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

Sierra Smith

October 13, 2025

Abstract

Households face rare but potentially catastrophic losses from hurricanes and must often rely on imperfect forecasts to guide protective action. This paper links hurricane forecasts to flood insurance uptake, using forecast errors as novel, exogenous reference points. The analysis reveals strong behavioral asymmetries: households respond sharply to unanticipated losses (false misses), weakly to anticipated ones (true hits), and negatively to false alarms. These patterns are consistent with reference-dependent preferences and loss aversion, suggesting that deviations between predicted and realized storm outcomes shape insurance behavior. The findings have implications for the design of forecast communication and disaster preparedness policy.

1 Introduction

Flood insurance uptake in the United States remains persistently low, even in high-risk regions such as coastal Florida. Many households remain uninsured, and coverage rates often decline once the memory of recent storms fades. This pattern presents a puzzle: if insurance protects against salient and potentially ruinous risks, why does demand remain limited and unstable? In this paper, I link hurricane forecasts to flood insurance uptake, arguing that forecast errors—the gap between predicted and realized storm outcomes—act as behavioral reference points that shape subsequent decisions to insure. These errors create expectation violations that trigger asymmetric responses consistent with loss aversion.

To illustrate, consider two similar households along the Florida coast, both who experience the effects of a major hurricane. For one, the storm makes landfall as predicted—a true hit that confirms the forecast. For the other, the storm hits contrary to forecast predictions—a false miss that was unexpected. While both households may become more likely to insure, the one experiencing the false hit is likely to respond more strongly due to unaligned expectations. The divergence in behavior reflects how forecast accuracy—or lack thereof—generates meaningful psychological contrasts that guide protective decision-making.

These dynamics build on foundational insights from behavioral economics. Prospect theory posits that individuals evaluate outcomes relative to reference points, with losses weighted more heavily than gains (Kahneman and Tversky, 1979). Expectation-based models formalize how anticipated outcomes anchor preferences and how surprises generate utility shocks (Kőszegi and Rabin, 2006, 2007, 2009). Salience theory further suggests that vivid or unexpected risks disproportionately capture attention and shape behavior (Bordalo et al., 2012). In the case of hurricanes, forecasts create salient expectations about risk, and deviations from those forecasts—whether in the form of a surprise impact or a false alarm—become psychologically meaningful.

Prior research documents a range of behavioral irregularities in insurance markets. Coverage tends to spike after disasters and decay over time, revealing cyclicality in demand (Kunreuther, 1978; Camerer and Kunreuther, 1989). In low-income settings, subsidized insurance products often see limited uptake (Giné et al., 2008; Cole et al., 2013), though simple nudges or direct experience increase enrollment (Karlan et al., 2014; Cai and Song, 2017). However, most studies proxy salience and experience using indirect measures, such as time since last disaster or aggregate losses. This paper instead uses forecast errors—quantifiable, exogenous, and widely observed—as a direct measure of behavioral reference points.

I merge NOAA hurricane forecasts from 2008–2023 with administrative records of flood insurance policies in Florida, focusing on a 90 day pre-policy window to identify how deviations between forecasted and actual storm paths shape insurance behavior. The analysis shows that unanticipated storm impacts (false misses) drive significantly larger increases in demand than expected ones (true hits), while false alarms (false hits) reduce take-up—consistent with loss aversion and salience.

This paper contributes to three literatures. First, it advances the study of disaster insurance by identifying behavioral frictions in coverage uptake. Second, it enriches models of reference-dependent preferences by introducing measurable forecast errors as anchors. Third, it speaks to the economics of salience, showing how public forecasts shape protective investment in high-stakes, real-world settings.

The approach also suggests broader applications. Forecast-based reference points may matter in other contexts—such as crop or health insurance—where risk expectations are formed in advance. In a changing climate, where storm frequency, severity, and forecast accuracy evolve, these behavioral dynamics may influence household adaptation. More broadly, the findings underscore how public information environments shape risk perception and decision-making.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 presents the behavioral framework and Section 4 formalizes the model. Section 5 describes the data and empirical strategy. Section 6 presents the results. Section 7 discusses extensions, and Section 8 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). Yet 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) shows that disaster insurance markets are characterized by low participation and pronounced cycles in coverage.

These anomalies have been attributed to a range of behavioral and institutional frictions. Homeowners 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 subsidized, take-up remains surprisingly low (Cole et al., 2013; Giné et al., 2008). Interventions such as reminders Karlan et al. (2014), framing Johnson et al. (1993), and default enrollment Ericson and Starc (2012); Handel (2013); Robinson et al. (2021) have been shown to shift coverage decisions, further highlighting behavioral sensitivities.

A complementary literature considers how experience shapes demand. Camerer and Kunreuther (1989) show that demand spikes in the aftermath of disasters but decays as memories fade. Barseghyan et al. (2011) document inertia in deductible choices and Gallagher (2014) shows that flood experience increases subsequent insurance purchases. Collectively, these studies underscore that insurance markets exhibit patterns poorly explained by static models of risk preferences, motivating closer attention to the role of expectations and reference points.

Behavioral economics provides a natural lens for understanding deviations from standard models. Kahneman and Tversky (1979) introduce prospect theory, which posits that outcomes are evaluated relative to a reference point and that losses loom larger than gains. Building on this, Kőszegi and Rabin (2006, 2007, 2009) formalize expectation-based reference dependence, where anticipated outcomes anchor preferences and surprises generate utility shocks. Models of salience Bordalo et al. (2012, 2013) further suggest that individuals overweight vivid, attention-grabbing attributes—such as disaster forecasts—relative to baseline risks.

Empirical studies have provided evidence of these mechanisms across diverse contexts. Sydnor (2010) shows that households overinsure modest risks in ways consistent with probability weighting. Wakker and Deneffe (1996) and Post et al. (2008) demonstrate how distorted probabilities and loss aversion shape choices in insurance and gambling. Abaluck and Gruber (2011) and Handel (2013) document choice inconsistencies and inertia in health insurance. In experimental settings, **?** show that reference points in effort provision systematically affect labor supply, while Pope and Schweitzer (2011) find that golfers treat par as a salient benchmark, leading to loss aversion in putts.

A closely related strand exploits forecast errors as exogenous shocks to reference points. Card and Dahl (2011) use unexpected NFL outcomes, relative to betting market forecasts, to show that forecast errors generate emotional responses with real behavioral consequences. Busse et al. (2015) find that unexpected weather conditions influence consumer car purchases, highlighting salience in market behavior. Allen et al. (2017) show how forecast errors in analyst predictions affect stock returns.

Within the salience tradition, Frydman and Mormann (2016) provide experimental evidence that attention and salience distort choice under risk. Additionally, Allcott and Taubinsky (2015) demonstrate that salience affects energy consumption decisions, highlighting broader domains where vivid signals shape protective investment. Together, these strands suggest that disaster forecasts may anchor expectations and that deviations from predicted outcomes constitute natural reference points for subsequent decisions.

Disaster contexts provide fertile ground for testing these ideas. Kunreuther and Pauly (2004) emphasize that households systematically under-insure against rare, high-loss risks. Browne and Hoyt (2000) and Michel-Kerjan and Kousky (2010) show that demand for catastrophe insurance is highly sensitive to recent losses. Gallagher (2014) and Hallstrom and Smith (2005) document how hurricane and flood experience alters household behavior. In developing-country settings, Giné et al. (2008) find low demand for rainfall insurance in India even when heavily subsidized, while Cai and Song (2017) show that prior disaster experience increases adoption of insurance in China.

