W-shaped implied volatility and investor learning about event risk

SIMEN GUTTORMSEN*

School of Economics and Business, Norwegian University of Life Sciences, Ås, Norway

(*April 2026*)

Implied volatility curves can occasionally exhibit a W-shape—a local maximum near at-the-money, rather than the standard U-shaped smile. This paper provides a structural explanation of this pattern, grounded in Bayesian learning about hidden regimes. The magnitude of potential jumps is unobservable, and investors must infer it from asset prices through Bayesian filtering. This investor learning creates belief uncertainty that inflates the volatility of the no-jump component of the risk-neutral density, which produces a three-component mixture that maps to a W-shaped IV curve. The main theorem establishes a necessary condition on beliefs for W-shapes to emerge. Comparative statics show that their likelihood and intensity peak at maximum uncertainty. The theoretical predictions are confirmed by empirical tests using U.S. equity options for 247 S&P 500 firms (2017–2025). W-shapes appear six times more often when options span an earnings announcement, with increasing likelihood closer to the event. Firms with stronger earnings reactions have more frequent W-shapes, while FOMC announcements do not produce them. Straddle returns for W-shaped volatilities do not differ from normal curves, suggesting the market correctly prices the elevated event risk.

Keywords: Implied volatility; W-shape; regime switching; Bayesian learning; earnings announcements; event risk

JEL Classification: G12; G13; D83

1. Introduction

Before an earnings announcement, there is often a buildup in implied volatility (IV) for at-the-money equity options that produces a W-shaped IV curve as it reaches its peak just prior to the announcement. The W-shape dissipates immediately after the announcement. This study demonstrates that W-shapes will form when investors are unsure if they will receive an earnings surprise or see their stock prices increase. The study tests five model assumptions using a large sample of U.S. equity options.

Unlike previous research into option prices—since Rubinstein (1985)'s non-parametric documentation of the symmetric IV "smile" and the post-1987 literature which developed the left-skewness of the "smirk" (Bates 1991, Rubinstein 1994)—this research examines the *middle* of the risk-neutral distribution (the *body*) rather than the *tails*. Tail dynamics like fat tails, excess kurtosis, and crash risk have been explained with various jump-diffusion models (Merton 1976), stochastic volatility models (Hull and White 1987, Heston 1993), and combinations of both (Bates 1996, Duffie *et al.* 2000, Bakshi *et al.* 1997). Further, a large body of research documents how option-implied tail risk is priced around macroeconomic announcements (e.g., Jacobs *et al.* 2025). It is similarly well

*Corresponding author. Email: simen.guttormsen@nmbu.no

established that implied volatility rises prior to scheduled earnings announcements and collapses shortly thereafter (Patell and Wolfson 1979, Ederington and Lee 1996). The W-shape, by contrast, reflects non-normality in the *body* of the distribution: a depletion of probability mass near the current price, precisely where most of the weight ordinarily resides. This body-level distortion has received far less theoretical attention.

The economic rationale behind this is clear. Prior to an earnings announcement, the market expects one of three possible outcomes: a "no surprise" middle ground in which there is little movement in the price of the stock, and two tails representing positive and negative surprises. This results in a risk-neutral distribution that is hollowed out at the center and has mass concentrated in moderately high levels of both positive and negative movement—the multi-modal shape that corresponds to a W-shaped implied volatility curve.

There exists some literature on parts of this phenomenon. Hull (2009) describes the inversion of a "U"-shaped IV curve, referred to as an inverted U, or a "frown", describing how it arises from a bimodal risk-neutral distribution when anticipating a single large jump. Glasserman and Pirjol (2023) describe mathematically how W-shaped IV curves arise from mixtures of at least three Gaussian distributions with the central component having greater volatility than the wings. Alexiou *et al.* (2025) find that firms exhibit concave IV curves prior to their earnings announcements and that these IV curves completely disappear immediately upon the announcement date. In addition, firms that exhibit concave IV curves demonstrate significantly higher absolute returns on the announcement date compared to firms that do not exhibit such IV curves. While Glasserman and Pirjol (2023) provide a mathematical description of how W-shaped IV curves can be generated via a mixture of distributions, the characterization of mixture weights as exogenously determined leaves unanswered questions regarding what generates the weights associated with the distributions and why such weights vary systematically over time. A connection exists between the shapes of IV curves and risk-neutral densities as demonstrated via the Breeden and Litzenberger (1978) formula. Mixture distributions have been used previously to estimate bimodal beliefs (Melick and Thomas 1997). However, previous research has failed to provide an explanation for why the mixture weights that generate W-shaped IV curves develop around events and then vanish post-event.

