When the Storm Breaks (Expectations): Reference Dependence and Demand for Flood Insurance

Sierra Smith

October 27, 2025

Abstract

Households in hurricane-prone areas face rare but potentially devastating losses, yet flood insurance coverage remains low and unstable. This paper links hurricane forecasts to flood insurance uptake, introducing forecast errors as novel, exogenous reference points for household expectations. Using administrative records of NFIP policies in Florida merged with storm forecasts, I show that deviations between predicted and realized storm outcomes generate strong behavioral asymmetries: unanticipated impacts ("false misses") drive large increases in demand, often exceeding the response to correctly predicted strikes ("true hits"), while predicted threats that fail to materialize ("false hits") reduce uptake, suggesting that false alarms erode perceived risk. These patterns cannot be explained by actuarial risk alone and are consistent with reference-dependent preferences and loss aversion, in which unexpected losses loom larger than avoided ones. Salience further amplifies these dynamics, as recent storms and hurricane-classified systems elicit the strongest responses, while near-miss "close calls" often depress demand. Together, the findings demonstrate that household insurance decisions are shaped as much by the psychological impact of forecast errors as by objective risk, with implications for forecast communication, policy timing, and disaster preparedness.

1 Introduction

Flood insurance take-up in the United States remains persistently low, even in regions most vulnerable to storm surge and flooding, such as coastal Florida. A large share of households remain uninsured, and coverage rates often decline as memories of past disasters fade. This pattern raises a central puzzle: why does demand remain limited and unstable, despite insurance being a primary tool for mitigating catastrophic risk?

In this paper, I argue that household insurance behavior depends not only on objective exposure but also on subjective expectations shaped by hurricane forecasts. Forecasts provide salient, widely publicized signals about risk. When the realized storm path diverges from those forecasts, the discrepancy itself becomes a behavioral reference point. Such expectation violations generate asymmetric responses consistent with loss aversion: households surprised by unexpected impacts increase coverage sharply, while those exposed to false alarms often reduce coverage. In this way, the accuracy of public forecasts plays a central role in shaping protective decisions.

The case of Hurricane Ian in 2022 illustrates this mechanism. Figure 1 shows NOAA's predicted track as of September 26 compared with Ian's realized trajectory. Residents along Florida's Gulf coast were accurately warned and experienced damage in line with expectations, a "true hit." By contrast, households on the Atlantic coast were initially told they would be spared, only to face major flooding as the storm crossed the peninsula, a "false miss." Both groups faced losses, but the surprise element created stark differences in their perception of risk.

These dynamics build on behavioral theories of decision-making under uncertainty. Prospect theory emphasizes that individuals evaluate outcomes relative to reference points, and losses loom larger than equivalent gains (Kahneman and Tversky, 1979). Expectation-based models formalize how anticipated outcomes generate utility shocks when violated (Kőszegi and Rabin, 2006, 2007, 2009). Salience theory further suggests that vivid, unexpected outcomes capture disproportionate attention and influence subsequent choices (Bordalo et al., 2012). Forecasts thus serve a dual role: they not only inform households about risk but also anchor the reference points that structure post-disaster behavior.

Previous research has documented irregularities in insurance demand: coverage increases after major events and decays over time (Kunreuther, 1978; Camerer and Kunreuther, 1989); subsidized products often suffer from low acceptance in developing countries (Giné et al., 2008; Cole et al., 2013); while enrollment can be pushed upward through direct experience or behavioral interventions (Karlan et al., 2014; Cai and Song, 2017). However, most studies proxy experience or salience using coarse variables such as elapsed time since the last disaster or aggregate damage in a region. This paper introduces hurricane forecast errors that are quantifiable, exogenous, and

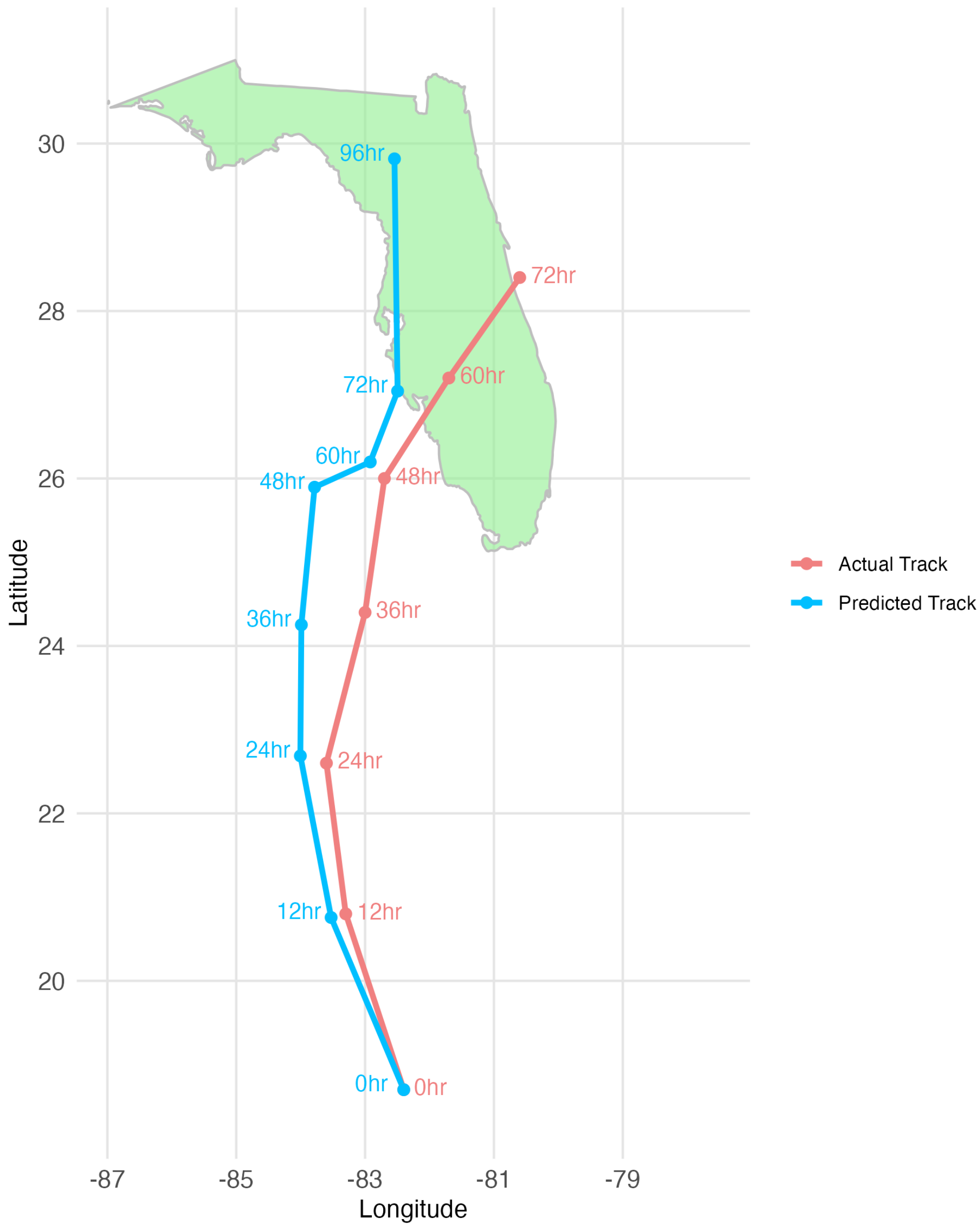


Figure 1: NOAA forecast track for Hurricane Ian (12 PM, September 26, 2022) compared with the observed track.

publicly distributed as a novel and powerful measure of expectation formation.

To examine these dynamics, I merge administrative data from the National Flood Insurance Program (NFIP) with storm forecast records from the National Oceanic and Atmospheric Administration (NOAA). I study how the deviations between the predicted and realized storm paths affect the issuance of new policies. The analysis shows three robust patterns: (i) unanticipated impacts (false misses) produce the largest increases in demand; (ii) expected impacts (true hits) also increase demand, though less strongly; and (iii) false alarms (false hits) reduce demand. Together, these findings suggest that insurance uptake is systematically mediated by the accuracy of forecasts, not only by the severity of the storm itself.

This paper makes three contributions. First, it advances the literature on disaster insurance by identifying forecast accuracy as a behavioral friction in coverage decisions. Second, it extends reference-dependent preference models by demonstrating that measurable forecast errors can serve as empirically tractable reference points. Third, it connects theories of salience to a real-world, high-stakes context, showing how widely available public information shapes private protective investments.

The policy implications are twofold. For insurers and regulators, understanding how households respond to forecast errors can improve the design of programs such as the NFIP, particularly in anticipation of climate change. For forecasting agencies, the results highlight the behavioral consequences of forecast communication: even well-intentioned false alarms may erode demand for protection. Beyond flood insurance, forecast-based reference points can matter for agriculture, health, and other settings where ex ante risk expectations shape coverage decisions.

The remainder of the paper proceeds as follows. Section 2 reviews related literature. Section 3 provides background information on NFIP and hurricanes. Section 4 describes the data. Section 5 presents the behavioral framework. Section 6 outlines the theoretical model. Section 7 details the empirical strategy. Section 8 reports results. Section 9 discusses extensions. Section 10 concludes.

2 Related Literature

Classical models of insurance demand predict that risk-averse individuals will purchase actuarially fair coverage to smooth consumption across states of the world (Arrow, 1963; Mossin, 1968; Arrow, 1971). However, real-world behavior departs from this benchmark. Browne et al. (2000); Kunreuther and Pauly (2004) document under-insurance in catastrophe contexts. While works by Kunreuther (1978, 1996) show that disaster insurance markets are characterized by low participation and pronounced cycles in coverage. These anomalies highlight that static expected-utility models alone cannot explain observed market outcomes.

These deviations have been attributed to range of frictions. Households often face liquidity and credit constraints (Grace et al., 2004; Gollier, 2005), limited understanding of contract terms (Schlesinger, 2000; Zweifel and Eisen, 2012), and distrust in insurers (Cole et al., 2013). Even when products are heavily subsidized, take-up remains low (Giné et al., 2008; Cole et al., 2013). Behavioral interventions such as reminders (Karlan et al., 2014), framing devices (Johnson et al., 1993), and default enrollment mechanisms (Ericson and Starc, 2012; Handel, 2013; Robinson et al., 2021) have been shown to raise coverage, emphasizing the role of psychological and institutional constraints.

Another important strand considers the role of experience. Coverage demand increases in the immediate aftermath of disasters and decays as memories fade (Camerer and Kunreuther, 1989). Deductible choices also exhibit inertia (Barseghyan et al., 2011), and direct experience of a flood increases subsequent insurance purchases (Gallagher, 2014). Collectively, these findings suggest that household behavior reflects dynamic processes of learning and forgetting, not just static risk preferences.

Behavioral models provide a unifying framework for these irregularities. Prospect theory posits that individuals evaluate outcomes relative to reference points, with losses weighted more heavily than equivalent gains (Kahneman and Tversky, 1979). Expectation-based models extend this insight, showing how anticipated outcomes shape preferences and how deviations from expectations generate utility shocks (Kőszegi and Rabin, 2006, 2007, 2009). Empirical work supports these mechanisms: households overinsure modest risks in ways consistent with probability weighting (Sydnor, 2010); distorted probabilities and loss aversion influence decisions in insurance and gambling contexts (Wakker and Deneffe, 1996; Post et al., 2008); and inertia in deductible choices appears even in competitive health insurance markets (Abaluck and Gruber, 2011; Handel, 2013). In experimental settings, reference dependence shapes effort provision (Abeler et al., 2011) and even sporting behavior (Pope and Schweitzer, 2011).

A closely related line of research exploits forecast errors as exogenous shocks to reference points. Card and Dahl (2011) use unexpected NFL outcomes to show that deviations from expectations drive emotional and behavioral responses. Forecast surprises in weather affect automobile purchases (Busse et al., 2015), and forecast errors in financial markets alter stock returns (Allen et al., 2017). These studies demonstrate that when expectations are measurable and forecasts are credible, deviations provide powerful natural experiments for studying reference dependence.

