks the year after a flood at a 9 percent increase relative to the year before a flood. Take-up in the years before a flood is economically small and statistically not different from 0 for all years except for 15 years before a flood.[32]

[31] Refer to online Appendix Section E and Tables 5 and 6 for further details regarding the 1990–2007 estimating panel. These tables include specifications that are balanced in event time (Table 6), exclude Louisiana communities (Table 5, column 6), and model the dependent variable in levels (Table 5, column 4).

[32] The point estimates are about 1–2 percentage points smaller and the duration of statistical significance is shorter when the specification does not control for the lagged effect of floods before 1980. This suggests that there is likely a downward bias in relying only on the estimation results from the 1990–2007 panel.

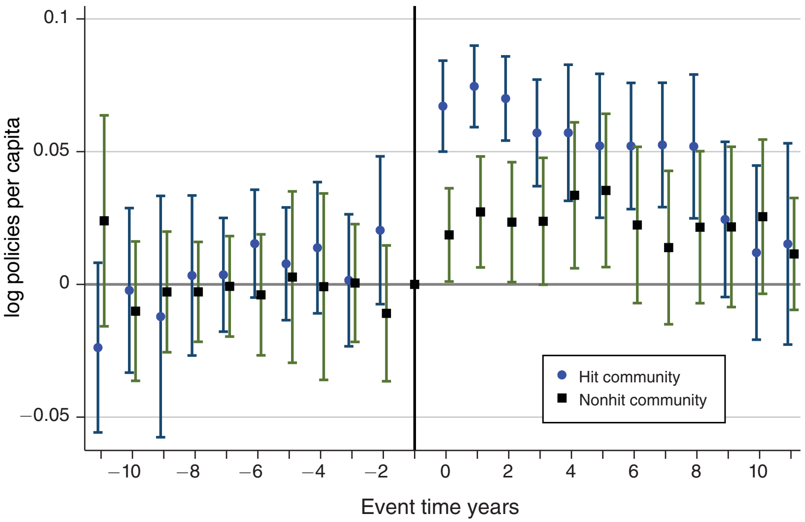

Figure 3. Flood Insurance Take-Up for Hit and Nonhit Communities within Presidential Disaster Declaration Flooded Counties 1990–2007

Notes: The figure plots event time insurance take-up coefficients from estimation of equation (2) on the 1990–2007 panel. The event study specification includes a set of indicators for nonhit communities within PDD counties. Please refer to the notes to Figure 2 and to Section III for more details on the event study specification.

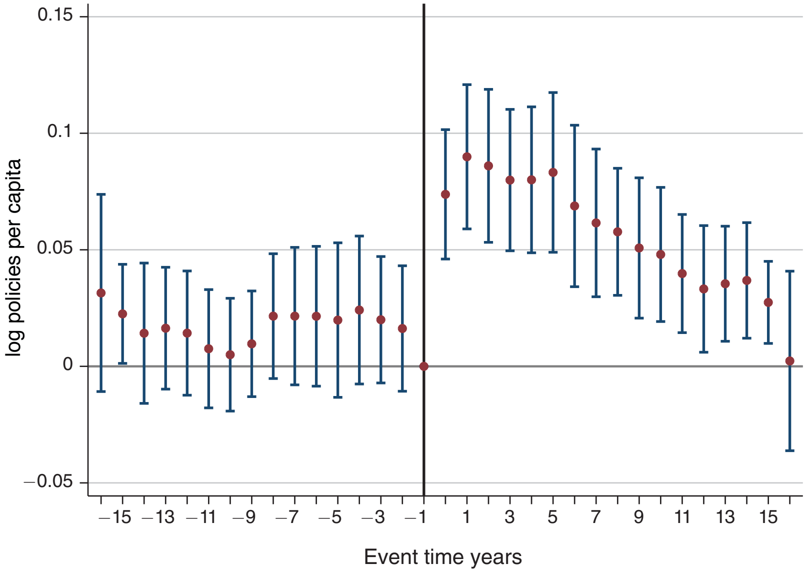

Figure 4. Flood Insurance Take-Up for Communities in Presidential Disaster Declaration Counties 1980–2007

Notes: The figure plots event time insurance take-up coefficients from estimation of equation (1) on the 1980–2007 panel. All estimated coefficients can be interpreted as the percent increase in flood insurance policies per capita for a hit community relative to the year before a flood (−1 on the x-axis). The end points on the graph are binned so that −16 (+16) is a bin for years −16 to −27 (+16 to +27). The vertical axis measures log per capita flood insurance take-up. The coefficient for the year before a flood is normalized to zero. The bars show the 95 percent confidence interval. Standard errors are clustered by state. There are 9,607 communities in the event study. A community is defined by the National Flood Insurance Program (NFIP) and corresponds to political jurisdictions: city, town, village, etc. A community is defined as hit if it is in a PDD flooded county.

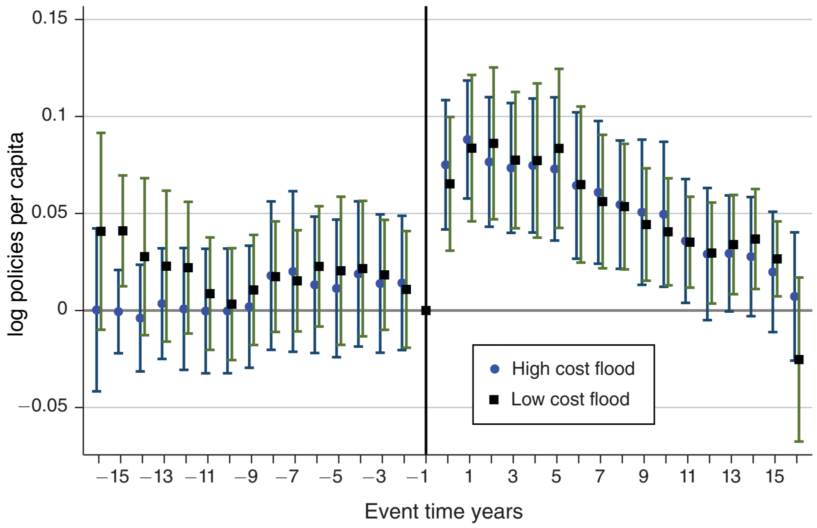

Figure 5. Flood Insurance Take-Up for Communities in PDD Counties after High- and Low- Cost Floods 1980–2007

Notes: The figure plots event time insurance take-up coefficients from estimation of a version of equation (1) that separately identifies floods as above (circles) or below (squares) per capita median cost on the 1980–2007 panel. Per capita cost is calculated over all 836 floods from 1980–2007 by dividing (a measure of) total PDD cost by the total population living in the effected counties in the year of a flood. The per capita cost ranges from less than \$1 to \$12,440, with a mean of \$70, and a median of \$20. Costs include all Public Assistance and Individual Assistance paid out after a flood. Please refer to the notes to Figure 4 for more details on the event study, and online Appendix Section D for a detailed cost data description.