Recent work explores mechanisms more directly tied to expectations. Bin and Landry (2013) find that floodplain residents adjust willingness to pay for coverage in response to updated risk perceptions. Kousky and Kunreuther (2014) emphasize the role of myopia and misperceived probabilities in shaping disaster preparedness. Robinson et al. (2021) demonstrate that default options increase flood insurance enrollment in the U.S., particularly after recent events. In parallel, Wagner (2022) develops a structural model of adaptation and adverse selection in disaster insurance markets, showing how expectation formation and selection interact dynamically.

This paper contributes to these literatures by introducing hurricane forecast errors—the gap between predicted and realized storm paths—as a novel measure of behavioral reference points. Forecasts are salient, credible, and widely disseminated, making them natural anchors for household expectations. Deviations from forecasts create psychologically meaningful contrasts that shape subsequent insurance demand. By merging forecast data with detailed Florida flood insurance records, the paper provides direct evidence on how expectation-based reference dependence and salience shape protective investment in a high-stakes, real-world market. In doing so, it advances the insurance demand literature, contributes to models of reference-dependent preferences, and deepens our understanding of how natural disasters influence economic behavior.

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 backed flood insurance to residents in flood-prone areas, thereby shifting post-disaster assistance from reactive aid to proactive risk pooling (Michel-Kerjan and Kousky, 2010; Kousky, 2018).

As of 2022, over 22,000 communities across 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), incorporating factors such as property elevation, location within or outside SFHAs, building age and structure type, 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 upfront, and coverage typically does not begin until 30 days after purchase—a rule explicitly designed to discourage last-minute purchases in anticipation of storms (Federal Emergency Management Agency (FEMA), 2023a). This 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 NFIP's structure as a commitment device, requiring ex ante recognition of risk and sustained participation—features that 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—referred to broadly 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 a 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 higher). 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 heightened potential for destruction (NOAA National Hurricane Center, 2020a).

Although hurricanes of any intensity can cause substantial damage—particularly through 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 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 National Hurricane Center (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 track, intensity, and spatial extent, and serves as a critical input to emergency management, media coverage, and household-level risk perception (NOAA National Hurricane Center, 2022b).

These forecasts are generated using a suite of numerical weather prediction models that solve complex physical equations governing 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 depiction of the probable path of the storm's center based on historical forecast errors over the past five years. The cone reflects only the uncertainty in the track forecast—not the size of the storm or the range of potential hazards—and it widens with forecast horizon, visually conveying the increasing uncertainty associated with longer lead times (NHC, 2023). 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).

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 may delay protective actions by individuals who perceive the information as ambiguous or unreliable. As such, forecast credibility, timing, and clarity are essential not just for scientific accuracy but for effective behavioral influence and insurance-related decision-making.

4 Behavioral Framework for Insurance Response

Insurance decisions in the face of hurricanes offer a unique window into how individuals perceive and react to risk under uncertainty. Because storms are both 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 individuals might interpret and respond to combinations of forecasted 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 experience interact to shape insurance demand.

4.1 Case 1: Experience-Driven Updating

The most straightforward intuition is that individuals respond primarily to the realized impact of a storm, rather than its forecasted trajectory. Because storms are rare and uncertain, homeowners may initially discount the risk and delay purchasing insurance. However, when a storm does make landfall nearby, the event becomes highly salient – prompting individuals to update their beliefs and seek coverage in anticipation of future threats. If this model holds, we would expect insurance demand to rise similarly following both predicted hits and predicted misses, as long as the storm ultimately hits. 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 storms Gallagher (2014); Kousky (2018), consistent with reactive, experience-driven responses. This may also reflect myopic risk assessment Kunreuther and Pauly (2004), wherein individuals underweight low-probability future threats until the risk becomes salient 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.

4.2 Case 2: Dual Sensitivity to Forecasts and Outcomes

A second possible interpretation is that individuals respond not only to outcomes but also to the 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 the risk (Tversky and Kahneman, 1973; Bordalo et al., 2012). Conversely, the latter should have the weakest effect, as neither the prediction nor the outcome signals 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 storm risk – either through forecast warning or unexpected landfall. However, the inconsistency between prediction and outcome may 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.

4.3 Case 3: Asymmetric Salience of Surprise Events

A related interpretation builds on Case 2, but places greater emphasis on the salience of realized outcomes. Individuals continue to respond to both forecasts 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 belief updating about both storm risk and the reliability of 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.

4.4 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 of the same magnitude. In the context of hurricane 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 on whether that outcome deviates from the forecasted 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. Conversely, 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.

5 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.

5.1 Insurance Demand Model

Consider a homeowner who, in each period, faces some risk of experiencing a damage-inducing storm. Let $i \geq 0$ denote the probability that the homeowner purchases flood insurance in a given period. This probability is influenced by the outcome of a contemporaneous storm, denoted 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 i^0 according to the psychological impact of the storm outcome, captured by gain-loss utility. Specifically, I assume:

$$i = i^o + \mu(y, p), \tag{1}$$

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 storm 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 elevates insurance demand, even in the absence of surprise. In contrast, the case $y = p = 0$- an expected miss- has no psychological salience and thus 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:

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

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

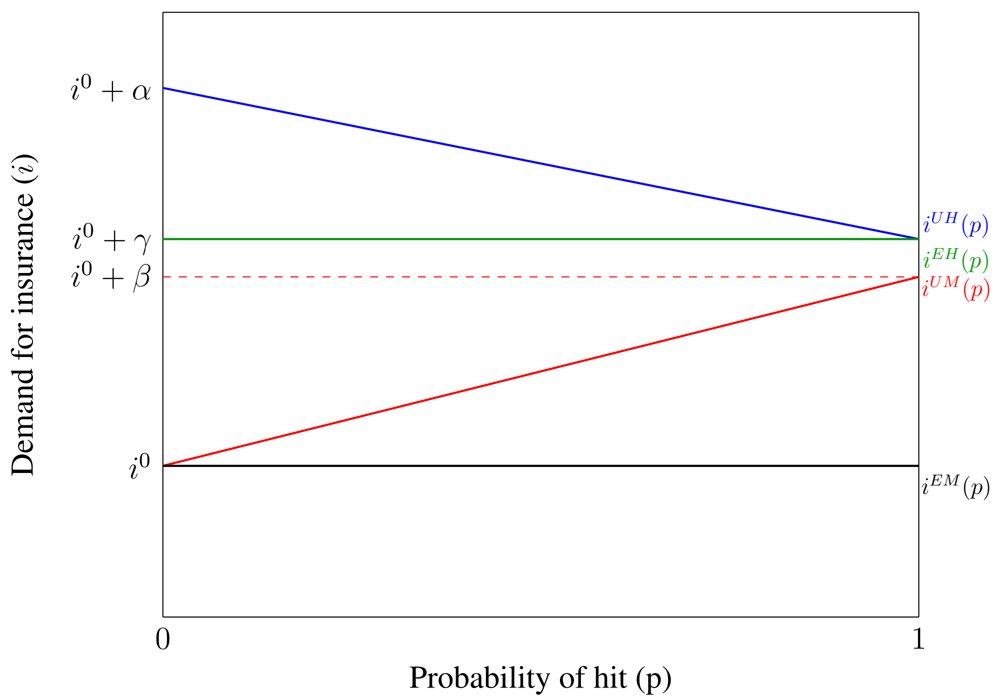


Figure 1: Demand Following a Hit or Miss

5.2 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 issuances in Florida. Storms are classified ex ante based on NOAA's 72-hour forecast: a storm is designated a predicted hit if it is expected to pass within 300 nautical miles of a location, and a predicted miss otherwise.

The empirical model includes interaction terms between these forecast classifications and actual storm outcomes (hit or miss), with no-storm periods serving as the reference group. This framework captures how forecast accuracy shapes insurance behavior—whether through correct predictions, unforeseen impacts, or overstated threats.