Salience-based models offer a complementary mechanism. Bordalo et al. (2012, 2013) argue that individuals overweight vivid or attention-grabbing attributes relative to baseline probabilities. Experimental evidence confirms that attention distortions shape risk-taking (Frydman and Mormann, 2016), and salient information has been shown to affect domains ranging from energy consumption (Allcott and Taubinsky, 2015) to protective investments.

Disaster contexts provide a natural testing ground for these theories. Households routinely under-insure against rare but catastrophic risks (Kunreuther and Pauly, 2004), and demand responds strongly to recent losses (Browne and Hoyt, 2000; Michel-Kerjan and Kousky, 2010; Gallagher, 2014; Hallstrom and Smith, 2005). In developing countries, the demand for rainfall insurance remains low even under large subsidies (Giné et al., 2008), although the experience with disasters increases adoption (Cai and Song, 2017). Recent studies explore expectation formation more explicitly: Bin and Landry (2013) show that willingness to pay adjusts with perceived flood risk; Kousky and Kunreuther (2014) emphasize myopia and misperceived probabilities; and Wagner (2022) develops a dynamic model that links expectation formation with adverse selection.

This paper contributes to these literatures by introducing hurricane forecast errors as a novel, measurable reference point. Forecasts are salient, credible, and widely disseminated, making them natural anchors for household expectations. Deviations from these forecasts create expectation shocks that are both psychologically meaningful and exogenous to individual households. By combining forecast data with detailed administrative records of flood insurance policies, this study provides direct evidence that forecast errors shape coverage decisions. In doing so, it advances the insurance demand literature, enriches models of reference-dependent preferences, and deepens our understanding of how salience and expectations influence protective investment in high-stakes, real-world markets.

3 Background

3.1 The National Flood Insurance Program

The National Flood Insurance Program (NFIP) was established by the U.S. Congress in 1968 through the National Flood Insurance Act (P.L. 90-448) in response to growing concerns over the widespread lack of private market coverage for flood risk. At the time, repeated flood disasters had imposed heavy financial burdens on both affected households and the federal government, which frequently resorted to ad hoc disaster relief. The NFIP was designed to fill this gap by offering federally supported flood insurance to residents in flood-prone areas, thus shifting post-disaster assistance from reactive aid to proactive risk pooling (Michel-Kerjan and Kousky, 2010; Kousky, 2018).

In 2022, more than 22,000 communities in all 50 states and territories had enrolled in the NFIP, supporting nearly 5 million active policies, representing roughly $1.3 trillion in total coverage (Congressional Research Service, 2023; Federal Emergency Management Agency (FEMA), 2023a). Policies are administered through the Write-Your-Own program, in which private insurers sell and service policies on behalf of the federal government while FEMA underwrites the risk and sets standardized premium rates. These premiums are calculated using FEMA's flood risk maps (known as Flood Insurance Rate Maps, or FIRMs), which incorporate factors such as property elevation, location within or outside SFHAs, building age and type of structure, and more recently actuarial variables under the new Risk Rating 2.0 pricing methodology (Federal Emergency Management Agency (FEMA), 2023b; Kousky and Kunreuther, 2014).

Notably, NFIP policies must be paid in full for the entire year in advance and coverage typically does not begin until 30 days after purchase. This rule was explicitly designed to discourage last-minute purchases in anticipation of storms (Federal Emergency Management Agency (FEMA), 2023a). The delay introduces important temporal frictions into insurance decision-making: individuals must assess and act on risk in advance, often months before hurricane season peaks. In behavioral terms, this creates room for salience, risk perception, and recent weather experiences to disproportionately shape demand.

Although policyholders may cancel at any time, premiums are generally non-refundable, which further reduces the attractiveness of speculative or short-term enrollment. This reinforces the structure of the NFIP as a commitment device, which requires ex ante recognition of risk and sustained participation. These features contrast sharply with many forms of post-disaster aid or private short-term insurance markets.

3.2 Hurricanes

3.2.1 Hurricanes and Storm Classifications

Tropical cyclones, broadly referred to as storms in this paper, are organized atmospheric systems that originate over warm tropical or subtropical ocean waters. These systems develop when warm, moist air rises from the ocean surface, generating convection and releasing latent heat that fuels further intensification. A system is initially designated a tropical disturbance when it exhibits sustained thunderstorm activity without well-defined circulation. If convection becomes organized and a closed low-level center forms, the system is upgraded to a tropical cyclone (NOAA National Hurricane Center, 2019; Landsea and Franklin, 2013).

Storms are classified according to their maximum sustained one-minute surface wind speeds using the Saffir-Simpson Hurricane Wind Scale (SSHWS). A system with wind speeds of 38 mph or less is classified as a tropical depression. Once wind speeds reach between 39 and 73 mph, the storm becomes a tropical storm and receives an official name. When sustained winds exceed 74 mph, the system is classified as a hurricane. Hurricanes are further subdivided into five categories: Category 1 (74–95 mph), Category 2 (96–110 mph), Category 3 (111–129 mph), Category 4 (130–156 mph) and Category 5 (157 mph or greater). The National Hurricane Center (NHC) designates storms in Categories 1 and 2 as minor hurricanes, while storms in Categories 3 through 5 are considered major hurricanes, due to their increased potential for destruction (NOAA National Hurricane Center, 2020a).

Although hurricanes of any intensity can cause substantial damage through wind impact, storm surge, and inland flooding, major hurricanes account for a disproportionate share of economic losses and fatalities. These distinctions play an important role in the empirical analyses that follow, where I examine how variation in storm intensity shapes insurance behavior and policy uptake.

Geographically, storms form in both the Atlantic and eastern Pacific basins, but only a subset of these systems pose direct threats to the United States, while the eastern Pacific sees a higher total number of tropical cyclones, prevailing wind patterns and cooler ocean temperatures often inhibit landfall. By contrast, the Atlantic basin, especially the Gulf of Mexico and the Caribbean Sea, provides ideal conditions for storm formation and intensification, with warmer waters and steering currents that frequently direct systems toward the U.S. mainland (Emanuel, 2005).

To delineate temporal patterns of risk, the National Hurricane Center defines the Atlantic hurricane season as extending from June 1 to November 30, a window that captures nearly all historical U.S. landfall events. On average, a typical season produces approximately 14 named storms, of which 7 become hurricanes and 3 escalate to major hurricanes (NOAA National Hurricane Center, 2020b). These storms generate a wide array of hazards, including high winds, torrential rainfall, coastal storm surge, and inland flooding, often affecting areas far beyond the immediate coastline.

3.2.2 Forecasting and Risk Communication

Once a tropical cyclone forms, the NHC initiates a continuous cycle of forecast updates, typically issued at six-hour intervals (at 00, 06, 12, and 18 UTC) and continuing until the storm dissipates or merges with another system. Each advisory includes predictions of the storm's future trajectory, intensity, and spatial extent, and serves as a critical input to emergency management, media coverage, and perception of risk at the home level (NOAA National Hurricane Center, 2022b).

These forecasts are generated using an array of numerical weather prediction models that solve complex physical equations that govern atmospheric motion. The NHC synthesizes outputs from various global and regional models, each differing in spatial resolution, physical parameterizations, and initial condition schemes. By integrating multiple model outputs, human forecasters develop a consensus track and intensity forecast for public release (Cangialosi et al., 2020; Tallapragada et al., 2014).

Each forecast includes projected storm center coordinates and maximum sustained wind speeds at regular intervals: every 12 hours from 0 to 72 hours ahead and every 24 hours from 72 to 168 hours. This structure yields up to 11 discrete time-point predictions per forecast cycle, facilitating time-sensitive decisions by households, businesses, and government agencies (NOAA National Hurricane Center, 2022a).

Central to public risk communication is the "cone of uncertainty", a graphical representation of the probable path of the center of the storm based on historical forecast errors over the past five years. The cone reflects only the uncertainty in the forecast track. It does not indicate the size of the storm or the range of potential hazards. The cone widens with forecast horizon, visually conveying the increasing uncertainty associated with longer lead times. Despite its technical limitations, the cone has become a widely recognized visual tool in media and public discourse, and plays a powerful role in shaping perceived risk and behavioral responses (Broad et al., 2007; Morss et al., 2010). Figure 2 provides an example forecast for Hurricane Ian created by NOAA which was subsequently distributed to media outlets National Hurricane Center (2022).

Forecast accuracy varies substantially with horizon. While 24- and 48-hour track forecasts are generally reliable, forecast errors, especially for intensity, increase sharply beyond 72 hours. This growing uncertainty at longer lead times can erode confidence in forecasts and can delay protective actions by individuals who perceive information as ambiguous or unreliable. As such, forecast credibility, timing, and clarity are essential not just for scientific accuracy but also for effective behavioral influence and insurance-related decision-making.

4 Data Sources and Sample Construction

4.1 NFIP Policies

The empirical analysis draws on administrative records from the NFIP spanning 2009 to 2023. These data, maintained by FEMA, provide comprehensive information on all flood insurance policies nationwide, including newly issued and renewed contracts. Each policy record contains detailed property characteristics (e.g., structure type, geographic coordinates, estimated replacement cost) as well as policy-level attributes (e.g., coverage limits, deductibles, premiums, and effective

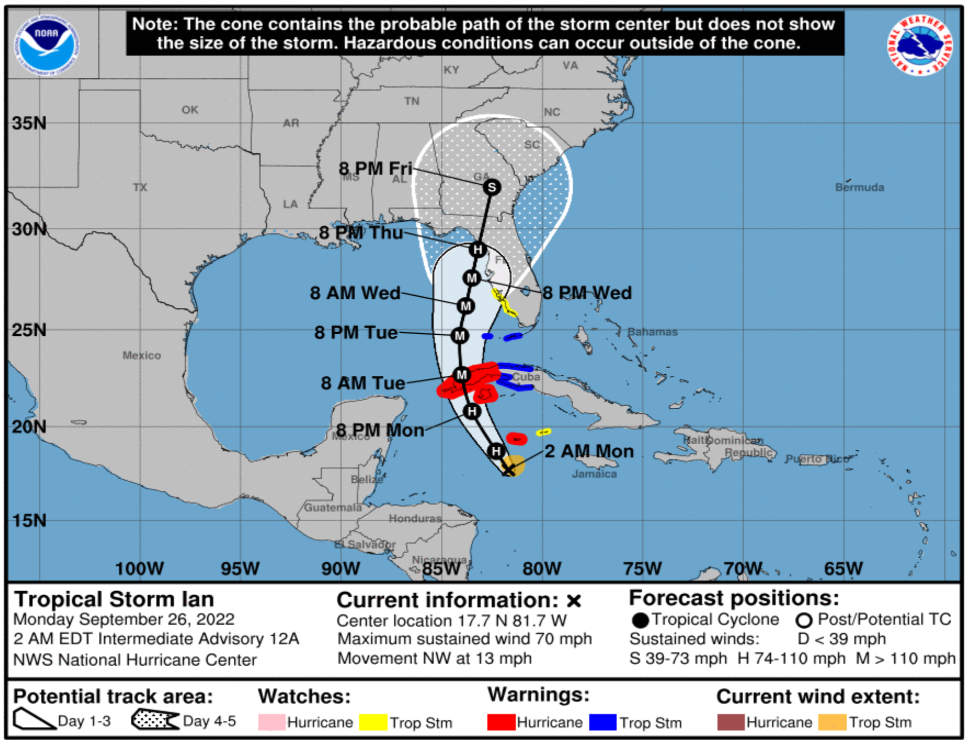


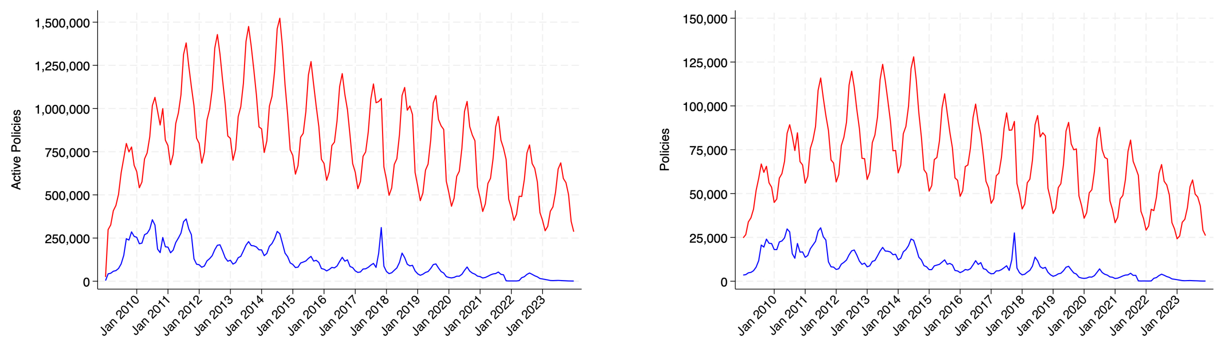
Figure 2: Example of a "cone of uncertainty" forecast graphic distributed by the National Hurricane Center.

dates).