Source: Public Entity Risk Institute

Flood Costs.—This paper assumes that homeowners use the new flood information to update their conditional yearly flood probability. It is also possible that homeowners use the new floods to update their expectations over flood damage. Figure 5 plots the take-up coefficients from the estimation of a version of equation (1) that separately identifies floods as above or below per capita median cost.[33] The dots (squares) plot above (below) median coefficients. Insurance take-up is very similar after a flood regardless of whether the flood is high or low cost. There is no statistically significant difference between any of the pairs of post-flood coefficients. Homeowners interpret the information provided by high- and low-cost floods the same and do not appear to use new floods to learn about expected flood damages.

Migration.—Migration is a potential explanation for the spike and dissipation of insurance after a new flood, but for migration to explain the pattern of insurance take-up two things must be true. First, there is enough population turnover so that there is always a pool of newer residents. Second, these newer residents are unaware

[33] Per capita cost is calculated over all 836 floods from 1980–2007 by dividing (a measure of) total PDD cost by the total population living in the effected counties in the year of a flood. The per capita cost ranges from less than \$1 to \$12,440, with a mean of \$70 and a median of \$20. Costs include all Public Assistance and Individual Assistance paid out after a flood (Public Entity Risk Institute). Please refer to online Appendix Section D for a detailed data description.

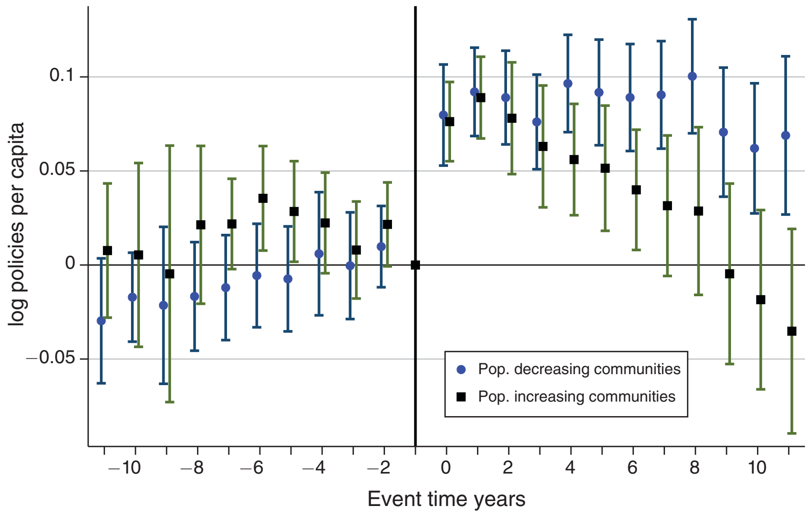

Figure 6. Flood Insurance Take-Up for Population Increasing and Population Decreasing Communities 1990–2007

Notes: The figure plots event time insurance take-up coefficients from estimation of a version of equation (1) that separately identifies floods that hit population increasing (circles) and decreasing (squares) communities on the 1990–2007 panel. A community is defined as having an increasing (decreasing) population if its population grew (shrank) between 1990 and 2007. 6,113 (56 percent) of the 10,841 communities have a growing population. Refer to the notes to Figure 2 for more regression details.

of the flooding history and there must be a sufficiently high cost to obtaining this information. Section IV shows that a Bayesian learning model that meets these two criteria could explain observed insurance take-up.[34]

There is mixed evidence on the role migration plays in accounting for the observed pattern of insurance take-up. On one hand, insurance take-up differs between population-increasing and -decreasing communities from 1990–2007.[35] Figure 6 plots estimates of insurance take-up using a version of equation (1) that divides communities into population-increasing (squares) or -decreasing (circles) communities from 1990–2007. The population-increasing communities have a larger share of newer residents relative to the population-decreasing communities. Insurance take-up jumps the same for both groups of communities after a flood. Take-up in the population-increasing communities quickly declines, while that in the population-decreasing communities remains relatively flat at the higher level.

On the other hand, communities in high-migration counties do not have a larger insurance take-up rate after a flood. We divide counties into quartiles based on the average yearly county in-migration rate from 1984–2007.[36] We run the same event

[34] Interestingly, it is not necessary that newer residents initially underestimate the true flood probability if these residents only consider the recent (shorter) flood history.

[35] Annual community-level migration data are not available.

[36] Counties in the first quartile have the lowest average annual in-migration rate, while counties in the fourth quartile have the highest annual in-migration rate. The migration data used in the event study analysis described are from the IRS county-to-county migration files. County-to-county migration files are not available for 1983. Thus, 1984–2007 is the longest uninterrupted panel. Please refer to online Appendix Section D for more details.

study model (equation 1), except that we use the estimation period 1984–2007 and include a separate set of event time indicator variables for high-migration and low-migration counties. The coefficient point estimates for post-flood insurance take-up are larger for the low-migration counties, but not statistically different than those of the high-migration counties.[37]

There is also no county-level evidence that flooding leads to greater migration. Again, we use equation (1) and the estimation period 1984–2007, and consider as our dependent variable both migration and log migration. This finding is consistent with another recent paper that fails to find evidence of migration from counties hit by hurricanes (Deryugina 2011).

Protective Measures.—Community-wide flood protective measures could potentially explain the observed pattern of flood insurance take-up shown in Figures 2 and 4. A community may initiate protective measures after being hit by a flood that reduce the likelihood of future floods. If this occurs, residents may be more inclined to self-insure in the years immediately following a flood before any community-wide structural changes are complete.[38]

Three pieces of evidence suggest that community-wide protective measures are not an important factor in explaining the observed pattern of insurance take-up. First, online Appendix Figure 3 shows that the same insurance take-up spike and decay pattern repeats for sequential floods that hit the same community. Second, the vast majority of the large-scale flood control projects were completed before 1980 (Graf 1999).[39] Third, very few communities participate in a NFIP program that seeks to incentivize better community flood plain management. Among those communities that do participate, there is no evidence that recent floods lead to increased participation.[40]

B. *Neighboring Communities*

The 1980–2007 panel is used to estimate the effect of a "nearby" flood on insurance take-up. We consider two definitions of proximity to a flood for nonflooded communities: geographic distance and media exposure. We vary the definition of a geographically neighboring community as one in either an adjacent (nonflooded) PDD county, or in the closest 1, 5, 10, or 20 (nonflooded) counties. For ease of exposition, the text focuses on communities in the closest five counties. The results are very similar regardless of the definition of a geographic neighbor.[41] A media market

[37] This result is not sensitive to whether we compare the top/bottom quartiles, or above/below median. Online Appendix Figure 2 plots the point estimates from the above/below median migration regression.