To assess robustness, I re-estimate the model using alternative forecast horizons (36–120 hours) and distance thresholds (100–500 nautical miles), and I incorporate a predicted close category (250–350 nm) to capture borderline cases. Additional robustness checks explore storm salience by testing whether insurance demand responds more strongly to recent storms (within 31–60 days) and to those officially classified as hurricanes. These extensions reveal that both temporal proximity and perceived severity amplify behavioral responses to risk.

6 Data Sources and Sample Construction

6.1 NFIP Policies

My empirical analysis draws on administrative records from the National Flood Insurance Program (NFIP) spanning the years 2009 to 2023. These data are provided by the Federal Emergency Management Agency (FEMA) and encompass comprehensive, nationwide information on both newly issued and renewed flood insurance policies. Each record includes details on both property characteristics—such as structure type, geographical coordinates (latitude and longitude), and estimated replacement costs—and policy-level attributes, including the effective date of coverage, original issue date, premium paid, and coverage limits for building and contents.

To focus on a more homogeneous policy and risk environment, I restrict the analysis to policies issued in the state of Florida, which consistently ranks among the most flood-prone and heavily insured states in the U.S. This geographic limitation allows for more precise alignment with Atlantic hurricane exposure and ensures consistency in regulatory context, floodplain management practices, and storm risk communication.

The NFIP covers a broad spectrum of property types, including single-family homes, multifamily residences, condominiums, apartment complexes, rental properties, and commercial structures. Because the central aim of this analysis is to investigate household-level insurance decisions, I restrict the sample to policies associated exclusively with residential houses—excluding policies written for commercial or multifamily properties.

To better isolate the determinants and timing of initial insurance uptake, I also exclude policy renewals and retain only newly issued policies. This distinction allows the analysis to focus on first-time purchase behavior, rather than ongoing coverage decisions, and better captures how households respond to dynamic risk factors—such as hurricane forecasts, storm experience, and policy salience—when making flood insurance enrollment decisions.

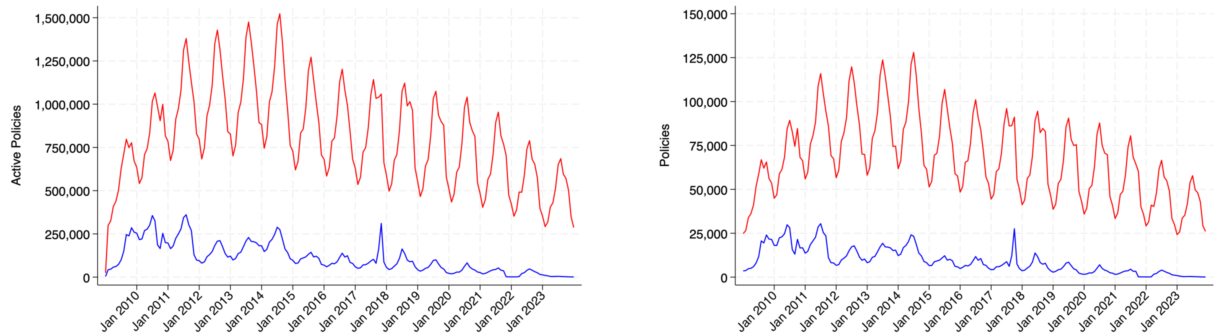
The resulting dataset comprises a high-resolution panel of new residential flood insurance purchases in Florida over a fifteen-year period, and forms the empirical foundation for the analyses that follow.

Figure 2 presents monthly trends in residential NFIP policy activity in Florida between 2009 and 2023. Panel A displays the total number of active policies, with the red line capturing the full stock of policies in force—including both renewals and new issuances—while the blue line isolates newly issued policies. In contrast, Panel B plots the same series but based solely on policy effective dates, thereby removing any noise associated with mid-cycle renewals or administrative delays.

Both panels reveal clear seasonal patterns, with policy activity peaking in late summer—typically around August and September—and falling to a low point in the winter months. This seasonality aligns with the timing of the Atlantic hurricane season and suggests that insurance demand is temporally responsive to perceived storm risk.

The total stock of active residential policies in Florida rose sharply in the early years of the sample, peaking around 2014–2015, before entering a gradual, persistent decline over the latter half of the period. This decline may reflect growing affordability concerns, increased private-sector competition, or declining perceived flood risk in certain areas.

New policy issuances, shown in blue, remain consistently lower in magnitude but display punctuated spikes, often occurring immediately before or after major hurricanes. These short-lived surges in uptake—followed by rapid drop-offs—suggest that salient events or forecasted threats may temporarily elevate demand, consistent with behavioral models of insurance uptake that emphasize recency and salience effects.



(a) Total Active Policies

(b) Effective Date Policies

Figure 2

Table 1 reports summary statistics for newly issued residential flood insurance policies in Florida between 2009 and 2023. The dataset includes over 1.66 million observations, reflecting first-time policy enrollments under the National Flood Insurance Program (NFIP). These statistics describe key financial and structural attributes of insured properties.

The typical insured home is a single-story primary residence, with an average of 1.35 floors and a mean age of 26 years. Roughly 83% of homes are designated as primary residences, while only 6% are classified as elevated structures, consistent with construction patterns in many older Florida neighborhoods. Home ages vary widely, ranging from new constructions to properties over two centuries old.

On the financial side, the average annual premium is approximately \$493, while the total policy cost—which includes additional surcharges and fees—averages around \$565. Households typically insure about \$205,000 in building coverage and approximately \$74,000 in contents coverage, though these amounts span a wide range, with some policies covering multimillion-dollar structures. Both building and contents deductibles are typically modest, averaging around \$1,450 and \$1,260, respectively.

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	1,663,846	25.96	18.21	0	267.71
Total Observations	1,664,216				

6.2 Storm Data

To capture household exposure to hurricane activity, I construct a storm-level panel using forecast and track data from the National Hurricane Center's (NHC) Tropical Cyclone Forecast/Advisory Archive, accessed via the NOAA Automated Tropical Cyclone Forecast (ATCF) system. This dataset includes all tropical cyclones in the North Atlantic basin between 2008 and 2023, and serves as the basis for measuring both realized storm characteristics and advance warnings available to households prior to landfall.

Each storm is observed across multiple forecast cycles, typically issued at six-hour intervals (00, 06, 12, 18), and includes both realized coordinates and intensity (the "best track") as well as projected paths, wind speeds, and radii at various 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.

For each forecast issuance, the dataset reports the geographic position (latitude and longitude) of predicted and observed storm centers, maximum sustained wind speed (in knots), which maps to hurricane categories via the Saffir-Simpson scale, and storm size estimates, including wind radii at 34-, 50-, and 64-knot thresholds in all four quadrants.

Table 2 summarizes annual storm counts by type and predicted intensity from 2008 to 2023, reflecting each storm's highest forecasted category. Tropical depressions and hurricanes are the most common types, with hurricanes defined as storms expected to exceed 74 mph. The number of predicted hurricanes varies widely, from just 2 in 2013 to 13 in 2020—a year of historically high activity. Category 1 storms are the most frequently forecasted, while Category 4 and 5 hurricanes are rare. Several years (e.g., 2013 and 2014) saw no storms predicted above Category 2, underscoring the episodic nature of severe hurricane threats. Total predicted storm counts range from 9 in 2014 to 31 in 2020, reflecting both underlying climatological cycles and variation in forecasted severity. Although rare, subtropical systems also appear in the dataset, highlighting the range of storm types communicated to the public and their potential influence on perceived risk.