For the purposes of this study, I focus on a more homogeneous risk and policy environment. The sample is restricted to Florida, which consistently ranks among the most flood-prone states in the U.S. and has a high concentration of NFIP policies. This geographic focus allows for a consistent regulatory context, shared floodplain management practices, and a direct link to Atlantic hurricane exposure. I further restrict the analysis to residential single-family houses, excluding commercial, condominium, and multifamily structures, in order to study household-level decisions. Finally, to isolate first-time purchase behavior, I retain only newly issued policies and exclude renewals. This ensures that the analysis captures responses to evolving risk perceptions, particularly those shaped by forecasts, rather than inertia in ongoing coverage.

The resulting data set consists of more than 1.66 million newly issued residential flood insurance policies in Florida over a fifteen-year period. Figure 3 presents monthly policy activity between 2009 and 2023. Panel A plots the total number of active policies (red line), including renewals and new issuance, while Panel B isolates effective-date issuance, filtering out mid-cycle adjustments. Both panels reveal pronounced seasonality: policy activity rises sharply in late summer, coinciding with the Atlantic hurricane season, and falls during the winter months. This cyclical pattern is consistent with the idea that salience and perceived storm risk strongly influence purchase decisions.

Beyond seasonality, the data reveal longer-run trends. The total stock of active policies increased steadily until peaking around 2014–2015, followed by a gradual decline in subsequent years. This decline may reflect affordability concerns, growth in private-market alternatives, or shifting perceptions of flood risk. Newly issued policies, shown in blue, are smaller in magnitude but exhibit sharp spikes around major hurricanes, temporary surges in demand that rapidly decay. Such dynamics align with behavioral models emphasizing recency and salience effects.



(a) Total Active Policies

(b) Effective Date Policies

Figure 3: Total stock of active NFIP residential policies (Panel A) and newly issued effective-date policies (Panel B)

Table 1 reports summary statistics for the final sample. The typical policy insures a single-story primary residence with an average home age of 26 years, though the distribution spans newly constructed homes to properties over two centuries old. Roughly 83% of insured properties are primary residences, while only 6% are elevated structures. Average coverage amounts are \$205,000 for buildings and \$74,000 for contents, but these values vary widely, reflecting the heterogeneity of Florida's housing stock. Premiums average \$493 annually, with total policy costs (including surcharges and fees) averaging \$565, although some high-value homes pay substantially more. Deductibles remain modest, averaging around \$1,450 for buildings and \$1,260 for contents.

These descriptive patterns confirm both the heterogeneity of insured properties and the volatility of insurance demand, providing a foundation for the empirical analysis of how forecast errors shape new policy purchases in subsequent sections.

Table 1: Summary Statistics for NFIP Policy Variables

Variable	Obs.	Mean	Std. Dev.	Min	Max
Building Deductible	1,655,539	1,449	1,202	500	50,000
Contents Deductible	1,546,265	1,262	929	500	25,000
Total Building Coverage	1,648,422	204,797	66,053	100	5,750,000
Total Contents Coverage	1,485,254	74,249	32,176	100	500,000
Total Premium	1,664,211	493	519	0	56,309
Total Policy Cost	1,664,208	565	557	0	56,349
Primary Residence (1 = Yes)	1,664,216	0.83	0.37	0	1
Elevated Building (1 = Yes)	1,664,216	0.06	0.24	0	1
Floors	1,664,207	1.35	0.70	1	6
Home Age (Years)	1,663,846	25.96	18.21	0	268
Total Observations	1,664,216				

4.2 Storm Data

To capture household exposure to storm activity, I construct a storm-level panel using forecast and track data from the National Hurricane Center's (NHC) hurricane database. This dataset includes all tropical cyclones in the North Atlantic basin between 2008 and 2023 and provides the foundation for measuring both realized storm characteristics and the advance warnings available to households prior to landfall.

Each storm is observed across multiple forecast cycles, typically issued every six hours (00, 06, 12, and 18 UTC). At each issuance, the database reports both realized storm positions and intensities as well as projected tracks, wind speeds, and storm radii at different lead times. Forecasts are structured as forward-looking predictions of storm location and strength at standard horizons—0, 12, 24, 36, 48, 72, 96, and 120 hours ahead—allowing households to form risk expectations at varying levels of advance notice.

For each forecast cycle, the data include the geographic position (latitude and longitude) of both predicted and observed storm centers, maximum sustained wind speed (mapped to hurricane categories using the Saffir–Simpson scale), and storm size estimates based on the radial extent of winds at 34-, 50-, and 64-knot thresholds in all four quadrants. These detailed records make it possible to compare the evolution of forecasted storm paths with subsequent realized outcomes, and thus to quantify the degree of forecast error faced by households in real time.

Table 2 summarizes annual storm counts by type and predicted intensity between 2008 and 2023, classifying each storm according to its highest predicted category during its life cycle. Tropical depressions and hurricanes are the most frequently forecast storm types. The number of predicted hurricanes varies considerably, ranging from only two in 2013 to thirteen in 2020, a year of historically high activity. Category 1 hurricanes are the most common forecast outcome, while Category 4 and 5 storms remain extremely rare. Several seasons (e.g., 2012 and 2013) featured no storms predicted above Category 2, highlighting the episodic nature of severe hurricane threats. Overall, the total number of predicted storms fluctuates substantially across years, from as few as nine in 2014 to as many as 31 in 2020, reflecting both climatological cycles and year-to-year variation in forecast severity. Although less common, subtropical systems also appear in the dataset, highlighting the diversity of storm types communicated to the public and their potential role in shaping household perceptions of risk.

Table 2: Annual Counts of Storm Types at Prediction, 2008–2023

Category	**Type**	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
Storm Type	Subtropical Depression	0	0	0	0	0	0	0	0	0	1	2	0	2	1	0	0
	Subtropical Storm	0	0	0	0	0	1	0	0	0	0	0	2	0	0	0	1
	Tropical Depression	8	6	8	11	9	9	3	7	7	6	5	8	14	10	6	10
	Tropical Storm	1	2	1	1	0	3	0	1	1	1	1	4	2	2	1	3
	Hurricane	8	3	12	7	10	2	6	4	7	10	8	6	13	7	9	7
Hurricanes	Category 1	3	1	7	3	8	2	4	2	3	4	6	3	7	3	6	4
	Category 2	1	1	1	2	2	0	1	1	2	2	0	1	1	2	1	1
	Category 3	4	1	3	2	0	0	1	0	1	1	0	0	2	0	1	0
	Category 4	0	0	1	0	0	0	0	1	1	3	2	1	3	2	1	2
	Category 5	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
All Storms		17	11	21	19	19	15	9	12	15	18	16	20	31	20	16	21

4.3 Linking Insurance and Storm Forecast Data

The NFIP policy data are merged with the NHC storm forecast database along both temporal and spatial dimensions to construct a panel that links household insurance decisions to storm expectations. Newly issued residential NFIP policies are first aggregated to the month–location level, recording the number of policies by effective date and geographic coordinates. To ensure complete coverage of potential insurance activity, I generate the full set of possible month–year and location combinations for Florida during the study period, yielding 1,497 unique geographic points observed across 180 months. Location–month instances with no new policies are coded as zeros, producing a balanced panel that captures both the presence and absence of enrollment activity.

This panel is then linked to storm forecasts within a forward-looking window. Specifically, each location–month observation is matched to all storms whose initial forecast issuance occurred between 31 and 120 days prior to the policy's effective date. This three-month horizon reflects the period during which households are most likely to form expectations and make enrollment decisions in response to forecast information. For every storm–location pair in this window, I calculate the great-circle distance between the predicted storm center and the policy location's coordinates. The predicted center is taken from the 72-hour lead forecast in the baseline specification, which provides a meaningful balance between forecast availability and accuracy.

When multiple forecast cycles exist for a given storm, the cycle in which the predicted center lies closest to the location is retained. To further capture the salience of the most relevant storm information, storms within the exposure window are ranked by their temporal proximity to the policy effective month, with the closest storm forecast treated as the primary exposure. This approach allows insurance uptake to be studied not only as a function of realized storm impacts, but also of forecasted threats communicated in advance to households.

4.4 Expected Outcomes

Figure 4 illustrates the relationship between predicted and realized storm distances at the 72-hour forecast horizon, measured in nautical miles from the Florida panhandle. Each point corresponds to a unique storm that was predicted to pass within 600 nautical miles of the region. The considerable vertical dispersion around the 45-degree line reflects forecast error: many storms deviate substantially from their predicted paths, with some missing entirely and others striking unexpectedly despite initial projections. This variation provides the empirical foundation for studying how expectation violations influence household insurance behavior.

To operationalize exposure, I classify storms using a binary "Hit" indicator that captures whether the predicted storm center fell within a specified distance of the policy location. The baseline threshold is set at 300 nautical miles, which approximates the average radial extent of tropical-storm-force winds in Atlantic hurricanes (typically ranging from 200–400 nautical miles; Kimball and Mulekar, 2004). This threshold balances sensitivity to meaningful forecast exposure with the need to exclude distant storms unlikely to influence household perceptions. The same classification is applied to realized storm tracks to measure actual exposure. Together, these indicators generate consistent measures of both predicted and realized storm impact.