[38] We focus on community-wide protective measures because it is unlikely that individual property owners can alter their property to avoid being flooded by the type of large regional floods evaluated in the paper.

[39] Graf (1999) examines the National Inventory of Dams and concludes: "*Water resource regions have experienced individualized histories of cumulative increases in reservoir storage* (*and thus of downstream hydrologic and ecologic impacts*), *but the most rapid increases in storage occurred between the late 1950s and the late 1970s. Since 1980, increases in storage have been relatively minor.*" [p.1] Importantly, Graf's definition of a dam includes flood control projects such as storm protection works in coastal Florida.

[40] Online Appendix Section D provides details on the Community Rating System (CRS) program, and online Appendix Section E discusses event study results that control for CRS community participation.

[41] Please refer to online Appendix Section E and Tables 8–13 for a detailed discussion of all geographic neighbor results, and to online Appendix Section D for details on the neighbor data sources.

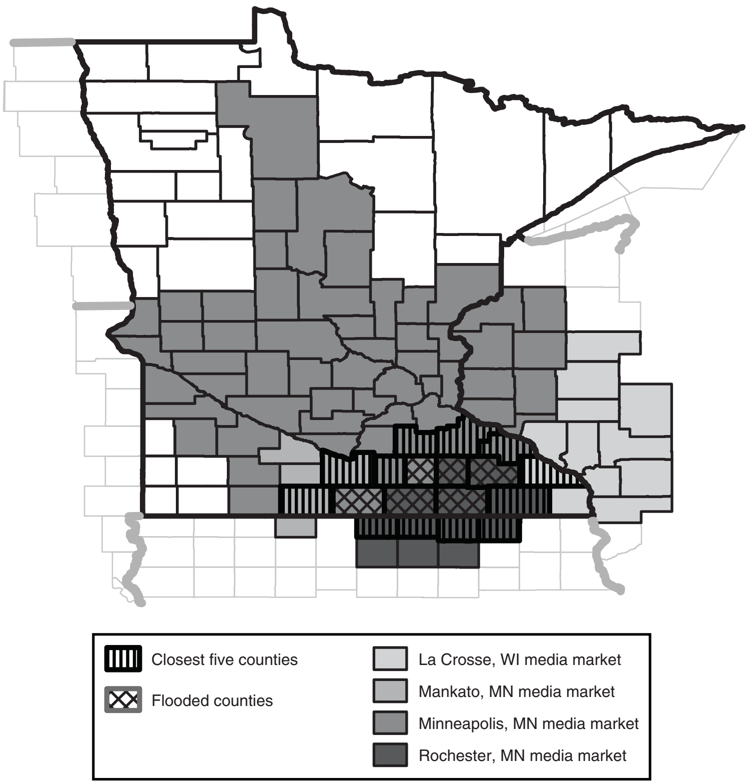

Figure 7. TV Media Market and Geographic Neighbor Minnesota Identification Example

Notes: The figure shows the state of Minnesota outlined in black. In 2004, six counties (marked by crossing lines) in Minnesota had a PDD flood. The parallel vertical lines indicate counties that are among the five closest counties to a flooded county and also not flooded. The flooded and (five closest) geographic neighbor counties are part of four different media markets. Counties in the four media markets are denoted by shades of grey. Closest counties are determined by Euclidean distance between county centroids. Nielson Media Research classifies each US county as belonging to a primary television media market. Please refer to Section IIIB for details.

neighbor is a nonflooded community that shares the same TV media market as a flooded community. Nielson Media Research classifies each US county as belonging to a primary TV media market. The 1980–2007 panel includes 212 Designated Media Markets (DMAs). Importantly, local news programming differs by media market.

There are at least two reasons why we may expect homeowners in geographically neighboring communities to increase insurance take-up after a flood. First, there is likely to be some flooding in the region surrounding the most severely impacted flood areas. Second, if geographic areas share similar flood risks, then homeowners could use nearby flooding to learn about their own flood risk. Local TV news is a potential source of general flood risk information, but also a mechanism to learn about new nearby floods.

Media Market and Geographic Neighbor Identification.—Figure 7 provides an example to help clarify the distinction between flooded counties, geographic neighbors, and media neighbors. Figure 7 shows the state of Minnesota outlined in black. In 2004, six counties (marked by crossing lines) in Minnesota had a PDD flood.

The parallel vertical lines indicate counties that are among the five closest counties to a flooded county and also not flooded. The flooded and (five closest) geographic counties are part of four different media markets. Counties in the four media markets are denoted by shades of grey. The white counties on the map are counties that are part of other media markets. In general, the media markets are spatially much larger than the flooded and five closest geographically neighboring counties. For example, the Minneapolis, MN media market ranges more than 400 miles from the border with Iowa in the south to nearly the Canadian border in the north.

We use the spatial mismatch between the geographic proximity to a flood and the coverage area of the TV media markets to estimate whether homeowners in neighboring nonflooded communities react to a nearby flood by purchasing insurance. We separately measure the insurance take-up effect for homeowners living in communities that are close to a flood, in the same TV media market as a flood, or close and in the same TV media market. The empirical strategy used to separately identify the role of local TV news media from that of the geographic proximity is similar to Snyder and Strömberg (2010).[42]

Neighbor Event Study Results.—Figure 8, panels A–D show postflood insurance take-up for flooded communities (circles), geographic neighbors (squares), and media neighbors (triangles) from four separate regressions using equation (2) and the 1980–2007 panel. For space considerations, only the event study coefficients corresponding to the year of a flood and the first ten post-flood years are displayed.[43] The bars around each neighbor point estimate show the 95 percent confidence interval. All of the flooded point estimates are significant at the 1 percent level.

Panel A plots the coefficient estimates (squares) from an event study that includes indicators for geographic neighbors. Insurance take-up peaks at 2.5 percent and is statistically significant at the 5 percent level for the first three years after a flood. Panel B plots the coefficient estimates (triangles) from an event study that includes indicators for media market neighbors. The media neighbor point estimates for the first 5 years after a flood range between 2.8 percent and 3.6 percent and are statistically significant at the 1 percent level. These point estimates are about one-third as large as those for flooded communities (circles). In panel C, the media take-up effect is virtually unchanged and the geographic neighbor effect mostly disappears when both sets of neighbor indicators are included in the same event study.

Panel D further explores these findings by isolating homeowner take-up in those communities that share a media market but are not geographically close to the flood (triangles), and the take-up effect in communities close to a flood but in a different media market (squares). The media coefficient estimates in panel D are very similar

[42] Snyder and Strömberg (2010) use the spatial mismatch between political jurisdictions and newspaper coverage to estimate how citizen knowledge affects politicians' actions. I thank James Snyder for sharing the DMA data (first used by Ansolabehere, Snowberg, and Snyder 2006), and Ansolabehere, Gerber, and Snyder (2001).