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

Category	**Type**	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
	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
Storm Type	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
	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
Hurricanes	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

6.3 Linking Insurance and Storm Forecast Data

I merge the NFIP policy data with the NOAA/NHC storm forecast dataset along both temporal and spatial dimensions. First, I aggregate newly issued residential NFIP policies to the monthly level, recording the number of policies by effective month and geographic coordinates. To construct a complete panel of potential insurance activity, I then generate all possible combinations of month–year and geographic locations across Florida during the study period which results in 1,497 unique coordinates containing 180 observations. Location–month cells with no observed policies are coded as having zero new issuances, yielding a balanced panel that captures both presence and absence of demand.

This panel is then matched to all storms whose initial forecast dates fall between 31 and 120 days prior to the effective date of each policy, forming a three-month window of potential forecast exposure. For each storm–location pair within this window, I calculate the great-circle distance between the predicted storm center and the property's coordinates, based on the storm forecast at a fixed horizon—72 hours ahead in the baseline specification. When multiple forecasts exist for a given storm, I retain the one in which the predicted center is closest to the property. Additionally, storms within the window are ranked by their temporal proximity to the effective month to identify the most recent relevant forecast.

6.4 Expected Outcomes

Figure 3 provides a sample of the relationship between predicted and actual storm distances at the 72-hour forecast horizon, measured in nautical miles from the Florida panhandle. Each point represents a unique storm that was predicted to come within 600nm of the Florida Panhandle. The considerable vertical dispersion around the trend line reflects forecast error, with some storms missing or striking unexpectedly, despite proximity predictions.

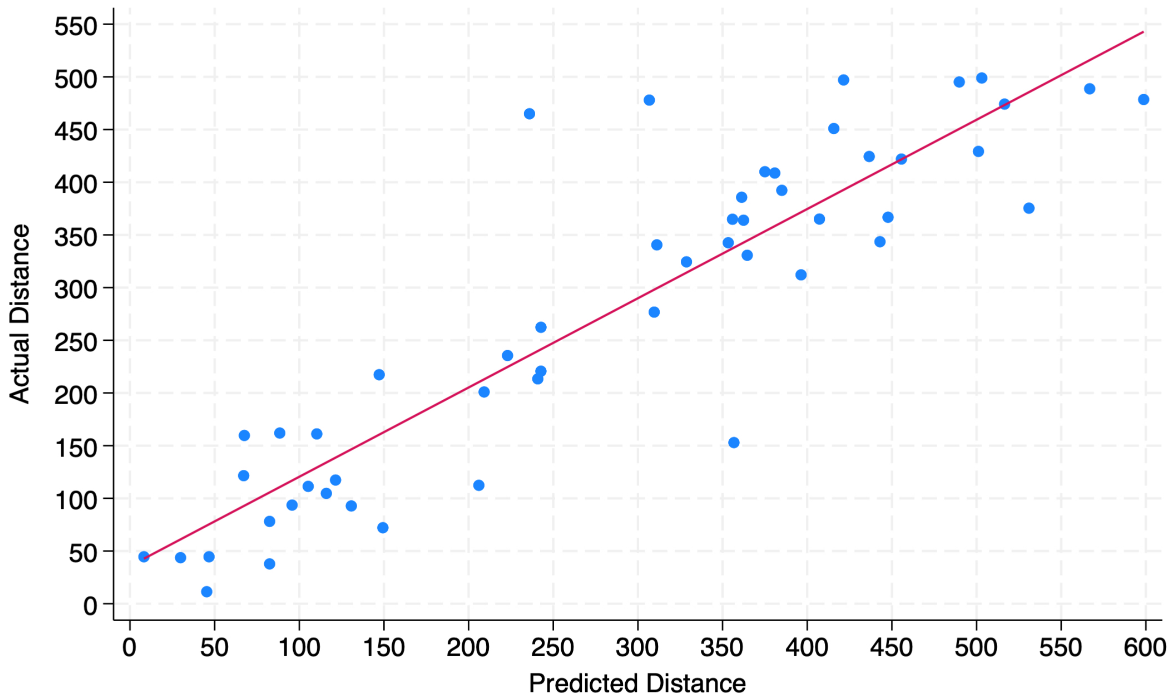


Figure 3: 72 Hour Predicted vs Actual Distance (Panhandle)

Each storm is further classified using a binary "Hit" indicator, defined as whether the storm's predicted center fell within a specified distance threshold—300 nautical miles in the baseline case—of the policy location. The 300 nautical mile threshold is chosen as the baseline hit definition because it approximates the average radial extent of tropical storm-force winds in Atlantic hurricanes, which typically range between 200 and 400 nautical miles from the storm center (Kimball and Mulekar, 2004). This distance captures a broad swath of plausible storm impact zones, balancing the need for sensitivity to forecast exposure with the goal of excluding distant, likely irrelevant systems. The same classification is applied using the storm's realized (actual) distance to capture whether the location was ultimately exposed to storm proximity. These classifications are used to construct measures of both forecasted and realized exposure.

Figure 4 plots the empirical probability of a realized hit as a function of the storm's predicted distance at the 72-hour forecast horizon. Vertical lines indicate classification thresholds: predicted hits (less than 250 nautical miles), close calls (250–350 nm), and predicted misses (greater than 350 nm). The curve reveals a steep decline in realized storm impact as forecasted distance increases. When a storm is predicted to pass within 250 nm, the actual probability of impact is approximately 70%, indicating relatively high forecast precision. In contrast, storms predicted to pass within 350 nm have only about a 20% chance of actually making impact.