Figure 5 plots the empirical probability of realized storm exposure against predicted distance at the 72-hour horizon. Vertical lines mark the 300nm threshold used to classify predicted hits (to the left) and predicted misses (to the right). The curve declines sharply as predicted distance increases: close forecasts are associated with higher realized probabilities of impact, while distant forecasts are rarely realized as strikes. Yet even within the predicted hit zone, forecast precision remains limited. When a storm is projected to pass within 300nm, the realized probability of actual impact is only about 42%, implying that the majority of predicted hits do not materialize. Conversely, when storms are forecast to miss by more than 300nm, the realized probability of an actual strike

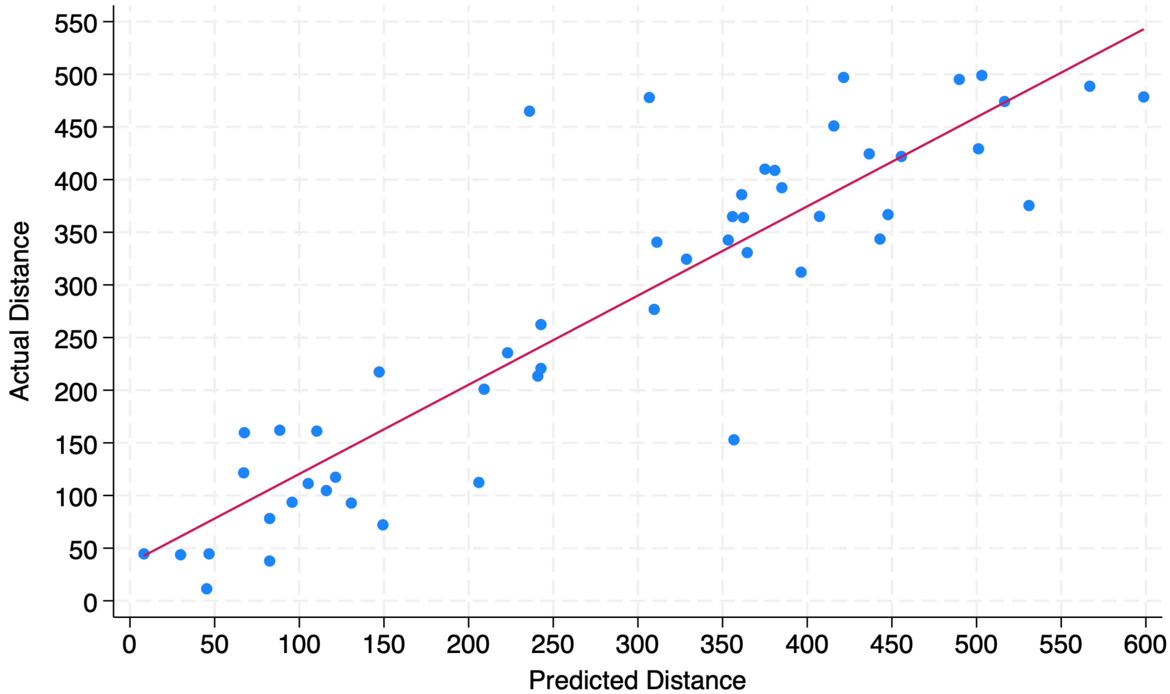


Figure 4: Scatterplot of predicted and realized storm distances from the Florida panhandle at the 72-hour horizon.

falls close to zero, but not entirely, some storms still deviate substantially from their projected paths.

This pattern highlights two important features of the forecast environment. First, forecasts are systematically informative: predicted distance strongly correlates with realized exposure, confirming that households receive meaningful signals about risk. Second, forecasts are also noisy: the wide gap between predicted hits and realized strikes generates frequent expectation violations.

Combining predicted and realized indicators yields a four-category taxonomy of forecast accuracy. A "True Hit" occurs when both predicted and realized distances fall within 300nm; a "True Miss" when both fall outside. A "False Hit" reflects storms predicted to strike that ultimately missed, while a "False Miss" captures storms predicted to spare a location that ultimately hit. In addition, location–month observations with no relevant storm activity during the 31–120 day exposure window are assigned a separate "No-Storm" indicator. This classification framework provides the key empirical link between prediction accuracy and subsequent insurance demand.

Table 3 reports the distribution of storm–location observations across categories. True hits and true misses together account for more than 60% of cases, with tropical storms and hurricanes comprising the majority within these groups. False hits and false misses are relatively rare—together

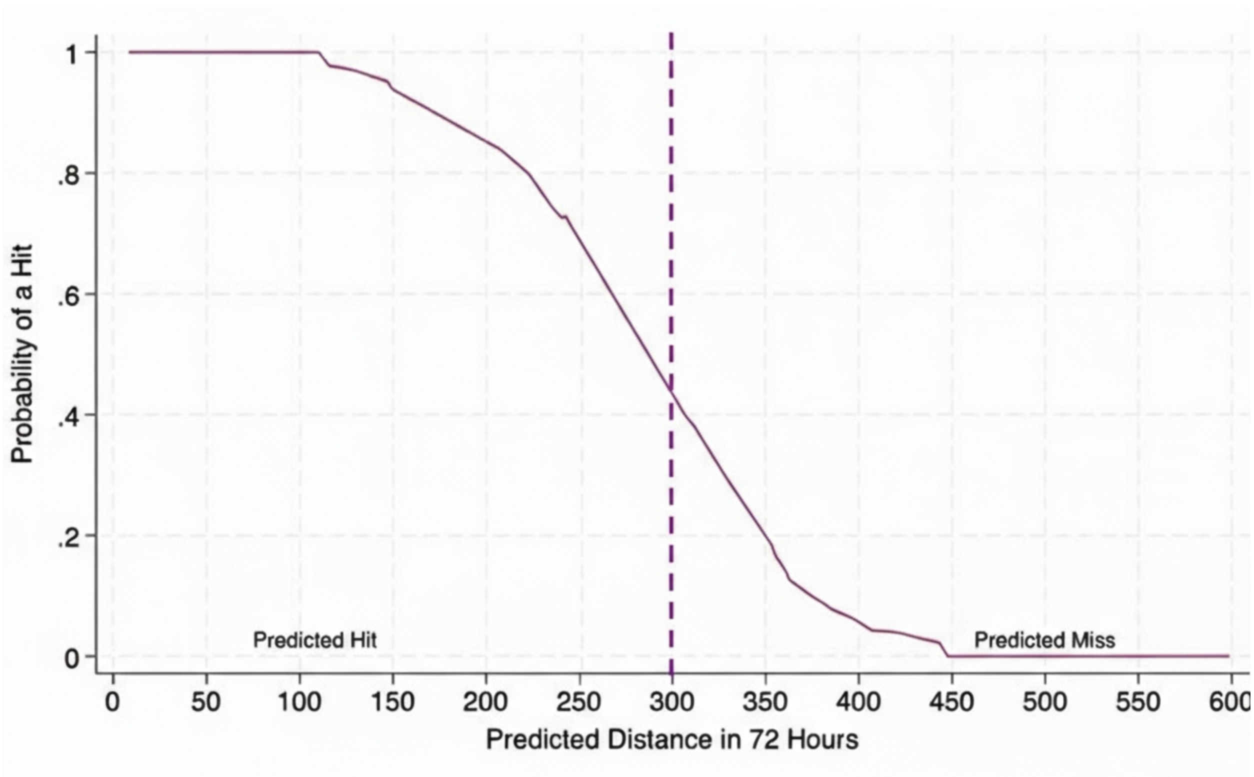


Figure 5: Empirical probability of realized storm impact as a function of predicted distance at the 72-hour horizon.

Table 3: Distribution of Forecast Accuracy Categories

Category	Frequency	Percent
True Hit	59,272	22.0%
Subtropical Storm	171	
Tropical Depression	7,347	
Tropical Storm	26,716	
Hurricane	25,038	
True Miss	113,778	42.2%
Subtropical Storm	86	
Tropical Depression	23,448	
Tropical Storm	43,619	
Hurricane	46,625	
False Hit	1,948	0.7%
Subtropical Storm	35	
Tropical Depression	862	
Tropical Storm	106	
Hurricane	945	
False Miss	4,850	1.8%
Subtropical Storm	0	
Tropical Depression	1,597	
Tropical Storm	3,223	
Hurricane	30	
No Storm	89,612	33.3%
Observations	269,460	

representing less than three percent of the sample. Roughly one-third of location–month observations fall into the no-storm category, reflecting the frequency with which households faced no proximate forecast activity during the relevant decision window.

5 Behavioral Framework for Insurance Response

Insurance decisions in the face of storms offer a unique window into how individuals perceive and react to risk under uncertainty. Because storms are rare and highly variable in their impacts, households must often rely on imperfect forecasts to guide protective behavior, such as whether to purchase flood insurance. The behavioral response to such forecasts likely depends not only on the realized outcome of a storm, but also on the alignment (or misalignment) between what was predicted and what ultimately occurred.

In what follows, I outline four conceptual cases that describe different ways that individuals might interpret and respond to combinations of predicted and realized storm exposure. These cases draw on established theories of salience, availability bias, ambiguity aversion, and reference-dependent utility. Together, they form a framework for understanding how predictive signals and lived experiences interact to shape insurance demand.

5.1 Case 1: Experience-Driven Updating

The most straightforward intuition is that individuals respond primarily to the realized impact of a storm, rather than its predicted trajectory. Because storms are rare and uncertain, homeowners can initially discount the risk and delay buying insurance. However, when a storm makes landfall nearby, the event becomes highly salient, prompting people to update their beliefs and seek coverage in anticipation of future threats. If this model holds, I would expect insurance demand to rise similarly following both predicted hits and predicted misses, as long as the storm ultimately makes impact. Conversely, if the storm misses, individuals would perceive a lower threat and show little response.

This behavior is consistent with availability bias Tversky and Kahneman (1973), where recent or vivid events are more likely to influence decision-making than abstract probabilities. Empirical evidence shows that insurance uptake often spikes after high-impact events Gallagher (2014); Kousky (2018), consistent with reactive responses driven by experience. This may also reflect myopic risk assessment Kunreuther and Pauly (2004), in which individuals underweight low-probability future threats until the risk becomes pronounced through direct exposure. However, this case contrasts with models of reference dependence in which individuals respond not only to outcomes, but also to how those outcomes compare to expectations.

5.2 Case 2: Dual Sensitivity to Forecasts and Outcomes

A second possible interpretation is that individuals respond not only to outcomes, but also to predictions themselves. Consider the two consistent cases: a predicted hit that results in a realized hit and a predicted miss that results in a realized miss. In this framework, we would expect the former to have the strongest effect on insurance demand, as the prediction reinforces the outcome and highlights risk (Tversky and Kahneman, 1973; Bordalo et al., 2012). In contrast, the latter should have the weakest effect, as neither the prediction nor the outcome signals an elevated risk.

The remaining two cases, predicted hit but realized miss and predicted miss but realized hit, introduce conflicting information. In both cases, the homeowner becomes aware of the risk of a storm, either through a forecast warning or unexpected landfall. However, inconsistency between prediction and outcome can reduce trust in forecasts or increase ambiguity (Ellsberg, 1961; Kunreuther, 1996). Despite this, the salience of the event may still elevate perceived vulnerability. Consequently, insurance demand is expected to increase in these cases, albeit less than in the scenario where both the prediction and realization align to signal high risk.

5.3 Case 3: Asymmetric Salience of Surprise Events

A related interpretation builds upon Case 2, but places greater emphasis on the salience of realized outcomes. Individuals continue to respond to both predictions and actual storm outcomes, but realized hits (whether expected or not), heighten salience more than realized misses. As in Case 2, the implications for the consistent scenarios remain unchanged: a predicted hit that results in a realized hit is expected to generate the highest increase in demand, while a predicted miss that results in a realized miss elicits the least response.

However, this interpretation introduces an important distinction between the two ambiguous cases. Specifically, a predicted miss that results in a realized hit is more salient than a predicted hit that results in a realized miss, as the former involves an unexpected impact and may prompt the updating of beliefs about both the risk of the storm and the reliability of the forecasts (Bordalo et al., 2012). In contrast, a false alarm may be discounted as noise. As a result, insurance uptake is likely to be higher after a surprise hit than after a false alarm, even though both involve inconsistencies between prediction and outcome.

5.4 Case 4: Reference Dependence and Forecast-Based Expectations

A final interpretation draws on the reference-dependent utility framework developed by Kőszegi and Rabin (2006, 2007, 2009), which incorporates expectation-based reference points and loss aversion. In this model, individuals assess outcomes relative to what they expect to occur, with losses weighted more heavily than gains. In the context of storm forecasts, the prediction establishes the reference point: a predicted hit sets the expectation of loss, while a predicted miss sets the expectation of safety or gain.

The psychological response depends not only on the realized outcome, but also on whether that outcome deviates from the prior expectation. A realized hit that was not predicted constitutes an unexpected loss and should therefore trigger the strongest increase in insurance demand. A predicted and realized hit is an expected loss and should still prompt insurance uptake, but to a lesser degree. In contrast, both types of misses represent gains. A predicted and realized miss (an expected gain) is the least likely to spur insurance behavior. A realized miss following a predicted hit is an unexpected gain; it may prompt modest demand due to residual salience, but loss aversion implies a muted response compared to the surprise hit.