[43] The preperiod neighbor indicators are not statistically significant. State-by-year fixed effects flexibly control for changing calendar year factors that might be correlated with insurance take-up, but exclude cross-state identification (shown to be an important source of variation in Figure 7). For this reason, the regressions that examine take-up in neighboring communities (Figure 8 and online Appendix Tables 8–13) use larger end bins to improve statistical power without changing the interpretation of the coefficients of interest.

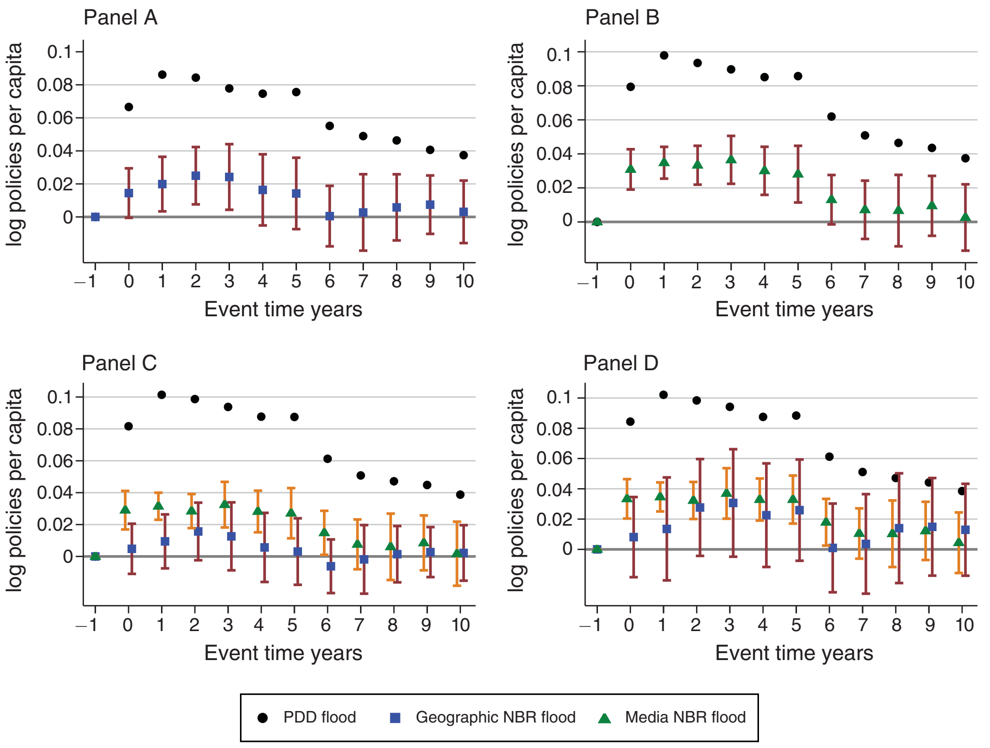

Figure 8. Flood Insurance Take-Up for Geographic and Media Neighbors

Notes: Each panel contains coefficients from a distinct event study regression using a version of equation (2) and the 1980–2007 panel. Panel A includes event time indicators for communities located in one of the five closest nonflooded counties. Panel B includes event time indicators for nonflooded communities located in the same TV media market as a flooded community. Panel C includes both geographic and media indicators. Panel D includes both geographic and media indicators, and their interaction (not displayed). Please refer to Section IIIB and online Appendix Section D for further details.

to panel C. There is no difference in take-up among nonflooded homeowners in the same media market based on geographic proximity to the flood. There is some evidence of increased take-up in geographically close communities not in the same media market. This take-up is driven exclusively by homeowners in communities just outside the PDD flooded counties.[44]

The results in Figure 8 do not depend on the definition of a geographic neighbor.[45] The estimates and statistical significance for the media neighbor coefficients are remarkably stable and always statistically significant at the 1 percent level up until the first 5 years after a flood. This is true regardless of whether the event study

[44] Online Appendix Table 13, columns 1 and 2 divide geographic neighboring communities into those in the closest nonflooded county and those in the closest two to five nonflooded counties. The point estimate is 4.6 percent and statistically significant for both the second and third years after a flood for communities in the closest county. The same estimates for communities in the closest two to five counties are 1.0 percent and 1.6 percent, and are not significant.

[45] Online Appendix Tables 8–10 show the results for the same event study specifications as in Figure 8, using the following geographic neighbor definitions: a community in either an adjacent county or the closest 1, 5, 10, or 20 (nonflooded) counties. Online Appendix Tables 11–13 show the results for geographic neighbor "rings" (1, 2–5, 6–10, or 11–20 counties). Online Appendix Section E provides a detailed discussion of these results.

controls for geographic neighbors or isolates media neighbors that are not also geographic neighbors. The postflood geographic neighbor coefficients also display a similar pattern as those in Figure 8, panel C: small coefficient estimates and no (or only marginal) statistical significance after controlling for the media market. The notable exception is for the communities in the single closest geographic county just outside the worst flooded counties. Take-up in these communities is similar to that of nonflooded communities in the same media market as a PDD flood.

We also estimate whether the TV media effect is greater for nonflooded homeowners when a greater share of the media market is flooded (Snyder and Strömberg 2010). The hypothesis is that if a larger share of the media market is flooded, then there is likely to be more flood information (e.g., news stories) conveyed through the local TV media. We create two new flooded media market "congruence" variables that range from zero to one based on the share of the media market counties (or population) that is flooded by a particular PDD flood.[46]

Panel A of Table 1 displays year of flood take-up coefficients from three separate regressions using equation (2) and the 1980–2007 panel. Each regression focuses on the media neighbor effect. Column 1 repeats the same specification as Figure 8, panel B. In the year of a flood, there is an estimated 7.9 percent increase in insurance for homeowners in flooded communities and 3.1 percent increase for media neighbors. Columns 2 and 3 add the media market congruence variables. The congruence variables are positive and statistically significant in both specifications. The greater the share of the media market covered by a PDD flood, the higher is flood insurance take-up in nonflooded areas of the media market. Panel B calculates the implied insurance take-up at the median. The median population in a media market flooded by a PDD flood is 36 percent. Summing the congruence effect at the median with the media market event study coefficient yields an implied media neighbor total effect of 3.4 percent. This implied effect is very similar to the baseline estimate of 3.1 percent.

TV News Story Evidence.—Local TV media markets provide variation in information about, and exposure to, large floods. In the five years from 2003–2007, local ABC, CBS, NBC, and FOX affiliate news stations in media markets that had at least one county included as part of a flooding PDD for the calender year had more than three times as many news stories on large floods relative to markets without a flood. There were 4.3 times as many news stories on floods where a larger share (above median) of the market population was flooded, and 2.3 times as many stories for floods where a smaller (below median) share of the population was flooded.[47]

Interpretation.—Homeowners could use new nearby floods to learn about their own flood risk. First, the new floods could change a homeowner's understanding of

[46] Measuring the "congruence" between the geographic area of the media market and the geographic area of the flood is a direct application of a strategy proposed by Snyder and Strömberg (2010). I thank the editor for recommending this analysis.