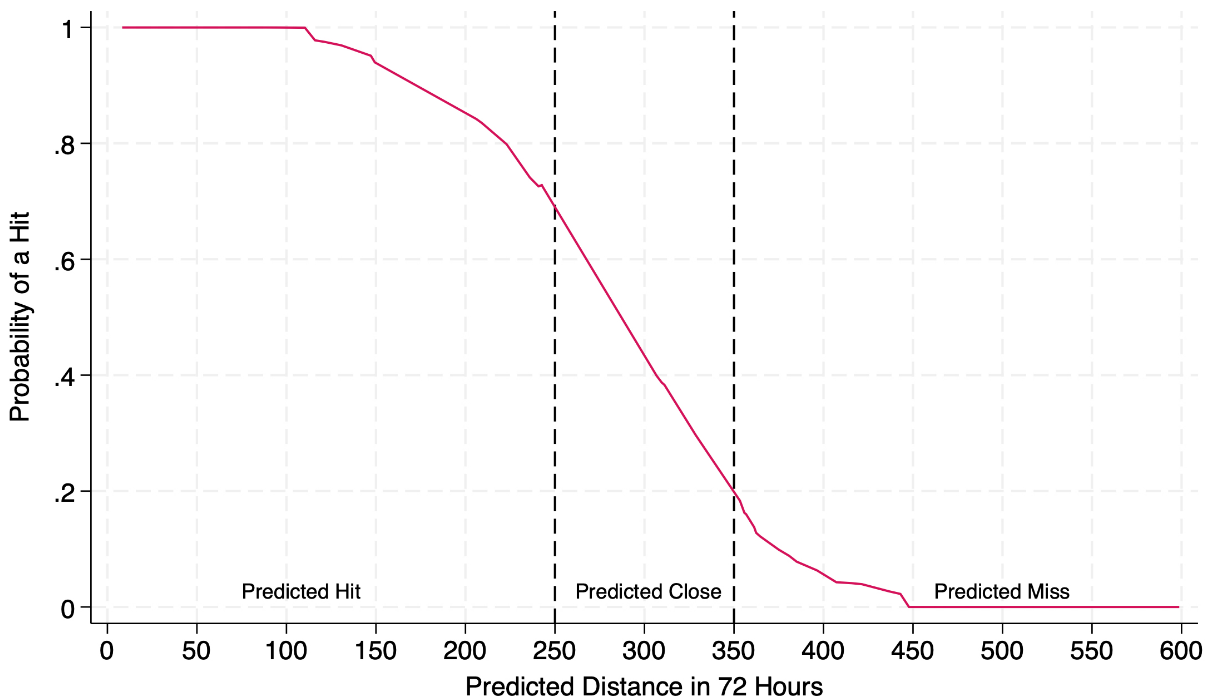


Figure 4: Lowess Graph

To classify storm exposure, I combine predicted and actual distance indicators into a four-category taxonomy of forecast accuracy. A True Hit occurs when both the forecasted and realized storm centers fall within 300 nautical miles of a location; a True Miss occurs when both fall outside this threshold. A False Hit refers to storms predicted to hit that ultimately miss, while a False Miss captures storms that were not forecasted to hit but eventually come within the threshold. Location–month cells with no storms during the 31–120 day exposure window are assigned a separate no-storm indicator. Table 3 presents the distribution of observations across these classifications. True Hits and True Misses account for over 60% of observations, with hurricanes and tropical storms representing the majority within these categories. By contrast, False Hits and False Misses constitute a small share of the sample, while periods without any nearby storms represent roughly one-third of observations.

Table 3: Distribution of Storm Storm Categories

Category	**Frequency**	**Percent**
True Hit	59,272	22%
Subtropical Storm	171	
Tropical Depression	7,347	
Tropical Storm	26,716	
Hurricane	25,038	
True Miss	113,778	42.22%
Subtropical Storm	86	
Tropical Depression	23,448	
Tropical Storm	43,619	
Hurricane	46,625	
False Hit	1,948	0.72%
Subtropical Storm	35	
Tropical Depression	862	
Tropical Storm	106	
Hurricane	945	
False Miss	4,850	1.80%
Subtropical Storm	0	
Tropical Depression	1,597	
Tropical Storm	3,223	
Hurricane	30	
No Storm	89,612	33.26%
Observations	269,460	

7 Empirical Methodology

The unit of analysis is a location–month, and the dependent variable is the number of newly issued NFIP residential policies in that location during month t. I estimate a Poisson regression model appropriate for count data, with the following general specification:

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

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 fixed effects, regional trends); and $f(p_{it}, a_{it}; \lambda)$ is a function of p_{it}, the perceived probability of a storm impacting location i at time t, and a_{it}, a binary indicator of the realized storm outcome, governed by behavioral parameter λ.

I assume that the perceived probability of storm impact, p_{it}, is a function of the forecasted distance, D_{it}, such that:

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

This formulation allows for the influence of actual outcomes a_{it} to vary depending on prior forecast salience. The coefficient of interest captures the behavioral effect of experiencing a hit or miss conditional on forecast proximity. I assume that forecasts disseminated by the National Hurricane Center (NHC) are viewed by households as informative and credible, and that conditional on the forecast, the realized outcome is treated as exogenous noise. This identification strategy allows me to distinguish between true positives, false alarms, false negatives, and true misses, and estimate their differential effects on flood insurance demand.

8 Baseline Empirical Results

Table 5 presents the results of three baseline Poisson regressions estimating flood insurance take-up in Florida between 2009 and 2023. In these models, I classify storm exposure based on the interaction between the storm's predicted proximity and its realized impact. Specifically, I define a predicted "hit" as any storm with a forecasted distance of 300 nautical miles or less, and a realized "hit" as any storm that ultimately came within that same threshold. The storm exposure function is specified 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 includes four mutually exclusive indicator variables: a true hit (λ_1) when the storm was both forecast and realized hit; a false hit (λ_2) when the storm was predicted to hit but ultimately missed; a true miss (λ_3) when no storm was predicted to hit and none did; and a false miss (λ_4) when no storm was predicted to hit but one did indeed hit. In addition, I include a fifth indicator for No Storm, defined as cases in which no storm occurred within the 31–60 day exposure window. This "no storm" category captures location–time observations without any relevant forecast or storm activity and serves as the reference group in the regression. Coefficients on the remaining categories are thus interpreted relative to this baseline of no forecasted or realized storm risk.

The simple model in Column (1) of Table 5 includes fixed effects for location, year, and month to account for spatial heterogeneity and temporal seasonality. Column (2) builds on this by introducing random effects to allow for unobserved variation across units. Columns (3) through (5) sequentially incorporate additional policy and property-level covariates, including average premium amounts, coverage limits, replacement values, and home age. Column (4) adds indicators for the type of storm associated with each observation—such as hurricane, tropical storm, or subtropical system. Finally, Column (5) includes a count variable for the total number of storms occurring within 600 nautical miles during the 90-day exposure window, controlling for broader variation in local storm activity.

Unsurprisingly, a storm hit results in a significant increase in insurance take-up. In the case of a True Hit—where a storm was both forecasted and realized within the threshold—insurance demand rises between 14% and 50%, depending on the model specification. The highest increase occurs in Column (5). Notably, all storm coefficients (except for True Miss) increase in magnitude under this specification.

A potential reason for this shift could be that the inclusion of total storm count captures background risk or informational saturation in high-storm periods. In such contexts, individual storms—may stand out more sharply against the backdrop of generalized risk, intensifying behavioral responses.

False Misses likewise exhibit substantial increases in insurance take-up. Across all specifications, the effect of a False Miss is consistently larger than that of a True Hit. Moreover, the difference between the two effects is statistically significant in every model, as shown in the lower panel of Table 5. This pattern provides empirical support for loss aversion: individuals appear to respond more strongly when a negative outcome (storm impact) violates their prior expectations, compared to when the outcome is both expected and realized. An unexpected loss likely sharpens the perceived need for future protection, thus elevating insurance demand beyond what a correctly forecasted storm elicits.

True Misses are associated with only modest increases in insurance take-up, ranging from 2% to 4%. This is perhaps unsurprising, as the realized outcome—no storm impact—closely resembles the experience of households that faced no storm at all. The result underscores the rarity of actual storm impacts and suggests that individuals may not adjust their perceived risk substantially in the absence of realized consequences.

False Hits, however, yield a more surprising finding: a consistent decrease in insurance demand, with effects ranging from an 18% to 31% reduction. Intuitively, one might expect that being warned

Table 4: Distribution of Storm Occurrence and 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

of a potential storm—even if it ultimately misses—would increase perceived risk and encourage precautionary behavior. Instead, the opposite appears to occur. One possible explanation is that households interpret the false alarm as confirmation that storm threats are overblown or unlikely to materialize. The incorrect forecast may reinforce a belief that their area is not truly at risk, thereby weakening the motivation to purchase coverage. In this way, a failed prediction may diminish trust in warnings and reduce demand for protection.

Because of this pattern, the effect of True Misses is consistently greater than that of False Hits, resulting in a negative contrast between the two, as shown in the loss aversion panel at the bottom of Table 5. While this difference is statistically significant across all specifications, it actually represents a violation of standard loss aversion predictions. Under loss aversion, we would expect that unexpected negative outcomes—such as a forecasted storm that fails to materialize (a False Hit)—would still evoke heightened concern and thus increase insurance demand more than the absence of any storm. However, the observed negative response suggests that some homeowners may interpret missed forecasts as overreactions or exaggerations, leading them to lower their perceived need for coverage rather than increase it.

The final row of Table 5 reports a difference-in-differences (DiD) estimate that directly contrasts the effect of a False Miss with that of a False Hit. Across all five specifications, this estimate is strongly positive and statistically significant. This indicates that the increase in insurance uptake following a False Miss exceeds the decline observed after a False Hit by a substantial margin. Put differently, unexpected storm impacts generate significantly stronger behavioral responses than unexpected near-misses, consistent with models of asymmetric loss sensitivity. This finding provides compelling empirical support for loss aversion in the context of insurance demand: individuals respond more intensely to realized threats they did not anticipate than to false alarms, even when both outcomes deviate equally from expectations.