6 Modeling the Effect of Forecasts and Insurance Demand

This section develops a simplified model of the effect of storm outcomes on the decision to purchase flood insurance and establishes the empirical framework for identifying the impact of storm forecasts. The central hypotheses are derived from Case 4 in the motivation section: specifically, that storm forecasts shape expectations in a manner consistent with gain-loss utility evaluated relative to a rational, expectation-based reference point.

6.1 Insurance Demand Model

Consider a homeowner who, in each period, faces some risk of experiencing a damage-inducing storm. Let $d \geq 0$ denote the probability that the homeowner purchases flood insurance in a given period. This probability is influenced by the outcome of a contemporaneous storm, indicated by $y \in \{0, 1\}$, where $y = 1$ indicates that the storm has a negative impact on the homeowner (i.e., a "hit") and $y = 0$ indicates no impact (i.e., a "miss").

Let $p = E[y]$, denote the homeowner's prior belief about the likelihood of a damaging storm. The decision to purchase insurance is assumed to deviate from a baseline level d^0 based on the psychological impact of the storm outcome, captured by gain-loss utility. Specifically, I assume:

$$d = d^0 + \mu(y, p), \tag{1}$$

where $\mu > 0$ is a piece-wise linear function defined as:

$$\mu(y,p) = \begin{cases} \alpha(y-p), & \text{if } y > p \\ \beta(p-y), & \text{if } y < p \\ \gamma, & \text{if } y = p = 1 \\ 0, & \text{if } y = p = 0 \end{cases}$$

for positive constants α, β, and γ. The assumption $\alpha > \gamma > \beta$ captures behavioral asymmetries in response to storm realizations. Specifically, the marginal effect of an unexpected hit exceeds that of an unexpected miss, consistent with loss aversion. The parameter γ reflects a salience effect: when an anticipated storm materializes, it reinforces perceived risk and increases insurance demand, even in the absence of surprise. In contrast, the case $y = p = 0$ (an expected miss) has no psychological salience and therefore does not affect insurance behavior.

Since storm outcomes are binary, the model implies four distinct expressions for insurance demand as a function of the probability of a hit, p:

$$\begin{aligned} d^{UH}(p) &= d^0 + \alpha(1-p) && \text{(Unexpected hit/loss)} \\ d^{UM}(p) &= d^0 + \beta p && \text{(Unexpected miss/gain)} \\ d^{EH}(p) &= d^0 + \gamma && \text{(Expected hit/loss)} \\ d^{EM}(p) &= d^0 && \text{(Expected miss/gain)} \end{aligned} \tag{2}$$

Figure 6 illustrates these four cases. The upper, downward-sloping line corresponds to $d^{UH}(p)$. When $p = 0$, a storm hit is entirely unexpected, leading to the highest level of insurance demand at $d^0 + \alpha$. As p increases, expectations and outcomes gradually align and demand declines until it converges with the expected hit level $d^{EH} = d^0 + \gamma$ at $p = 1$. Thus, d^{UH} is decreasing in p. The upward-sloping line represents $d^{UM}(p)$. When $p = 0$, a storm miss is fully anticipated, so demand aligns with the expected miss level $d^{EM} = d^0$. As p rises, the miss becomes increasingly unexpected, reaching a maximum at $d^0 + \beta$, when $p = 1$. Unlike the other two cases, both $d^{EH}(p)$ and $d^{EM}(p)$ are constant with respect to p, as they reflect scenarios where outcomes match prior expectations.

6.2 Evaluating the Effect of Forecast Information

I evaluate the behavioral effects of storm forecasts on flood insurance uptake using a Poisson count model of new policy issuance in Florida. The Poisson framework is well-suited to the data structure, where the dependent variable is the number of new policies issued at the location–month level. Storms are classified ex ante based on NOAA's 72-hour forecasts: a storm is designated as a predicted hit if its forecast track passes within 300 nautical miles of a location, and as a predicted miss otherwise.

The empirical specification incorporates interaction terms between these forecast classifications and realized storm outcomes (hit or miss), with months that experience no relevant storm activity serving as the omitted baseline category. This structure provides a direct test of how forecast accuracy conditions insurance demand, capturing differences in uptake following correct predictions (true hits and true misses), unforeseen impacts (false misses), and overstated threats (false hits). By explicitly modeling these contrasts, the analysis isolates the behavioral responses generated by

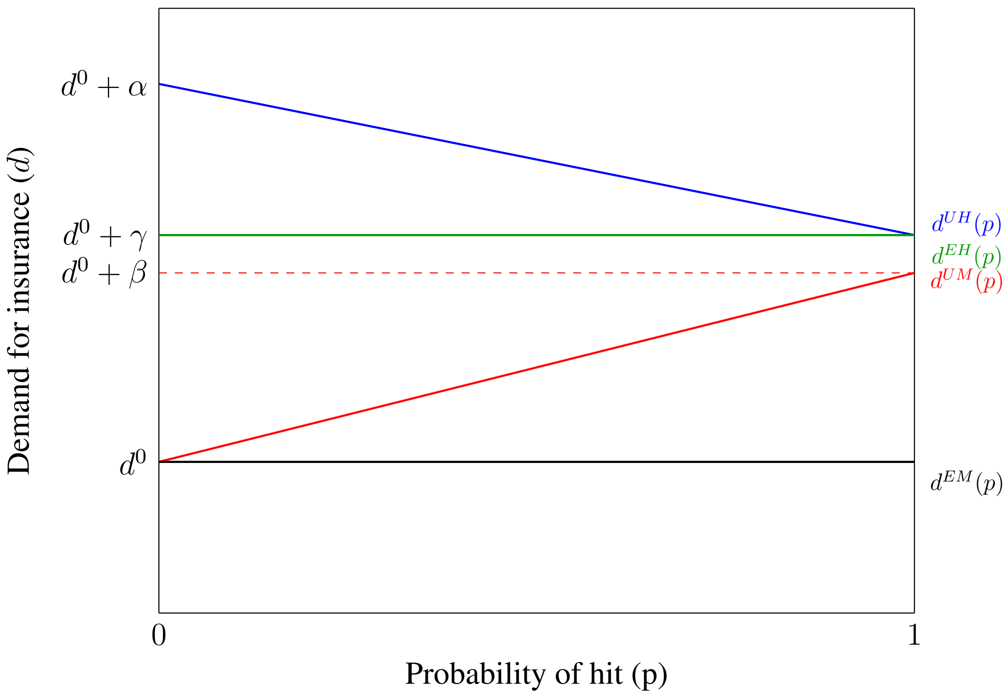


Figure 6: Stylized model of flood insurance demand under reference dependence

expectation violations relative to periods without storm exposure.

To assess robustness, I re-estimate the model under alternative definitions of exposure. First, I alter the distance threshold, testing cutoffs ranging from 100 to 500 nautical miles, to examine sensitivity to alternative definitions of a predicted hit. I also consider a "close-call" category, defined as storms predicted to pass between 250 and 350 nautical miles from a location, in order to capture borderline cases where exposure risk is ambiguous. Second, I vary the forecast horizon, considering predictions at 36, 48, 96, and 120 hours ahead, to evaluate whether shorter- or longer-range forecasts generate different behavioral effects.

In addition, I conduct robustness checks designed to probe storm salience. These include restricting the forecast window to storms occurring within the past 31–60 days to assess whether more recent events exert stronger influence on demand, and interacting the forecast categories with storm severity, distinguishing between systems formally classified as hurricanes and weaker storms. Together, these extensions evaluate whether temporal proximity and perceived intensity amplify the behavioral effects of forecast information, thereby testing the durability and scope of the baseline results.

7 Empirical Methodology

The unit of analysis is the location–month, defined as a specific geographic coordinate i observed in month–year t. The dependent variable is the number of newly issued NFIP residential policies in that location–month. Because the outcome is a non-negative count, I estimate a Poisson regression model, which is well-suited to the distributional properties of the data. The general specification is:

$$\log(Y_{it}) = \theta + X_{it}\gamma + f(p_{it}, a_{it}; \lambda) \tag{3}$$

where Y_{it} denotes the expected number of new policies in location i at time t; X_{it} is a vector of controls (e.g., month/year fixed effects, regional trends, and location-level covariates); and $f(p_{it}, a_{it}; \lambda)$ is a function that links insurance demand to both forecast information and realized storm outcomes. Here, p_{it} denotes the perceived probability of a storm impacting location i at time t, a_{it} is a binary indicator for whether a realized storm impact occurred, and λ is a behavioral parameter capturing how expectation violations condition household responses.

I assume that perceived strike probability p_{it} is a function of the predicted storm distance, D_{it}, which operationalizes forecast salience. Substituting into the model yields:

$$\log(Y_{it}) = \theta + X_{it}\gamma + g(D_{it}, a_{it}; \lambda) \tag{4}$$

This specification allows the effect of realized outcomes a_{it} to vary systematically with prior forecasts D_{it}. In other words, the impact of experiencing a storm is conditioned by whether households were led to expect one. The coefficients of interest capture the behavioral consequences of true hits, true misses, false hits, and false misses.

The identification strategy rests on two key assumptions. First, households view National Hurricane Center forecasts as informative and credible public signals of risk. Second, conditional on forecast proximity, the realized storm outcome represents exogenous variation from the household perspective. Under these assumptions, the framework isolates the behavioral effects of forecast accuracy by comparing insurance uptake across categories of forecast–outcome pairs. This design thus enables a direct test of whether expectation violations produce asymmetric shifts in demand for flood insurance.

8 Baseline Empirical Results

Table 4 presents estimates from five baseline Poisson regressions of new flood insurance issuance on storm exposure classifications. In these specifications, storms are defined by the interaction between predicted and realized proximity. A storm is classified as a predicted hit if NOAA's 72-hour forecast placed the storm center within 300 nautical miles of a location, and as a realized hit if the storm ultimately came within the same threshold. The storm exposure function is formalized as:

$$
\begin{aligned}
g(D_{it}, a_{it}, \lambda) = \lambda_1 \cdot 1(D_{it} \leq 300)(a_{it} = 1) \\
+ \lambda_2 \cdot 1(D_{it} \leq 300)(a_{it} = 0) \\
+ \lambda_3 \cdot 1(D_{it} > 300)(a_{it} = 0) \\
+ \lambda_4 \cdot 1(D_{it} > 300)(a_{it} = 1)
\end{aligned}
\tag{5}
$$

This specification yields four mutually exclusive categories: a true hit (λ_1), a false hit (λ_2), a true miss (λ_3), and a false miss (λ_4). In addition, a fifth category, no storm, is defined as location–months with no storms forecast or realized within the 31–120 day exposure window. This "no storm" group serves as the omitted reference category, so all coefficients are interpreted relative to periods without storm activity.

Column (1) includes location, year, and month fixed effects to account for spatial heterogeneity and temporal seasonality. Column (2) adds random effects to capture unobserved variation across units. Columns (3) through (5) sequentially introduce additional controls: property and policy characteristics (e.g., coverage limits, premiums, replacement values, home age); storm-type indicators (hurricane, tropical storm, subtropical system); and finally a count of all storms within 600 nautical miles during the 90-day exposure window. The latter control helps account for periods of high storm activity, where individual forecasts may carry different salience against a backdrop of multiple threats.

The results highlight three main patterns. First, true hits increase new insurance uptake by 14–50% across specifications. Demand rises most strongly when storm counts are controlled (Column 5), suggesting that direct impacts become more salient in busy storm seasons.