[47] Flood-related news stories are determined by a text-based search of the transcriptions of the local news broadcasts. Online Appendix Section D provides more details. As a robustness check, I also consider whether there are fewer flood news stories when a flood occurs at the same time as other newsworthy events. Panel regression estimates suggest some crowding out of flood news when floods and important national media events occur in the same month (Appendix Section E.9 and Tables 14 and 15).

Table 1—TV Media Market and Presidential Disaster Declaration Flood Congruence

	Baseline (1)	Pop. congruence (2)	County congruence (3)
Panel A. Event study coefficients			
Year of flood: flooded	0.079***	0.064***	0.062***
	(0.012)	(0.012)	(0.015)
Year of flood: media neighbor	0.031***	0.024***	0.022***
	(0.006)	(0.007)	(0.006)
Population TV media market congruence		0.028*	
		(0.016)	
County TV media market congruence			0.041**
			(0.020)
R^2	0.206	0.206	0.206
Communities	268,996	268,996	268,996
Panel B. Implied insurance takeup			
Median congruence (conditional on media market flood)		0.36	0.28
Implied flooded effect at the median		0.087	0.073
Implied media neighbor effect at the median		0.034	0.033

Notes: Panel A, columns 1–3, display select coefficients from estimation of equation (2) with media neighbors on the 1980–2007 panel. Column 1 reproduces the specification of Figure 8, panel B. Columns 2 and 3 run the same specification, except add the population and county congruence variables. The population (county) congruence variable measures the proportion of the population (counties) in the media market hit by a PDD flood. Panel B calculates the total hit and media neighbor implied insurance take-up effects by summing the congruence effect at the median with the relevant year of flood effect.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

the general background risk of a flood. Second, we might expect a nearby flood to be of differential importance to residents living in communities that share similar flood characteristics or flood histories. For example, a coastal flood due to a storm surge after a hurricane would be more informative about a nonflooded coastal community's flood risk than a community many miles from the ocean.

Two pieces of evidence suggest that homeowners do not update their flood expectations based on the relevancy of a nearby flood. First, Figure 8, panels C and D show that nonflooded homeowners react to a flood in the media market the same even when they live geographically "far" from the flooded community.[48] The point estimates for postflood media neighbor take-up are virtually identical in panels C and D. This result is surprising if geographically close communities are more likely to share similar flood characteristics. If geographically close communities share similar flood characteristics then differences in flooding are likely due to randomness. We would expect homeowners in geographically close communities to have larger take-up rates after a flood.

[48] This result is even more striking when using the 20 closest county geographic neighbor definition (comparing column 4 of onlineAppendix Tables 9 and 10). Communities in the same media market that are not among the 20 closest counties are often hundreds of miles from the flood, yet insurance take-up is the same.

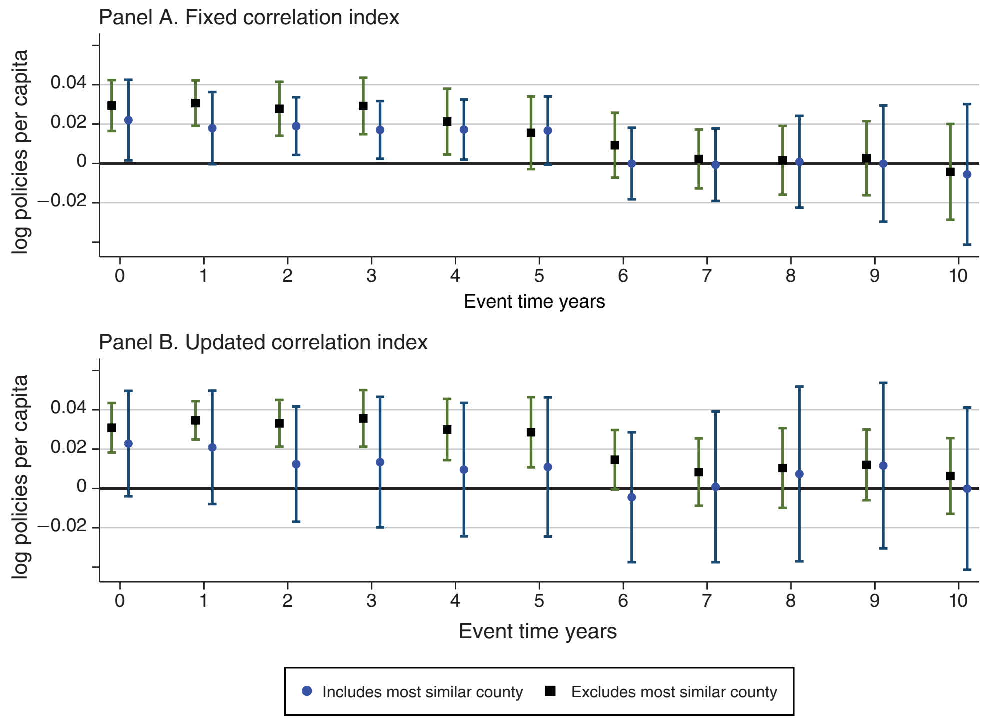

Figure 9. Insurance Take-Up for Media Neighbors by whether the Flood Includes the County with the Most Similar Flood History

Notes: Each panel plots media market neighbor insurance take-up coefficients from a distinct regression using a version of equation (2). The regression specification in the figure is the same as that of Figure 8, panel B, except that media market floods are divided into two types. The dots (squares) indicate insurance take-up for floods that include (exclude) the county with the most similar flood history. Panel A determines the county with the most similar flood history for each neighbor county using a flood correlation index and the years 1958–2007. Panel B defines the historical flood correlation index using only those years before the year of the current flood. Please refer to Section IIIB for further details.

Second, Figure 9 tests whether homeowners in a nonflooded community take up insurance after a flood at greater rates if a flooded community shares a similar flood history. I divide media market floods into two groups by asking the following question: is the county with the most similar PDD flood history to the nonflooded homeowner's county flooded? I estimate a version of equation (2) that separately considers the two types of media market floods. The two panels of Figure 9 use two distinct historical county correlation measures. Panel A uses the 50-year period (1958–2007). The assumption is that these 50 years are a representative time period that approximates the true underlying yearly flood correlation between counties in the same media market. Panel B only considers years before the most recent flood and therefore allows homeowners to learn about which county shares the most similar flood risk. Each panel displays the postflood insurance take-up point estimates for media market floods that include (dots) or do not include (squares) the county with the most similar flood history. Overall, the point estimates are higher for floods that do not include the most similar county. However, there is no statistical difference between any of the point estimate pairs for the two types of media market floods. Again, this is surprising if we believe that the difference between flooding in two communities with very similar flood histories is due to randomness, and not differences in community flood risk characteristics.