9 Extensions and Robustness Checks

9.1 Distance and Expectations

The baseline specification rests on two core assumptions: first, that a storm "hit" is defined as occurring when the predicted or actual center of the storm falls within 300 nautical miles of a location; and second, that expectations are formed based on the 72-hour forecast. The models presented in this section relax these assumptions seperately by varying the distance threshold used to define a hit and the forecast hour used as a reference point. For consistency and comparability, all regressions adopt the structure of model (4) from Table 5, which includes fixed effects, storm type indicators, and key policy-level covariates. Full regression results for each specification are provided in the appendix. Table 6 displays the results of alternative hit definitions.

The estimates as distance is adjusted by 100nm (relative to the 300nm baseline) remain relatively stable. With behavior to the baseline, varying between 19% and -2% differences. Perserve behavior emerges as distances are relaexed to the extreme of 100nm. In the former specification, all coefficients, save for True Misses, experience large jumps in effects, with False misses switching signs entirely. This is perhaps not unexpected as 100nm distance is quite restrictive in term of beinging impacted by a storm. Many storm effects extend well past 100nm which leads to misclassification in the model.

Relative to the baseline specification at 300 nm, estimated effects remain relatively stable. However, more pronounced behavioral shifts emerge at the extreme threshold of 100 nm. In this restrictive case, several coefficients—most notably False Misses—exhibit large swings in magnitude and even reverse sign. This pattern likely reflects the overly narrow scope of the 100 nm threshold: storm impacts often extend well beyond that distance, leading to systematic misclassification of storm exposure and reduced reliability of the forecast taxonomy. As a result, the model may erroneously treat genuinely impacted locations as uninformed or unaffected, distorting behavioral responses.

As the distance threshold expands beyond the baseline, the insurance response to storm classifications weakens. True Hits decline from a 31% increase at 300nm to just 12% at 500nm, while False Misses drop from 75%to 68%, suggesting that distant storms are perceived as less salient. False Hits also become less negative (-25% to -9%), indicating that erroneous forecasts of far-off storms are less likely to reduce demand. Meanwhile, True Misses turn slightly negative, reinforcing the notion that homeowners largely ignore non-events when storms are distant.

Tests of loss aversion likewise exhibit patterns consistent with the 300nm baseline results and the broader findings in Table 5. Across the specifications, unexpected hits (False Misses) continue to generate significantly stronger effects than anticipated hits (True Hits), reinforcing the presence of loss aversion and asymmetric behavioral sensitivity. An exception occurs at the 100nm threshold, where the pattern breaks down—likely due to the forementioned misclassification.

To further examine the effects of hit definition, I expand the classification scheme to include a "close call" category, defined as storms whose forecasted or actual proximity falls between 250 and 350 nautical miles. Under this scheme, a storm is classified as a hit if within 250 nm, a miss if beyond 350 nm, and close if it falls in the intermediate range. This refinement allows for a more granular exploration of how expectations and outcomes—especially when ambiguous or marginal—shape insurance behavior.

The results, presented in Table X, reveal several key patterns. First, True Hits continue to generate strong and consistent increases in flood insurance uptake, with estimated effects ranging from 19% to 66% across model specifications. Similarly, False Misses—cases where storms were not forecasted to hit but ultimately did—produce sizable positive effects on demand, reaffirming the importance of unanticipated losses in driving behavior. Notably, False Hits are not included in this specification, as no observations meet the criteria under the refined distance-based classification scheme. This omission is expected, as earlier models already revealed a relatively small number of such cases.

Turning to the new "close call" classifications, a rich pattern emerges. True Close storms are associated with a negative effect on uptake, between -10% and -16%, suggesting that when a storm is accurately forecasted to pass nearby but ultimately causes no damage, it may dull risk perceptions and reinforce beliefs in forecast overreaction.

More notably, Close × Hit observations—storms forecasted to pass nearby but that actually strike—are associated with strong increases in demand, between 45% and 83%. These "stealth hits" violate expectations and appear to generate especially salient learning. By contrast, Close × Miss storms—forecasted to pass near but ultimately irrelevant—consistently depress demand (-7% to -17%), potentially reflecting a desensitization effect or overconfidence in benign outcomes.

Other combinations show more muted effects. For example, Predicted Hits that turn out to be Close (i.e., overshot predictions) produce mixed and generally small effects, while Missed predictions of Close storms still raise uptake slightly (21%–24%), reflecting modest surprise or elevated attentiveness.

While the expanded classification incorporating "close calls" adds interpretive richness, it largely reinforces the core patterns identified in the baseline model. True Hits and False Misses remain the primary drivers of insurance demand, and the new proximity-based categories serve more to validate and nuance the original findings than to challenge them.

Table 7 examines how the choice of forecast horizon influences estimates. Each column corresponds to a different forecast hour—from 36 to 120 hours—and uses the 300nm distance threshold to define a "hit." Across horizons, True Hits consistently yield positive effects on insurance de-

Table 5: Distribution of Storm Occurrence and Model Estimates by Hit Definition Threshold

Storm Classification	**100 NM**	**200 NM**	**300 NM**	**400 NM**	**500 NM**
(a) Hit × Predicted Hit (True Hit)	0.79 (0.053)	0.30 (0.034)	0.31 (0.035)	0.15 (0.027)	0.12 (0.016)
(b) Hit × Predicted Miss (False Miss)	0.15 (0.040)	0.46 (0.082)	0.75 (0.089)	0.65 (0.063)	0.68 (0.056)
(c) Miss × Predicted Miss (True Miss)	0.02 (0.009)	0.02 (0.009)	0.04 (0.008)	-0.02 (0.008)	-0.03 (0.007)
(d) Miss × Predicted Hit (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) Row (b) = row (a)	-0.36	0.13	0.44	0.50	0.93
p-value	0.000	0.090	0.000	0.000	0.000
(2) Row (d) = row (c)	0.44	-0.13	-0.32	-0.15	-0.06
p-value	0.000	0.000	0.000	0.000	0.080
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

Table 6: Close calls

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 (0.26)	-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 (0.07)	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

mand, though the magnitude varies. The largest effects are observed at 72- and 96-hour forecasts (31% and 35%, respectively), suggesting these windows best capture the anticipatory behavior of homeowners. False Misses, where storms unexpectedly hit, are generally associated with even larger increases in take-up—particularly at the 60- and 72-hour marks—highlighting a strong behavioral response to unanticipated losses. However, the direction of this effect is less stable across shorter (36-48hr) or longer (120hr) horizons, where coefficients fluctuate in sign and significance.

By contrast, True Misses exhibit near-zero effects throughout, reinforcing their similarity to a no-storm baseline. The pattern for False Hits is more mixed: while generally negative around the 60-72hr window (-25%), some horizons such as 48hr and 120hr even show positive values, indicating that predictive errors may occasionally reinforce rather than reduce perceived risk.

Tests of loss aversion further confirm asymmetric behavioral responses. For most horizons—especially 60hr, 72hr, and 96hr—the difference between False Misses and True Hits is positive and statistically significant, indicating stronger reactions to unanticipated hits. Similarly, False Hits tend to depress demand more than True Misses, though this relationship weakens at the longest and shortest horizons. The DiD test (Row 1 - Row 2) is largest at 60hr and 72hr (1.35 and 0.84, respectively), suggesting that these windows best capture the psychological salience of missed expectations.