Second, false misses generate even larger increases in demand. Across all models, the effect of a false miss is statistically greater than that of a true hit, with differences ranging from 29 to 55 percentage points (bottom panel). This finding is consistent with loss aversion: unexpected losses trigger stronger behavioral adjustments than anticipated ones. A household surprised by an

Table 4: Baseline Model Estimates

Storm Classification	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
(a) Hit × Predicted Hit (True Hit)	0.14 (0.024)	0.25 (0.038)	0.25 (0.038)	0.31 (0.035)	0.50 (0.040)
(b) Hit × Predicted Miss (False Miss)	0.72 (0.083)	0.61 (0.055)	0.61 (0.069)	0.75 (0.089)	1.05 (0.144)
(c) Miss × Predicted Miss (True Miss)	0.03 (0.008)	0.02 (0.008)	0.02 (0.008)	0.04 (0.008)	0.04 (0.012)
(d) Miss × Predicted Hit (False Hit)	-0.31 (0.048)	-0.31 (0.055)	-0.31 (0.055)	-0.25 (0.053)	-0.18 (0.043)
Location & Time Fixed Effects	X	X	X	X	X
Random Effects		X	X	X	X
Home & Policy Variables			X	X	X
Storm Type Variable				X	X
Storm Count					X
Tests of Loss Aversion					
(1) Row (b) = row (a):	0.29	0.29	0.29	0.34	0.55
p-value	0.00	0.00	0.00	0.00	0.00
(2) Row (d) = row(c)	-0.32	-0.32	-0.32	-0.32	-0.35
p-value	0.00	0.00	0.00	0.00	0.00
Row (1) - Row (2)	0.90	0.90	0.90	0.84	0.72
p-value	0.00	0.00	0.00	0.00	0.00

unpredicted strike perceives a sharper need for protection than one that was warned in advance.

Third, false hits produce the opposite effect, a decline in demand of 18–31%. Rather than prompting precautionary behavior, false alarms appear to undermine risk perception, possibly by reducing trust in forecasts or by reinforcing beliefs that threats are exaggerated. As a result, the effect of a true miss is consistently greater than that of a False Hit, yielding a negative contrast that runs counter to standard loss aversion predictions. Instead of heightened caution, households may update downward their perceived need for coverage when faced with a predicted storm that never materializes.

The bottom panel of Table 4 reports difference-in-differences style contrasts across categories. In particular, the difference between false misses and false hits is large, positive, and highly significant in all specifications, ranging from 0.72 to 0.90. This result highlights the asymmetric sensitivity to expectation violations: unanticipated storm impacts drive sharp increases in demand, while unanticipated non-impacts reduce it, though by a smaller margin. The net effect is that forecast errors generate powerful behavioral responses, but in systematically different directions depending on whether the error exposes households to unexpected losses or unexpected relief.

9 Extensions and Robustness Checks

9.1 Distance and Expectations

The baseline specification relies on two central assumptions: first, that a storm hit is defined as occurring when the predicted or actual storm center falls within 300 nautical miles of a location; and second, that households form expectations based on the 72-hour forecast. In this section, I relax these assumptions by varying the distance threshold used to define a hit and later the forecast horizon. For consistency, all regressions adopt the structure of Model (4) from Table 4, which includes fixed effects, storm-type indicators, and policy-level covariates. Full regression tables are reported in the appendix; Table 5 presents results for alternative distance thresholds.

The distance experiments confirm the robustness of the baseline. When the hit definition is tightened to 200 nautical miles, results remain close to those at 300nm: true hits increase demand by roughly 30%, False Misses by nearly 50%, and False Hits reduce uptake. At the baseline 300nm threshold, the estimates reproduce the earlier pattern: true hits (31%) and false misses (75%) strongly raise demand, while false hits reduce it by 25%.

At the extremes, however, patterns weaken. When the threshold is narrowed to 100nm, the model produces unstable and sometimes counterintuitive estimates: true hits surge to 79%, while false misses shrink and even change sign. This misclassification is expected, many storms generate damaging effects well beyond 100nm, so the 100nm cutoff understates true exposure. Conversely, as the distance threshold expands, storm classifications become less informative and responses attenuate. At 500nm, true hits drop to a 12% effect, false misses fall to 68%, and false hits turn only mildly negative (-9%). True misses even shift slightly negative, reinforcing that distant non-events do little to alter perceived risk.

Tests of loss aversion remain consistent with the baseline across most thresholds. False misses continue to generate significantly stronger effects than true hits, emphasizing the asymmetric behavioral sensitivity to expectation violations. The exception is at 100nm, where misclassification obscures the underlying behavioral pattern.

9.2 Close Calls

To further examine the sensitivity of results to the definition of exposure, I introduce a "close call" category, defined as storms whose predicted or realized proximity lies between 250 and 350 nautical miles. In this scheme, storms within 250nm are classified as hits, those beyond 350nm as misses, and those in between as close. This refinement captures borderline cases where expectations are more ambiguous and risk perceptions less clear-cut.

Results are reported in Table 6. Consistent with the baseline, true hits generate strong increases in insurance uptake (19–66%), and false misses produce even larger effects (44–111%), reaffirming that unanticipated losses are the most powerful driver of demand. False hits do not appear in this specification, as few observations meet the criteria under the refined scheme which is consistent with earlier evidence that such cases are relatively rare.

The new "close" categories provide additional nuance. True close storms (predicted and realized near but non-striking) are associated with modest reductions in demand (-10% to -16%), suggesting that accurate forecasts of benign near-misses may dampen risk perception and foster complacency. By contrast, Close × Hit cases (storms predicted to skirt nearby but ultimately striking) produce large increases (45–83%), reflecting the salience of unexpected but proximate impacts. Conversely, Close × Miss cases consistently depress demand (-7% to -17%), consistent with desensitization or diminished trust in warnings when near threats fail to materialize.

Other combinations show weaker effects. Predicted hits that resolve as Close yield mixed, small coefficients, while missed predictions of close storms produce modest positive responses (21–24%), reflecting limited surprise.

Overall, the expanded classification confirms the central baseline result: true hits and false

Table 5: Estimates by Hit Definition Threshold

Storm Classification	**100 NM**	**200 NM**	**300 NM**	**400 NM**	**500 NM**
(a) True Hit	0.79 (0.053)	0.30 (0.034)	0.31 (0.035)	0.15 (0.027)	0.12 (0.016)
(b) False Miss	0.15 (0.040)	0.46 (0.082)	0.75 (0.089)	0.65 (0.063)	0.68 (0.056)
(c) True Miss	0.02 (0.009)	0.02 (0.009)	0.04 (0.008)	-0.02 (0.008)	-0.03 (0.007)
(d) False Hit	0.59 (0.103)	-0.11 (0.041)	-0.25 (0.053)	-0.35 (0.027)	-0.09 (0.056)
Tests of Loss Aversion					
(1) False Miss vs. True Hit	-0.36	0.13	0.44	0.50	0.93
p-value	0.000	0.090	0.000	0.000	0.000
(2) False Hit vs. True Miss	0.44	-0.13	-0.32	-0.15	-0.06
p-value	0.000	0.000	0.000	0.000	0.080
(3) Row (1) – Row (2)	-0.41	0.30	0.84	0.69	0.60
p-value	0.000	0.000	0.000	0.000	0.000

misses remain the dominant behavioral drivers of insurance demand. The "close call" categories highlight that ambiguous or marginal threats can themselves influence coverage decisions, but their effects primarily serve to nuance rather than overturn the baseline finding that expectation violations generate the largest responses.

9.3 Forecast Horizons

Table 7 examine how the choice of forecast horizon influences the estimated effects of storm classifications on insurance demand. Each column corresponds to a different forecast hour (36 through 120), with the hit threshold held constant at 300 nautical miles. This exercise explores whether behavioral responses are tied most strongly to short-, medium-, or longer-range forecasts.

Across horizons, true hits continue yield positive effects on policy uptake, though the magnitude varies. The largest effects occur at the 72-hour (31%) and 96-hour (35%) forecasts, suggesting that households are particularly responsive to medium-range predictions. These windows appear to strike a balance between credibility and salience: forecasts are recent enough to motivate action but sufficiently advanced to allow households time to respond.

False misses generally produce even stronger increases in demand, consistent with loss aversion. Effects are largest at the 60-hour (46%) and 72-hour (75%) horizons, emphasizing the salience of unanticipated losses. At shorter horizons (36–48 hours) and at 120 hours, however, coefficients fluctuate in sign and significance, reflecting the lower precision of very near-term or long-range forecasts. These instabilities highlight that expectation violations matter most when forecasts are both credible and actionable.

True misses show near-zero effects across horizons, reinforcing their similarity to no-storm baselines. By contrast, false hits are more variable: at 60–72 hours, they reduce demand by roughly 25%, consistent with the interpretation that false alarms erode risk perceptions. Yet at 48 and 120 hours, coefficients turn positive, suggesting that forecast errors at unusual horizons may occasionally reinforce rather than weaken perceived threat.

Tests of loss aversion confirm asymmetric responses. For the 60-, 72-, and 96-hour forecasts, false misses generate significantly larger effects than True Hits, while false hits reduce demand more than true misses. The difference-in-differences tests are largest at 60 and 72 hours (1.35 and 0.84, respectively), pinpointing these horizons as especially behaviorally relevant.

Taken together, the results suggest that the 72-hour forecast window provides the most consistent behavioral signal. At this horizon, all four storm classifications generate large and statistically significant effects, and the loss-aversion pattern is strongest. Medium-range forecasts thus appear

Table 6: Close Call Classifications

Storm Class	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
Predicted Hit × Actual Hit (True Hit)	0.19 (0.026)	0.19 (0.026)	0.37 (0.042)	0.41 (0.039)	0.66 (0.048)
Predicted Miss × Actual Hit (False Miss)	0.46 (0.085)	0.46 (0.085)	0.44 (0.085)	0.62 (0.114)	1.11 (0.185)
Predicted Miss × Actual Miss (True Miss)	0.03 (0.009)	0.03 (0.009)	0.01 (0.009)	0.03 (0.008)	0.04 (0.012)
Predicted Hit × Actual Miss (False Hit)	-	-	-	-	-
Predicted Close × Actual Close (True Close)	-0.10 (0.026)	-0.10 (0.026)	-0.16 (0.038)	-0.11 (0.037)	-0.07 (0.037)
Predicted Close × Actual Hit	0.52 (0.080)	0.52 (0.080)	0.45 (0.070)	0.56 (0.087)	0.83 (0.130)
Predicted Close × Actual Miss	-0.13 (0.029)	-0.13 (0.029)	-0.17 (0.034)	-0.10 (0.038)	-0.07 (0.039)
Predicted Hit × Actual Close	-0.02 (0.055)	-0.02 (0.055)	-0.05 (0.062)	0.08 (0.075)	0.19 (0.087)
Predicted Miss × Actual Close	0.24 (0.050)	0.24 (0.049)	0.24 (0.052)	0.23 (0.051)	0.21 (0.053)
Location & Time Fixed Effects	X	X	X	X	X
Random Effects		X	X	X	X
Home & Policy Variables			X	X	X
Storm Type Variable				X	X
Storm Count					X

to be the most salient anchor for household expectations, shaping flood insurance demand in ways consistent with reference-dependent models of decision-making.

9.4 Saliency

The baseline models treat all storms within the 90-day exposure window as equally relevant, implicitly assuming that households respond uniformly regardless of when the storm occurred. Yet salience is likely to vary across storms, with more recent events commanding greater attention and exerting stronger influence on behavior. Recency is a natural dimension of salience: storms occurring closer to the insurance decision date are more memorable and emotionally vivid, while the influence of earlier events may fade as households discount their relevance.

To examine this possibility, I re-estimate the model using 30-day bins to distinguish storms by their temporal proximity to policy inception. Specifically, the specification takes the form:

$$\log(Y_{it}) = \theta + Z_{i,31-60}\beta_1 + Z_{i,61-90}\beta_2 + Z_{i,91-120}\beta_3, \tag{6}$$

where θ includes the full set of location and time fixed effects, and each $Z_{i,\cdot}$ is a binary indicator equal to one if a storm fell within 300 nautical miles of location i during the designated 30-day window prior to the effective date of the policy. Observations without a qualifying storm in the 31–120 day window are assigned to a "No Storm" reference category. The model structure otherwise parallels that of Table 4, ensuring comparability.