IV. Discussion

A large and immediate change in beliefs after a disaster could be consistent with the common Bayesian learning model (Viscusi 1991). In the "Full Information" Beta-Bernoulli Bayesian Model, homeowners observe whether there is a flood in a given year and update their expectation of a future flood (DeGroot 1970; Davis 2004; Card 2010). Each community's yearly flood draw is assumed to be independently drawn from a stationary flood distribution with parameter p. The probability of a flood in a given year, p, is assumed to be distributed $Beta(\alpha, \beta)$.[49] A homeowner's conditional expectation of their yearly flood probability p is

$$E[p\,|\,S_t, t] \;=\; \frac{S_t + \alpha}{t + \alpha + \beta}, \tag{3}$$

where t is the number of yearly observations (time periods), $S_t = \sum_{s=1}^{t} y_s$ is the number of observed floods, and α and β are fixed parameters and determine the initial belief over flooding. Equation (3) implies that, as the stock of information increases, the effect of a new observation will become small (and eventually zero).

The large spike in insurance take-up after a flood combined with the relatively fast decay of this effect suggests that homeowners may not be considering all of the past flood information. There are two possibilities for why homeowners do not consider all of the past flood information: homeowners don't observe the whole history, or homeowners forget. One way to model this pattern is with a weighting parameter that discounts past information (Camerer and Ho 1999; Malmendier and Nagel 2011). In such a model, the stock of information never becomes so large as to rule out a large jump in the conditional expectation of a future flood. While the immediate impact of new information can be large, its impact on expectations quickly lessens, implying a steeper post-flood slope.

Equation (4) is a learning model that allows homeowners to discount past floods:

$$E[p\,|\,S'_t, t'] \;=\; \frac{S'_t + \alpha}{t' + \alpha + \beta}. \tag{4}$$

$S'_t = \sum_{s=1}^{t} y_s \delta^{t-s}$ are weighted flood observations and $t' = \sum_{s=1}^{t} \delta^{t-s}$ is the number of yearly observation "equivalents." $\delta \in [0, 1.05]$ is a weighting parameter. When $\delta < 1$, older floods are weighted less than more recent floods when updating conditional beliefs about a future flood. Equation (4) reduces to the Full Information model (equation 3) when $\delta = 1$.

[49] The Beta distribution is the conjugate prior for the Bernoulli distribution (DeGroot 1970) and used in most Bernoulli Bayesian models for convenience. PDD county-level flooding in the United States from 1958–2007 closely fits the Beta distribution (see online Appendix Figure 4).

The parameters α and β determine a homeowner's initial belief over the probability of a flood. I consider three different approaches to setting initial beliefs. The first (second) approach assumes that homeowners set their initial flood expectation equal to the mean flood probability of a county from the national (state) county flood distribution. These two approaches match the first two moments of the empirical flood distribution with the first two moments of the Beta Distribution (Davis 2004). The third approach assumes that homeowners only consider the flooding history of their county and allows the certainty of their prior to vary.[50]

I simulate conditional flood probabilities using the panel of PDD floods under both the Full Information and Discounting models. The purpose of the simulation is to provide evidence on how well each model matches the observed pattern of insurance take-up.[51] I use equation (4) to simulate probability time series given different values for the weighting parameter. I then select the time series of flood probabilities, $p(\delta)_{ct}$, that minimizes the mean square error of: $\ln(takeup_{ct}) = \alpha + \beta_t \ln p(\delta)_{ct} + \alpha_c + \gamma_{st} + \epsilon_{ct}$. This equation is the same as the baseline estimating equation, except here we replace the event time dummy variables with log flood probability. A minimum distance estimator is used to gauge the learning model fit (Abowd and Card 1989; Chamberlain 1982; Farber and Gibbons 1996).

The fit of each model is determined by observing how well the changes in simulated probabilities match the changes in insurance take-up in the years preceding and following a flood. It is important to remember that the event study framework controls for the different flooding histories for each community, while focusing attention on how conditional flood probabilities change after a new flood under each assumed learning model.

The Full Information Bayesian model does a poor job of matching insurance take-up. The model cannot match both the size of the immediate jump in insurance purchases and the speed of the decline back to baseline. In general, the Full Information Model predicts a smaller jump and a slower decline.[52] Among those starting value model parameterizations that provide an acceptable fit at the 5 percent significance level, the best fitting model for each parameterization is always one where homeowners discount older information ($\delta < 1$).

Finally, I also consider a second type of incomplete information Bayesian model where homeowners only have access to flood information if they reside in the county at the time of a flood. I calibrate this second incomplete information Bayesian model using national county migration flows. I use IRS county migration data and calculate the average migration rate (across both counties and years) from 1980–2007 to be 5.5 percent. I then use this migration rate to create a cohort-based migration

[50] The third approach sets the first moment of the Beta Distribution equal to the mean yearly probability of a flood $\left(E[p] = \frac{\alpha}{\alpha+\beta}\right)$ for each county for the years 1958–2007. We consider α, β combinations that fit this equation, by varying $\alpha \in (0, 15]$. No model simulation with an $\alpha > 10$ provides a statistical fit for the observed pattern of insurance take-up. By construction, a smaller α implies a smaller β. Together, α and β close to zero imply a highly uncertain prior belief. When α and β are small, the initial beliefs are "weak" and homeowners will almost ignore their initial beliefs when updating expectations.

[51] What follows is a short overview of the learning model probability simulation and insurance take-up comparison. Please refer to the online Appendix F for a detailed discussion.

[52] For example, see online Appendix Figure 5.

profile for the "typical" community. The calibration is meant as a benchmark and not to accurately account for migration differences over time or between counties.

Flood probabilities for the migration-calibrated model are simulated using equation (4), except that homeowners only consider flood information from their years of residence.[53] Again, the best-fitting model is always one where $\delta < 1$. However, for many of the parameterizations, a model with $\delta = 1$ can no longer be statistically rejected. A learning model without discounting, but where homeowners still have incomplete information due to migration can match the spike and decay pattern of flood insurance take-up.