Overall, the 72-hour forecast appears to be the most behaviorally relevant reference point—generating large and significant effects in all four storm classifications, and consistently supporting a loss-averse interpretation of insurance demand.

To capture variation in the timing of forecasted storm exposure, I also construct a categorical variable that assigns each location–month to one of four mutually exclusive groups: a predicted storm "Hit" within the 31–60 day, 61–90 day, or 91–120 day window prior to the policy's effective date, or no storm within the full 31–120 day period. Table 4 presents the distribution of observations across these categories. The majority of cases (69.4%) fall into the No Storm category, indicating that most policy periods are unaffected by nearby storm forecasts. By contrast, 5.6% of observations experienced a predicted storm hit in the most recent 31–60 day window, 7.2% in the 61–90 day window, and 17.8% in the 91–120 day window. The storm-type breakdown further shows that tropical storms represent the largest share of exposures across all bins, with hurricanes comprising only a small fraction of storm hits. This structure provides a basis for examining whether the recency of forecasted storm activity differentially shapes insurance demand.

Table 7: Distribution of Storm Occurrence and Model Estimates by Forecast Reference Point

Storm Classification	**36 hr**	**48 hr**	**60 Hour**	**72 Hour**	**96 Hour**	**120 Hour**
(a) Hit $\times$ Predicted Hit (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) Hit $\times$ Predicted Miss (False Miss)	-0.16 (0.069)	-0.13 (0.079)	0.46 (0.161)	0.75 (0.089)	-0.120 (0.085)	0.08 (0.045)
(c) Miss $\times$ Predicted Miss (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) Miss $\times$ Predicted Hit (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) Row (b) = row (a)	-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) Row (d) = row (c)	0.02	0.30	-0.22	-0.32	0.05	0.29
p-value	0.706	0.000	0.000	0.000	0.152	0.00
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	0.00

9.2 Saliency

So far, the empirical models have been relatively void toward the saliency of storms—neglecting to account for the possibility that some storms may command more attention or exert disproportionate influence on behavior. One natural dimension of salience is recency: storms that occur closer in time to the policy decision may be more memorable or emotionally salient than those that occurred further in the past. In the current empirical design, the first storm within the 90-day exposure window is used to characterize storm proximity. However, this may obscure heterogeneity in responsiveness across the exposure window.

To test whether more recent storms elicit stronger behavioral responses, I estimate the effect of storms occurring within 30-day bins using the following specification:

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

where θ captures the full set of location and time fixed effects, and each $Z_{i,\cdot}$ is a binary indicator equal to 1 if a storm fell within 300 nm of location i during the specified 30-day window prior to policy effectiveness. Observations with no qualifying storm within the 31–120 day window are assigned to a "No Storm" reference category. Table 8 displays the results using the specifications used in table 5.

The results reveal meaningful heterogeneity in insurance demand across the three temporal exposure windows. Demand increases substantially in response to storms that occur within 31–60 days of policy inception, with increases ranging from 46% to 113% across specifications. This pattern reinforces the idea that recency heightens salience, making more recent storm threats more influential in shaping insurance decisions.

By contrast, storms that occurred 61–90 days and 91–120 days prior are associated with modestly negative or flat effects on demand. While these estimates are small in magnitude, the consistently negative direction suggests a decay in behavioral responsiveness as storm events become more temporally distant. That is, the perceived relevance or urgency of past storm exposure appears to fade over time.

In a complementary test, I examine whether salience is also shaped by storm classification labels. Specifically, I re-estimate the model using a stricter exposure criterion: only storms officially designated as hurricanes are treated as predicted threats, while all other storm types (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 only in

Table 8: Binned days

Storm Class	**(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

response to hurricane-classified events, while less severe systems are discounted despite similar spatial proximity.

Table 9 presents model estimates under two saliency-based treatment variants. Column (1) reproduces the baseline specification from Table 5, which includes location and time fixed effects, random effects, and a full set of home and policy characteristic controls. Column (2) restricts the sample to storms classified exclusively as hurricanes, while Column (3) focuses on storms that occurred within 31–60 days prior to policy issuance. Column (4) applies both saliency filters, limiting the analysis to hurricanes that occurred in the most recent 31–60 day window.

Across all models, true hits and false misses remain positively associated with increased insurance uptake, though magnitudes vary substantially by treatment. In the baseline model (Column 1), true hits increase demand by 25%, and false misses by 61%. When the sample is limited to hurricanes (Column 2), the estimated effects shrink to 11% and 31%, respectively, suggesting that broader storm classifications may play a stronger role in shaping behavior. In contrast, when only storms in the most recent 31–60 day window are considered (Column 3), the estimated effects increase to 69% for true hits and 36% for false misses—highlighting the powerful role of temporal salience. The largest magnitudes are observed in Column (4), where hurricane-classified storms within 31–60 days are associated with a 125% increase in uptake for true hits and a 147% increase for false misses. This dramatic response supports the hypothesis that salient storm features—such as recency and perceived severity—amplify behavioral responses to risk.

The lower panel of Table 9 reports tests of loss aversion. In the baseline model (Column 1), the difference between false misses and true hits is statistically significant and positive (29 percentage points, p ¡ 0.01), consistent with loss aversion. A similar, though smaller, gap is observed when restricting to hurricanes (Column 2). However, in Column (3)—focusing on recent storms—the direction of the difference reverses, and in Column (4), the gap narrows further and is no longer statistically significant. These patterns suggest that while loss aversion holds under broader definitions, it weakens as storms become more salient—potentially because salience itself dominates the effect of misaligned expectations.

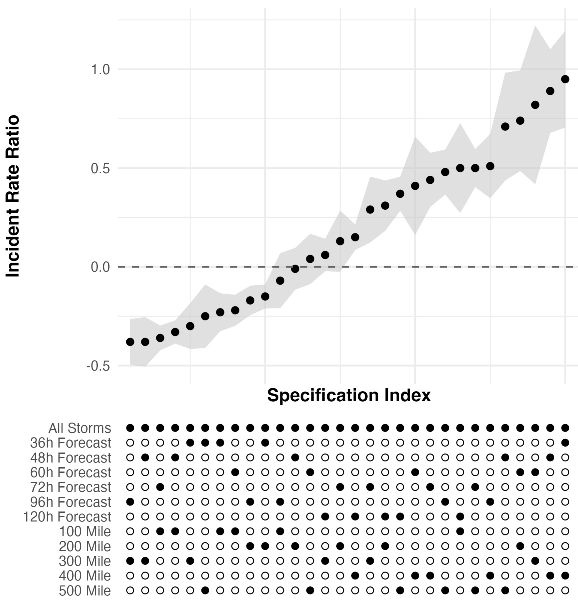
Together, the extensions and robustness checks presented in this section confirm the central findings of the baseline model: insurance demand responds asymmetrically to storm exposure, shaped by expectation violations, loss aversion, and salience. Across alternative hit definitions (100–500 nm), forecast horizons (36–120 hours), storm classification schemes (e.g., "close calls"), and saliency treatments (e.g., recent hurricanes only), the core behavioral patterns remain intact. Unexpected hits ("False Misses") consistently generate the largest increases in uptake, while anticipated non-events ("True Misses") have negligible effects. Demand is further amplified by psychologically salient cues—such as recency and storm severity—highlighting the behavioral dimensions of

Table 9: Distribution of Storm Occurrence and Model Estimates (Saliency Treatments)