Results are presented in Table 8. The estimates reveal substantial heterogeneity across exposure windows. Storms occurring within 31–60 days of policy issuance generate large and statistically significant increases in demand, with effects ranging from 46% to 113% across specifications. This finding reinforces that recency amplifies salience: households are especially responsive to threats that remain fresh in memory, consistent with models of limited attention and availability heuristics.

By contrast, storms that occurred 61–90 days earlier are associated with flat or modestly negative effects, while storms 91–120 days prior consistently reduce demand. Although the magnitudes are smaller than for recent storms, the direction of the coefficients suggests a decay in responsiveness as events become temporally distant. Put differently, the behavioral relevance of storm exposure diminishes over time, reflecting both fading memory and the declining psychological salience of past threats.

In a complementary test, I examine whether salience is also shaped by storm classifications. Specifically, I re-estimate the model using a stricter exposure criterion: only storms officially des-

Table 7: Estimates by Forecast Hour Reference Point

Storm Classification	**36 hr**	**48 hr**	**60 hr**	**72 hr**	**96 hr**	**120 hr**
(a) True Hit	0.21 (0.016)	0.20 (0.016)	-0.20 (0.010)	0.31 (0.035)	0.35 (0.021)	0.02 (0.065)
(b) False Miss	-0.16 (0.069)	-0.13 (0.079)	0.46 (0.161)	0.75 (0.089)	-0.12 (0.085)	0.08 (0.045)
(c) True Miss	-0.04 (0.012)	-0.03 (0.010)	-0.03 (0.014)	0.04 (0.008)	0.03 (0.013)	-0.008 (0.012)
(d) False Hit	-0.03 (0.044)	0.26 (0.041)	-0.25 (0.055)	-0.25 (0.053)	0.08 (0.034)	0.18 (0.069)
Tests of Loss Aversion						
(1) False Miss vs. True Hit	-0.31	-0.27	0.82	0.44	-0.38	0.06
p-value	0.000	0.000	0.000	0.000	0.000	0.13
(2) False Hit vs. True Miss	0.02	0.30	-0.22	-0.32	0.05	0.29
p-value	0.706	0.000	0.000	0.000	0.152	0.000
(3) Row (1) – Row (2)	-0.32	-0.54	1.35	0.84	-0.41	-0.17
p-value	0.000	0.000	0.000	0.000	0.000	0.000

Table 8: Estimates by Temporal Distance of Storm Exposure

Storm Window	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**
31–60 Days	0.53 (0.026)	0.54 (0.026)	0.46 (0.027)	0.54 (0.029)	1.13 (0.107)
61–90 Days	-0.08 (0.022)	-0.08 (0.022)	-0.08 (0.022)	-0.09 (0.022)	0.19 (0.040)
91–120 Days	-0.07 (0.015)	-0.07 (0.015)	-0.27 (0.018)	-0.25 (0.016)	-0.04 (0.023)
Location & Time Fixed Effects	X	X	X	X	X
Random Effects		X	X	X	X
Home & Policy Variables			X	X	X
Storm Type Variable				X	X
Storm Count					X

ignated as hurricanes are treated as predicted threats, while all other systems (e.g., tropical storms, subtropical depressions) are reclassified as equivalent to no storm. This binary distinction allows me to assess whether policyholders respond more strongly to high-salience hurricane warnings than to weaker storms. If salience matters, one would expect demand to rise primarily in response to hurricane-classified events, while less severe systems are discounted despite similar spatial proximity.

Table 9 presents estimates under four treatment variants. Column (1) reproduces the baseline specification from Table 4, which includes location and time fixed effects, random effects, and a full set of policy and property controls. Column (2) restricts the exposure definition to hurricanes only. Column (3) limits the window to storms occurring within 31–60 days prior to policy issuance, highlighting recency effects. Column (4) applies both filters simultaneously, focusing on hurricane-classified storms in the most recent 31–60 day period.

Across all models, true hits and false misses remain positively associated with increased insurance uptake, though magnitudes vary sharply by treatment. In the baseline (Col. 1), true hits raise demand by 25% and false misses by 61%. Restricting to hurricanes (Col. 2) attenuates these effects to 11% and 31%, suggesting that broader storm classifications play an important role in shaping behavior. In contrast, focusing on the most recent 31–60 days (Col. 3) amplifies effects: true hits increase demand by 69% and false misses by 36%. The strongest responses occur when both filters are applied (Col. 4): hurricane-classified storms within the most recent 31–60 days produce a 125% increase in uptake for true hits and a 147% increase for false misses. These dramatic magnitudes support the hypothesis that recency and severity interact to create particularly salient reference points.

The lower panel of Table 9 reports tests of loss aversion. In the baseline model, False misses exceed true hits by 29 percentage points, consistent with asymmetric loss sensitivity. The gap narrows to 18 points when restricted to hurricanes (Col. 2). When only recent storms are considered (Col. 3), however, the sign reverses: true hits dominate false misses. Finally, in the most restrictive specification (Col. 4), the difference is small and statistically insignificant. These results suggest that while loss aversion is evident under broad classifications, it weakens as storms become more salient. In highly salient contexts the psychological impact of severity and immediacy may overwhelm the incremental effect of expectation violations.

Together, these saliency-focused analyses reinforce the central conclusions. False misses consistently generate the largest increases in insurance uptake, while true misses remain negligible. At the same time, the magnitude of behavioral responses is strongly conditioned by cues of salience such as recency and storm severity. When storms are both recent and officially classified as hurricanes, demand surges dramatically, suggesting that salience can dominate or even displace loss-

Table 9: Model Estimates under Saliency Treatments

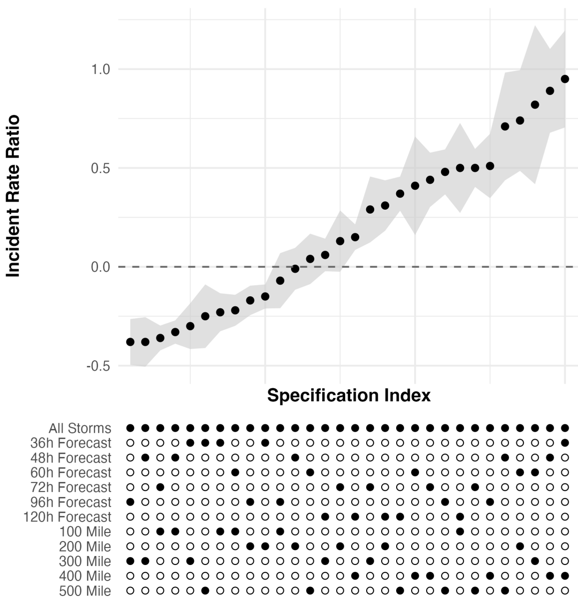
Storm Classification	**(1)**	**(2)**	**(3)**	**(4)**
(a) True Hit	0.25 (0.038)	0.11 (0.024)	0.69 (0.031)	1.25 (0.063)
(b) False Miss	0.61 (0.069)	0.31 (0.074)	0.36 (0.068)	1.47 (0.163)
(c) True Miss	0.02 (0.008)	-0.15 (0.009)	0.09 (0.013)	0.12 (0.018)
(d) False Hit	-0.31 (0.055)	-0.25 (0.009)	-0.19 (0.058)	0.00 (0.073)
Location & Time Fixed Effects	X	X	X	X
Random Effects	X	X	X	X
Home & Policy Variables	X	X	X	X
Hurricanes Only		X		X
31–60 Day Storms			X	X
Tests of Loss Aversion				
(1) False Miss vs. True Hit	0.29	0.18	-0.20	0.10
p-value	0.00	0.00	0.00	0.18
(2) False Hit vs. True Miss	-0.32	-0.12	-0.25	-0.11
p-value	0.00	0.04	0.00	0.10
(3) Row (1) – Row (2)	0.90	0.34	0.07	0.24
p-value	0.00	0.00	0.52	0.04

aversion dynamics. Across all robustness checks the evidence consistently highlights the asymmetric and psychologically mediated nature of insurance demand in the face of forecast uncertainty.

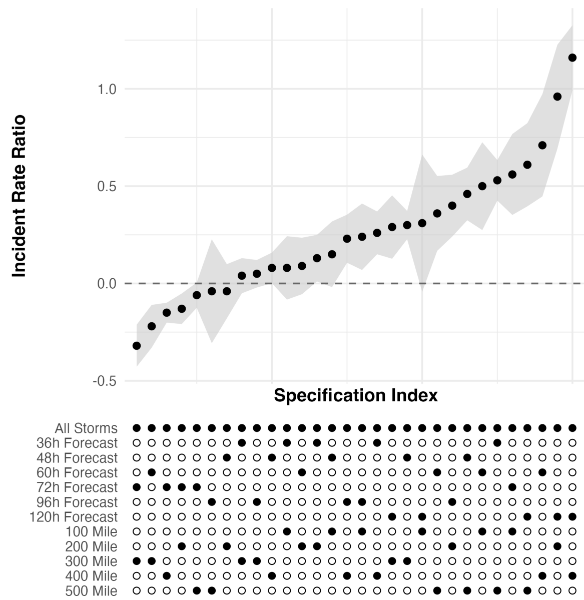
As a final analysis to assess the robustness of these behavioral findings, I conduct a specification curve analysis that systematically explores the sensitivity of results to modeling choices. Figure 7 reports the estimated effects for three core contrasts: (a) False misses versus true hits, (b) False hits versus true misses, and (c) the asymmetry between the two. Each point in the figure represents a distinct regression specification, varying forecast horizon, hit-distance threshold, and sample restrictions.

The results are strikingly stable. Across nearly all specifications, False misses produce strong and statistically significant increases in insurance uptake. Panel (a) shows a consistently positive and steep slope, reinforcing the effect of unanticipated storm impacts. By contrast, the effects of False Hits are weaker and more variable, as shown in Panel (b). In many cases households appear to discount over-predicted threats, interpreting false alarms as noise rather than signals.

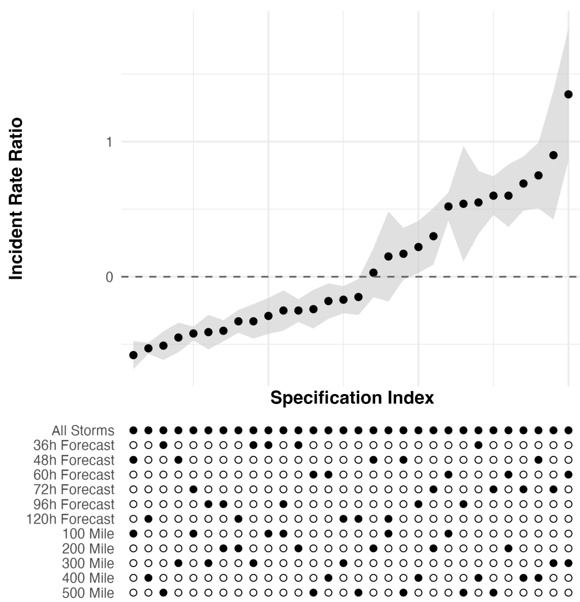
Panel (c) synthesizes these comparisons and reveals a clear asymmetry: demand responses are systematically stronger when expectations are violated by an unexpected hit than when a predicted hit fails to materialize. This asymmetry persists across horizons, thresholds, and subsamples, underscoring the robustness of the loss-aversion mechanism. Taken together, the specification curve analysis demonstrates that the paper's central claim that insurance demand is reference-dependent and shaped by the psychological salience of violated expectations holds across a wide spectrum of empirical choices.



(a) False Miss vs. True Hit



(b) False Hit vs. True Miss



(c) Asymmetric Behavior

Figure 7: Specification Curve Analysis of Forecast Accuracy Effects

10 Discussion

This paper provides new evidence that household responses to storm risk are mediated not only by objective exposure but also by the accuracy of forecasts that anchor expectations. By linking administrative NFIP records to hurricane forecasts, the analysis shows that deviations between predicted and realized storm outcomes generate powerful behavioral asymmetries. When households experience unexpected impacts, demand for flood insurance rises sharply, often more than after an anticipated strike. In contrast, false alarms depress demand, suggesting that warnings that fail to materialize erode trust and reduce perceived risk. These effects, which cannot be explained by actuarial exposure alone, demonstrate that insurance decisions are fundamentally reference-dependent.