There are several possible underlying interpretations. Availability Bias is consistent with the available evidence and is a non-Bayesian interpretation. There are also at least three learning model interpretations. First, homeowners could have the mistaken belief that past floods are less relevant for understanding their current flood risk.[54] One reason why past floods could be perceived as less important by homeowners is that they are less likely to have had personal experience with past floods. That is, the experience of being flooded leads homeowners to interpret the statistical information differently (e.g., Haselhuhn et al. 2012). Second, the same homeowners could be learning and forgetting (e.g., Agarwal et al. 2008). Third, if accessing past information involves a high cost, it could be completely rational to ignore this information (e.g., Sims 2010; Maćkowiak and Wiederholt 2012).[55] For example, in the migration-calibrated incomplete information model, floods that occur before a homeowner arrives carry so little weight in the decision-making process that they can actually be ignored.[56] While the county-level migration event study results do not support this interpretation, there is evidence that insurance take-up in communities with longer-tenured residents is more persistent than in communities with shorter-tenured residents.

V. Conclusion

We provide new evidence on how individuals update their beliefs over an uncertain and infrequent risk using a new panel dataset of large regional floods and the take-up of flood insurance in the United States. We find that, after controlling for calendar time trends and location fixed effects, the take-up of insurance is completely flat in the years before a flood, spikes immediately following a flood, and then steadily declines back to baseline. Robustness checks of the model show that changing insurance prices, changing homeowner income, potential serial correlation

[53] For example, a homeowner from a recently migrated cohort who has lived in the county for five years will only consider the past five years when updating expectations (events more than five years ago will have a weight of $\delta = 0$).

[54] Past floods would be less important if there were either annual correlation in floods, or a nonstationary flood probability (neither of which are true for PDD floods from 1958–2007). Please refer to online Appendix Sections B.6 and C for an extended discussion.

[55] The difference between the second and third interpretations is whether the same homeowners are both learning and forgetting (second interpretation), or there are different cohorts of homeowners that respond differently (third interpretation). One way to test the learning and forgetting interpretation would be to compare new and renewing policy holders, but unfortunately these data are not available.

[56] The migration evidence is also consistent with the first interpretation, where what is most important is experience with a flood.

in floods, and different flood costs are unlikely to explain the observed pattern in insurance take-up.

We also show that the news media affects how information on environmental risks is acquired and processed by homeowners not directly impacted by a flood. Those homeowners not flooded, but in the same TV media market as a flooded community, exhibit a spike in insurance purchases that is one-third as large as the spike in flooded communities. Nonflooded homeowners in the same TV media market take-up insurance at the same rate regardless of how relevant the TV flood news is towards understanding their own flood risk.

The large jump in insurance take-up implies that homeowners do not make a one time decision of whether to purchase flood insurance based, for example, on FEMA maps or engineering estimates. The large jump combined with the quick decay to baseline levels can not be explained by a Bayesian model where homeowners have full information of historical flooding and weigh each past flood observation equally. Overall, a learning model that discounts past floods does a good job of describing the observed pattern of flood insurance take-up. There are several possible underlying interpretations, including availability bias. There is modest support for the role of migration in any learning model interpretation. Either homeowners don't know about floods that occurred before they arrive in a community, or the experience of living through a flood leads homeowners to treat recent floods differently.

REFERENCES

Abowd, John M., and David Card. 1989. "On the Covariance Structure of Earnings and Hours Changes." *Econometrica* 57 (2): 411–45.

Agarwal, Sumit, John C. Driscoll, Xavier Gabaix, and David Laibson. 2008. "Learning in the Credit Card Market." National Bureau of Economic Research (NBER) Working Paper 13822.

American Insurance Association. 1956. "Studies of Floods and Flood Damage, 1952–1955. American Insurance Association: New York.

Anderson, Dan R. 1974. "The National Flood Insurance Program—Problems and Potential." *Journal of Risk and Insurance* 41 (4): 579–99.

Ansolabehere, Stephen, Alan S. Gerber, and James M. Snyder, Jr. 2001. "Does TV Advertising Explain the Rise of Campaign Spending? A Study of Campaign Spending and Broadcast Advertising Prices in US House Elections the 1990s and the 1970s." http://economics.mit.edu/files/1211.

Ansolabehere, Stephen, Eric C. Snowberg, and James M. Snyder, Jr. 2006. "Television and the Incumbency Advantage in U.S. Elections." *Legislative Studies Quarterly* 31 (4): 469–90.

Bin, Okmyung, Jamie Brown Kruse, and Craig E. Landry. 2008. "Flood Hazards, Insurance Rates, and Amenities: Evidence from the Coastal Housing Market." *Journal of Risk and Insurance* 75 (1): 63–82.

Bin, Okmyung, and Craig E. Landry. 2013. "Changes in Implicit Flood Risk Premiums: Empirical Evidence from the Housing Market." *Journal of Environmental Economics and Management* 65 (3): 361–76.

Browne, Mark J., and Robert E. Hoyt. 2000. "The Demand for Flood Insurance: Empirical Evidence." *Journal of Risk and Uncertainty* 20 (3): 291–306.

Camerer, Colin, and Teck Hua Ho. 1999. "Experience-Weighted Attraction Learning in Normal Form Games." *Econometrica* 67 (4): 827–74.

Card, David. 2010. "Models of Thinking, Learning, and Teaching in Games." Lecture Notes. University of California, Berkeley.

Chamberlain, Gary. 1982. "Multivariate Regression Models for Panel Data." *Journal of Econometrics* 18 (1): 5–46.

Davis, Lucas W. 2004. "The Effect of Health Risk on Housing Values: Evidence from a Cancer Cluster." *American Economic Review* 94 (5): 1693–1704.

DeGroot, Morris H. 1970. *Optimal Statistical Decisions.* 1st ed. New York: McGraw-Hill.

DellaVigna, Stefano, and Ethan Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *Quarterly Journal of Economics* 122 (3): 1187–1234.

Deryugina, Tatyana. 2011. "The Role of Transfer Payments in Mitigating Shocks: Evidence from the Impact of Hurricanes." http://deryugina.com/Deryugina_SafetyNet_Dec2011.pdf.

Deryugina, Tatyana. 2013. "How Do People Update? The Effects of Local Weather Fluctuations on Beliefs about Global Warming." *Climatic Change* 118 (2): 397–416.

Dixon, Lloyd, Noreen Clancy, Seth A. Seabury, and Adrian Overton. 2006. *The National Flood Insurance Program's Market Penetration Rate*. Santa Monica: RAND Corporation.

Downton, Mary W., and Roger A. Pielke, Jr. 2001. "Discretion without Accountability: Politics, Flood Damage, and Climate." *Natural Hazards Review* 2 (4): 157–66.

Driscoll, John C., and Aart C. Kraay. 1998. "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data." *Review of Economics and Statistics* 80 (4): 549–60.

Farber, Henry S., and Robert Gibbons. 1996. "Learning and Wage Dynamics." *Quarterly Journal of Economics* 111 (4): 1007–47.

Federal Emergency Management Agency, and Federal Insurance and Mitigation Administration. 2002. *National Flood Insurance Program, Program Description*. Federal Emergency Management Agency, and Federal Insurance and Mitigation Administration. August.