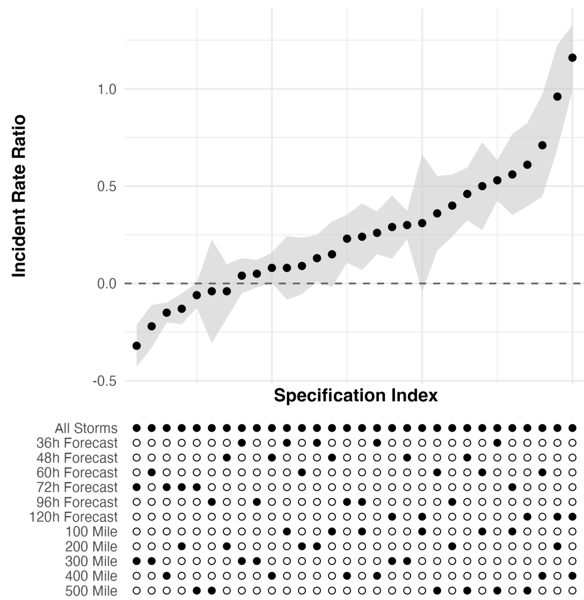
Storm Classification	**(1)**	**(2)**	**(3)**	**(4)**
(a) Hit × Predicted Hit (True Hit)	0.25 (0.038)	0.11 (0.024)	0.69 (0.031)	1.25 (0.063)
(b) Hit × Predicted Miss (False Miss)	0.61 (0.069)	0.31 (0.074)	0.36 (0.068)	1.47 (0.163)
(c) Miss × Predicted Miss (True Miss)	0.02 (0.008)	-0.15 (0.009)	0.09 (0.013)	0.12 (0.018)
(d) Miss × Predicted Hit (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) Row (b) = row (a):	0.29	0.18	-0.20	0.10
p-value	0.00	0.00	0.00	0.18
(2) Row (d) = row(c)	-0.32	-0.12	-0.25	-0.11
p-value	0.00	0.04	0.00	0.10
Row (1) - Row (2)	0.90	0.34	0.07	0.24
p-value	0.00	0.00	0.52	0.04

decision-making under risk.

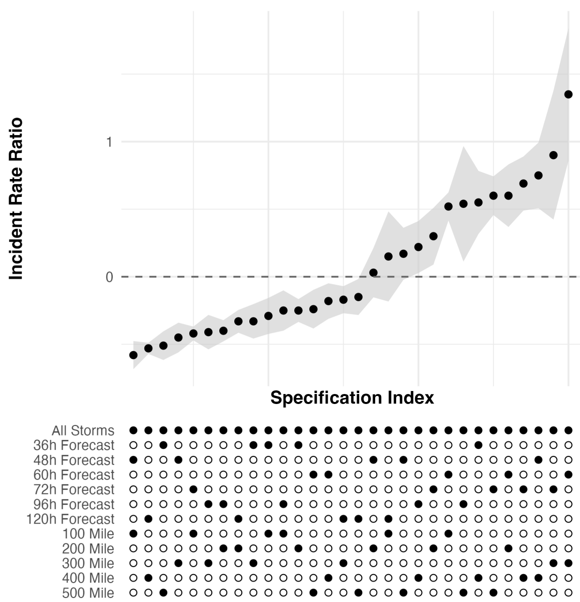
To assess the robustness of the core behavioral findings, I conduct a specification curve analysis across a wide range of modeling choices. Figure X presents the estimated effects for three central contrasts: the difference between the difference between False Misses and True Hits, False Hits and True Misses, and asymmetric responses between the two. Each dot in the figure represents a unique regression specification, varying systematically in forecast horizon (36h to 120h), hit distance threshold (100 to 500 nautical miles), and sample restrictions (e.g., hurricanes only). Across all panels, False Misses consistently produce strong and statistically significant increases in insurance uptake, regardless of model specification. This effect is especially robust in panel a, which compares False Misses to True Hits and shows a steep upward slope, highlighting the salience of unanticipated storm impacts. By contrast, False Hits generate weaker and more variable effects when compared to True Misses, suggesting that households may discount overpredicted risks. Panel c synthesizes these contrasts to reveal a clear asymmetry: demand responses are systematically stronger when expectations are violated by an unanticipated hit than when a forecasted hit fails to materialize. These results, which remain stable across a wide array of specifications, reinforce the interpretation that insurance behavior is reference-dependent, driven less by actuarial risk and more by the psychological salience of violated expectations.



(a) False Miss vs True Hit



(b) False Hit vs True Miss



(c) Asymmetric Behavior

Note: The analysis uses a 90-day forecast window and includes storm type controls. Confidence intervals at the 95% level are shown.

10 Discussion

This paper provides new empirical evidence on how individuals respond to forecasted and actual storm exposure in the context of flood insurance uptake. The most striking and consistent finding is the asymmetry between False Misses and False Hits. Storms that were not forecasted to hit but ultimately did lead to large and significant increases in insurance demand—often exceeding the effects of True Hits. By contrast, forecasted impacts that never materialize produce negative or null effects. This pattern cannot be explained by objective risk alone, as both types involve forecast errors of similar magnitude. Instead, it supports models of reference-dependent preferences, where outcomes are evaluated relative to prior expectations. The "pain" of being unexpectedly hit appears to outweigh the "relief" of being unexpectedly spared.

These dynamics are further highlighted in the nonlinear difference-in-difference tests that quantify asymmetric responses to gains and losses. The fact that insurance uptake is consistently more sensitive to unanticipated losses than to similarly unexpected gains is a hallmark of loss aversion, extending its empirical relevance to the domain of natural hazard preparation.

The salience of storm events also shapes behavioral responses. Storms that occur more recently (within 31–60 days of policy issuance) or that are classified as hurricanes (rather than weaker storm types) have disproportionately large effects on insurance uptake. These patterns are consistent with models of availability and salience, such as Bordalo et al. (2012), in which attention is selectively drawn to vivid, recent, or severe events. Interestingly, storms classified as "close calls" (250–350 nm away) that do not ultimately impact the area often reduce demand, potentially reinforcing a sense of invulnerability or diminishing the credibility of future warnings.

The muted effects of True Misses and the negative effects of False Hits present an interesting behavioral puzzle. A standard Bayesian updating model would predict little change in demand following a non-event, especially if the forecast was accurate. Yet, the decline in uptake after False Hits suggests a backfire effect—households may become less inclined to insure after being warned of a threat that fails to materialize. This pattern implies a form of forecast fatigue or erosion of trust, whereby repeated false alarms reduce the perceived usefulness of risk communication.

At the same time, True Misses produce only small increases in demand—consistent with both rational updating and with the notion that "nothing happened" confirms prior beliefs. These nuances help explain why insurance uptake remains stubbornly low even in high-risk regions, and why coverage tends to spike after disasters but then quickly decay.

These findings carry important implications for disaster risk management and insurance market design. First, they highlight the need for better calibration between forecast communication and public interpretation. If individuals systematically underreact to accurate warnings and overreact to surprises, insurance markets may suffer from volatility and underparticipation. Communication strategies that emphasize the uncertainty surrounding forecasts, or that frame messages in terms of potential regret or loss, may be more effective.

Second, the results suggest that policy interventions (such as subsidies or nudges) may be most effective after unanticipated storm events, when attention and perceived risk are heightened. This timing dimension could be critical for expanding coverage in at-risk areas.

Finally, from a theoretical standpoint, the paper contributes to the literature on reference dependence under uncertainty by demonstrating how forecast-based expectations operate as behavioral reference points, even in high-stakes, real-world decisions like disaster insurance.

Several limitations deserve mention. The analysis focuses on new policy uptake, rather than renewals or lapses, which could also reflect behavioral responses to risk. Additionally, while the observational design allows for rich classification of exposure, it cannot definitively disentangle causal mechanisms (e.g., whether salience drives attention, or whether attention alters beliefs). Future research could explore experimental or survey-based methods to more directly measure expectations, perceived accuracy, and trust in forecasts.

11 Appendix

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.

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.

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.