The results extend models of reference-dependent preferences and loss aversion into the domain of natural hazard preparedness. Forecasts anchor household expectations; deviations generate predictable gain–loss responses. When realized events deviate from these expectations, behavioral responses follow predictable gain–loss patterns. Importantly, the paper shows that these dynamics operate in a real-world, high-stakes setting where financial consequences are large and the institutional framework is well defined. This helps bridge the gap between experimental evidence on reference points and actual household investment in protection against rare disasters.

Salience further conditions these effects. Recent storms and those officially classified as hurricanes produce especially strong responses, while accurate near-miss forecasts often depress demand. These patterns align with theories of availability and salience, in which vivid, severe, or recent events dominate attention and decision-making. The interaction between salience and reference dependence is particularly noteworthy: when storms are both recent and severe, the amplifying effect of salience can overwhelm loss-aversion dynamics, suggesting that attention can substitute for, or even dominate, expectation-based evaluation.

The findings carry important implications for risk communication and insurance market design. First, they highlight the behavioral consequences of forecast errors. If households under-react to accurate warnings yet overreact to surprises, coverage will remain volatile and chronically low. Forecast communication strategies could be refined to address these biases, for example by emphasizing uncertainty bands, probabilities, or regret-framed messaging that prepares households for a range of outcomes rather than a single trajectory. Communicating confidence intervals or past forecast accuracy could temper expectations and reduce the sharp behavioral swings that accompany surprises or false alarms.

Second, the results suggest that the timing of interventions matters. Salience is highest immediately after unexpected storm impacts. Targeted subsidies, reminders, or enrollment nudges during these behavioral windows may be more effective at expanding insurance coverage than ongoing, general outreach. Similarly, policies that reinforce the credibility of forecasts by reducing false alarms, or by contextualizing them as an feature of prediction could help sustain trust in forecast-based warnings.

Third, the evidence has implications for the NFIP itself. The program's structure, including the 30-day waiting period before coverage begins, creates a natural lag that may limit the extent to which heightened salience translates into enrollment. Aligning policy rules with behavioral responses for example, by offering temporary waivers or expedited enrollment after major storms could improve coverage when households are most motivated to insure.

The muted effect of true misses and the negative effect of false hits raise a behavioral puzzle. A Bayesian updating framework would predict little change in demand following a non-event. Instead, repeated false alarms appear to generate "forecast fatigue," eroding trust in risk communication. This backfire effect suggests that credibility is as central as accuracy in sustaining protective behavior. Another open question concerns heterogeneity: do these responses differ by income, geography, or prior exposure? Households with greater resources may react differently to forecast errors than more vulnerable populations, with important equity implications.

Several limitations should be acknowledged. The analysis focuses on new policy uptake, leaving unexplored the dynamics of renewals, lapses, and coverage changes. These behaviors may also reflect expectation violations but could operate differently, especially if inertia or commitment interacts with salience. The observational design, while rich in classification of exposure, cannot fully disentangle mechanisms whether households update beliefs about risk, shift attention, or revise trust in forecasts. Survey and experimental work could help measure expectations more directly, complementing the observational results.

Future research could also extend the analysis beyond Florida to other coastal states or to private flood insurance markets, where behavioral responses may differ in the absence of NFIP rules. In addition, comparative work could explore whether similar mechanisms operate in other contexts where forecasts shape protective action such as rainfall insurance in agriculture, epidemic risk in health, or financial risk management.

References

Abaluck, J. and Gruber, J. (2011). Choice inconsistencies among the elderly: Evidence from plan choice in the medicare part d program. *American Economic Review*, 101(4):1180–1210.

Abeler, J., Falk, A., Goette, L., and Huffman, D. (2011). Reference points and effort provision. *American Economic Review*, 101(2):470–492.

Allcott, H. and Taubinsky, D. (2015). Evaluating behaviorally motivated policy: Experimental evidence from the lightbulb market. *American Economic Review*, 105(8):2501–2538.

Allen, E. J., Dechow, P. M., Pope, D. G., and Wu, G. (2017). Reference-dependent preferences: Evidence from marathon runners. *Management Science*, 63(6):1657–1672.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5):941–973.

Arrow, K. J. (1971). *Essays in the Theory of Risk-Bearing*. North-Holland Publishing Company, Amsterdam.

Barseghyan, L., Prince, J., and Teitelbaum, J. C. (2011). Are risk preferences stable across contexts? evidence from insurance data. *American Economic Review*, 101(2):591–631.

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

Bordalo, P., Gennaioli, N., and Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics*, 127(3):1243–1285.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843.

Broad, K., Leiserowitz, A., Weinkle, J., and Steketee, M. (2007). Misinterpretations of the "cone of uncertainty" in florida during the 2004 hurricane season. *Bulletin of the American Meteorological Society*, 88(5):651–667.

Browne, M. J., Chung, J., and Frees, E. W. (2000). International property-liability insurance consumption. *Journal of Risk and Insurance*, 67(1):73–90.

Browne, M. J. and Hoyt, R. E. (2000). The demand for flood insurance: Empirical evidence. *Journal of Risk and Uncertainty*, 20(3):291–306.

Busse, M. R., Pope, D. G., Pope, J. C., and Silva-Risso, J. (2015). The psychological effect of weather on car purchases. *The Quarterly Journal of Economics*, 130(1):371–414.

Cai, J. and Song, C. (2017). Do disaster experience and knowledge affect insurance take-up decisions? *Journal of Development Economics*, 124(C):83–94.

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

Cangialosi, J. P., Landsea, C. W., and Blake, E. S. (2020). The historic atlantic hurricane season of 2020: A look back. *Weatherwise*, 73(6):28–35.

Card, D. and Dahl, G. B. (2011). Family violence and football: The effect of unexpected emotional cues on violent behavior. *The Quarterly Journal of Economics*, 126(1):103–143.

Cole, S., Giné, X., Tobacman, J., Topalova, P., Townsend, R. M., and Vickery, J. (2013). Barriers to household risk management: Evidence from india. *American Economic Journal: Applied Economics*, 5(1):104–135.

Congressional Research Service (2023). Introduction to the national flood insurance program (nfip). Technical Report R44593, Congressional Research Service.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, 75(4):643–669.

Emanuel, K. A. (2005). Increasing destructiveness of tropical cyclones over the past 30 years. *Nature*, 436(7051):686–688.

Ericson, K. M. and Starc, A. (2012). Heuristics and heterogeneity in health insurance exchanges: Evidence from the massachusetts connector. *American Economic Review*, 102(3):493–497.

Federal Emergency Management Agency (FEMA) (2023a). *NFIP Flood Insurance Manual*.

Federal Emergency Management Agency (FEMA) (2023b). Risk rating 2.0: Equity in action.

Frydman, C. and Mormann, M. (2016). The role of salience in choice under risk: An experimental investigation. *SMU Cox: Marketing / Working Paper*.

Gallagher, J. (2014). Learning about an infrequent event: Evidence from flood insurance take-up in the united states. *American Economic Journal: Applied Economics*, 6(3):206–233.

Giné, X., Townsend, R. M., and Vickery, J. (2008). Patterns of rainfall insurance participation in rural india. *The World Bank Economic Review*, 22(3):539–566.

Gollier, C. (2005). Some aspects of the economics of catastrophe risk insurance. Technical Report CESifo Working Paper No. 1409, CESifo / SSRN.

Grace, M. F., Klein, R. W., and Kleindorfer, P. R. (2004). Homeowners insurance with bundled catastrophe coverage. *Journal of Risk and Insurance*, 71(3):351–379.

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

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7):2643–2682.

Johnson, E. J., Hershey, J., Meszaros, J., and Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty*, 7(1):35–51.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292.

Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, 129(2):597–652.

Kimball, S. K. and Mulekar, M. S. (2004). A 15-year climatology of north atlantic tropical cyclones. part i: Size parameters. *Journal of Climate*, 17(18):3555–3575.

Kőszegi, B. and Rabin, M. (2006). A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4):1133–1165.

Kőszegi, B. and Rabin, M. (2007). Reference-dependent risk attitudes. *American Economic Review*, 97(4):1047–1073.

Kőszegi, B. and Rabin, M. (2009). Reference-dependent consumption plans. *American Economic Review*, 99(3):909–936.

Kousky, C. (2018). Financing flood losses: A discussion of the national flood insurance program. *Risk Management and Insurance Review*, 21(1):11–32.

Kousky, C. and Kunreuther, H. (2014). Addressing affordability in the national flood insurance program. *Journal of Extreme Events*, 1(1):1450001.

Kunreuther, H. (1978). *Disaster Insurance Protection*. John Wiley & Sons, New York.

Kunreuther, H. (1996). Mitigating disaster losses through insurance. *Journal of Risk and Uncertainty*, 12(2/3):171–187.

Kunreuther, H. and Pauly, M. (2004). Neglecting disaster: Why don't people insure against large losses? *Journal of Risk and Uncertainty*, 28(1):5–21.

Landsea, C. W. and Franklin, J. L. (2013). Atlantic hurricane database uncertainty and presentation of a new database format. *Monthly Weather Review*, 141(10):3576–3592.

Michel-Kerjan, E. O. and Kousky, C. (2010). Come rain or shine: Evidence on flood insurance purchases in florida. *Journal of Risk and Insurance*, 77(2):369–397.

Morss, R. E., Demuth, J. L., and Lazo, J. K. (2010). Communicating uncertainty in weather forecasts: A survey of the u.s. public. *Weather and Forecasting*, 25(2):357–371.

Mossin, J. (1968). Aspects of rational insurance purchasing. *Journal of Political Economy*, 76(4, Part 1):553–568.

National Hurricane Center (2022). 5-day forecast cone for hurricane ian, 8:00 am edt advisory.

NOAA National Hurricane Center (2019). Tropical cyclone definitions.

NOAA National Hurricane Center (2020a). Saffir-simpson hurricane wind scale.

NOAA National Hurricane Center (2020b). Tropical cyclone climatology.

NOAA National Hurricane Center (2022a). How to read nhc forecast/advisory products.

NOAA National Hurricane Center (2022b). National hurricane center products and services.

Pope, D. G. and Schweitzer, M. E. (2011). Is tiger woods loss averse? persistent bias in the face of experience, competition, and high stakes. *American Economic Review*, 101(1):129–157.

Post, T., van den Assem, M. J., Baltussen, G., and Thaler, R. H. (2008). Deal or no deal? decision making under risk in a large-payoff game show. *American Economic Review*, 98(1):38–71.

Robinson, P. J., Botzen, W. J. W., Kunreuther, H., and Chaudhry, S. J. (2021). Default options and insurance demand. *Journal of Economic Behavior & Organization*, 183(3):39–56.

Schlesinger, H. (2000). The theory of insurance demand. In Dionne, G., editor, *Handbook of Insurance*, pages 131–151. Springer, Dordrecht / New York / London.

Sydnor, J. (2010). (over)insuring modest risks. *American Economic Journal: Applied Economics*, 2(4):177–199.

Tallapragada, V. S. et al. (2014). Hurricane weather research and forecasting (hwrf) model: 2014 scientific documentation. Technical report, NOAA / National Weather Service / Hurricane Research Division.

Tversky, A. and Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5(2):207–232.

Wagner, K. R. H. (2022). Adaptation and adverse selection in markets for natural disaster insurance. *American Economic Journal: Economic Policy*, 14(3):380–421.

Wakker, P. P. and Deneffe, D. (1996). Eliciting von neumann-morgenstern utilities when probabilities are distorted or unknown. *Management Science*, 42(8):1131–1150.

Zweifel, P. and Eisen, R. (2012). *Insurance Economics*. Springer, 1st edition.

11 Appendix