Federal Emergency Management Agency (FEMA). 2002a. *A chronology of major effects affecting the national flood insurance program.* Technical Report 28298-0029. Federal Emergency Managment Agency. October.

Federal Emergency Management Agency (FEMA). 2002b. *National Flood Insurance Program, Program Description.* Federal Emergency Managment Agency. August.

Federal Emergency Management Agency (FEMA). 2007. *National Flood Insurance Program, Mandatory Purchase of Flood Insurance.* Federal Emergency Management Agency. September.

Federal Emergency Management Agency (FEMA). 2008. *National Flood Insurance Program Flood Insurance Manual.* Federal Emergency Management Agency. May.

Ferraz, Claudio, and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics* 123 (2): 703–45.

Gallagher, Justin. 2014. "Learning about an Infrequent Event: Evidence from Flood Insurance Take-Up in the United States: Dataset." *American Economic Journal: Applied Economics.* http://dx.doi.org/10.1257/app.6.3.206.

Gentzkow, Matthew, Jesse M. Shapiro, and Michael Sinkinson. 2011. "The Effect of Newspaper Entry and Exit on Electoral Politics." *American Economic Review* 101 (7): 2980–3018.

Gerber, Alan S., Dean Karlan, and Daniel Bergan. 2009. "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions." *American Economic Journal: Applied Economics* 1 (2): 35–52.

Graf, William L. 1999. "Dam Nation: A Geographic Census of American Dams and Their Large-Scale Hydrologic Impacts." *Water Resources Research* 35 (4): 1305–11.

Haselhuhn, Michael P., Devin G. Pope, Maurice E. Schweitzer, and Peter Fishman. 2012. "The Impact of Personal Experience on Behavior: Evidence from Video-Rental Fines." *Management Science* 58 (1): 52–61.

Ho, Teck-Hua, and Juin-Kuan Chong. 2003. "A Parsimonious Model of Stockkeeping-Unit Choice." *Journal of Marketing Research* 40 (3): 351–65.

Insurance Research Council, and Insurance Institute for Property Loss Reduction. 1995. *Coastal Exposure and Community Protection: Hurricane Andrew's Legacy*. Insurance Institute for Property Loss Reduction. Boston, April.

Jaffee, Dwight, Howard Kunreuther, and Erwann Michel-Kerjan. 2008. "Long Term Insurance (LTI) for Addressing Catastrophe Risk." National Bureau of Economic Research (NBER) Working Paper 14210.

Kahneman, Daniel. 2011. *Thinking, Fast and Slow.* New York: Farrar, Straus and Giroux.

Kousky, Carolyn. 2010. "Learning from Extreme Events: Risk Perceptions after the Flood." *Land Economics* 86 (3): 395–422.

Kriesel, Warren, and Craig Landry. 2004. "Participation in the National Flood Insurance Program: An Empirical Analysis for Coastal Properties." *Journal of Risk and Insurance* 71 (3): 405–20.

Kunreuther, Howard. 1976. "Limited Knowledge and Insurance Protection." *Public Policy* 24 (2): 227–61.

Kunreuther, Howard. 1995. *Coastal Exposure and Community Protection.* Boston: Hurrican Andrew's Legacy Insurance Institute for Poverty Loss Reduction: Boston.

Kunreuther, Howard. 1996. "Mitigating Disaster Losses Through Insurance." *Journal of Risk and Uncertainty* 12 (2–3): 171–87.

Kunreuther, Howard C., Ralph Ginsberg, Louis Miller, Philip Sagi, Paul Slovic, Bradley Borkan, and Norman Katz. 1978. *Disaster Insurance Protection: Public Policy Lessons.* New York: John Wiley.

Kunreuther, Howard, Robert Meyer, and Erwann Michel-Kerjan. 2009. "Overcoming Decision Biases to Reduce Losses from Natural Catastrophes." http://opim.wharton.upenn.edu/risk/library/C2009_HK,RJM,EMK.pdf.

Maćkowiak, Bartosz, and Mirko Wiederholt. 2012. "Inattention to Rare Events." http://www.crei.cat/conferences/behavioral-decision-theory/wiederholt.pdf.

Malmendier, Ulrike, and Stefan Nagel. 2011. "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?" *Quarterly Journal of Economics* 126 (1): 373–416.

Michel-Kerjan, Erwann O. 2010. "Catastrophe Economics: The National Flood Insurance Program." *Journal of Economic Perspectives* 24 (4): 165–86.

Michel-Kerjan, Erwann, and Carolyn Kousky. 2008. "Come Rain or Shine: Evidence on Flood Insurance Purchases in Florida." University of Pennsylvania Wharton School Working Paper 2008-02-29.

Michel-Kerjan, Erwann, Sabine Lemoyne de Forges, and Howard Kunreuther. 2012. "Policy Tenure Under the U.S. National Flood Insurance Program (NFIP)." *Risk Analysis* 32 (4): 644–58.

Palm, Risa. 1995. "The Roepke Lecture in Economic Geography Catastrophic Earthquake Insurance: Patterns of Adoption." *Economic Geography* 71 (2): 119–31.

Sims, Christopher A. 2010. "Rational Inattention and Monetary Economics." In *Handbook of Monetary Economics*, Vol. 3, edited by Benjamin M. Friedman and Michael Woodford, 155–81. Amsterdam: North Holland.

Slovic, Paul, Howard Kunreuther, and Gilbert F. White. 1974. "Decision Processes, Rationality, and Adjustments to Natural Hazards: A Review of Some Hypotheses." In *Natural Hazards: Local, National and Global*, edited by Gilbert White, 187–205. New York: Oxford University Press.

Snyder, James M., and David Strömberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy* 118 (2): 355–408.

Swed, Frieda S., and C. Eisenhart. 1943. "Tables for Testing Randomness of Grouping in a Sequence of Alternatives." *Annals of Mathematical Statistics* 14 (1): 66–87.

Sylves, Richard, and Zoltán I. Búzás. 2007. "Presidential Disaster Declaration Decisions, 1953–2003: What Influences Odds of Approval." *State and Local Government Review* 39 (1): 3–15.

Tversky, Amos, and Daniel Kahneman. 1982. "Judgment under uncertainty: Heuristics and biases." In *Judgment under uncertainty: Heuristics and biases*, edited by Daniel Kahneman, Paul Slovic, and Amos Tversky, 3–22. Cambridge, UK: Cambridge University Press.

U.S. Government Accountability Office. 2008. *Flood Insurance: FEMA's Rate-Setting Process Warrants Attention*. U.S. Government Accountability Office (GAO). October.

Viscusi, W. Kip. 1991. "Economic Theories of Decision Making Under Uncertainty: Implications for Policy Analysis." In *Policy Analysis and Economics: Developments, Tensions, Prospects*, edited by David L. Weimer, 85–109. Boston: Kluwer Academic Publishers.