This study presents an investor-learning-based explanation for why W-shaped IV curves form prior to earnings announcements and then disappear immediately after the announcement. The basic premise underlying the study is that investor beliefs concerning whether a jump regime is operative determine the weights associated with each component of the mixture, while the *uncertainty* surrounding those beliefs increases the effective volatility of the central (no-jump) component, thereby meeting the conditions required for W-shaped IV curves as defined by Glasserman and Pirjol (2023). Immediately preceding an announcement date, investors assign some degree of belief π to a jump regime being operative. This reflects a state in which the stock will either experience a sharp upward or downward movement based on whether earnings exceed or fall short of analyst forecasts. Conversely, investors believe with probability $1 - \pi$ that a calm regime is in operation—i.e., that the announcement does not represent a significant departure from analysts' expectations and therefore experiences minimal price movement. Investor beliefs evolve according to the Wonham filter (Wonham 1965)—the optimal Bayesian filter for a hidden Markov state observed through a counting process. Prior to the announcement, non-arrival of jumps represents passive evidence against the event regime; hence, beliefs update deterministically as opposed to actively extracting information from the diffusion path. The closer investors get to the event date, and thus the higher their uncertainty as to whether a surprise will occur, the greater will be the degree to which the risk-neutral density is hollowed out at the center, thereby producing a W-shaped IV curve. Following the announcement, investor beliefs collapse into certainty and the W-shaped IV curve disappears. The study's methodology relies on prior literature demonstrating that Bayesian learning about unobservable states leads to excessive volatility and overreaction to new information (David 1997, Veronesi 1999, Pástor and Veronesi 2003). Hu *et al.* (2022) demonstrate that heightened uncertainty about the *magnitude* of an upcoming announcement produces its own separate risk premium—apart from the news itself—this mechanism works at the variance

level and this study extends it to the shape of the IV curve. Additionally, the model relies on prior work using regime-switching models in option pricing (Hamilton 1989, Naik 1993, Bollen 1998), as well as the double-sided jump models of Kou (2002).

The major theoretical result characterizes when W-shapes will occur: they require investor belief π to fall in the interval $(\pi_{\text{crit}}, 1 - \pi_{\text{crit}})$. In addition to identifying when W-shapes will appear, the model yields several testable hypotheses: (i) W-shapes should appear more often close to events that investors expect may increase their belief about the value of an asset; (ii) W-shapes should cease to exist after an event has occurred as investor beliefs become more certain; (iii) the distribution of realized outcomes for those events classified as W-shapes should have fatter tails than other events because there are three possible sources of outcomes (the two extremes and the middle); (iv) the frequency of appearance of W-shapes should vary inversely with the magnitude of event reactions among firms, since firms that experience large event reactions are expected to have less ambiguity associated with them; and (v) systematic events that produce no firm-specific bimodal risk should never produce W-shapes. Empirical tests conducted on U.S. equity options for 247 S&P 500 constituents over earnings announcements support each of these five hypotheses. A primary economic conclusion is that W-shapes signal high perceived risk of an event: earnings events preceded by W-shapes exhibit average absolute returns of 5.9%, while non-W-shape events yield average absolute returns of 4.3% ($p < 0.001$). Further, the outcomes for W-shape events display the bimodal nature of the three-component mixture. Additionally, when W-shapes appear, the implied volatility is significantly greater (70% vs. 51%) but the realized volatility does not differ (31% vs. 30%), which connects W-shapes to the variance risk premium.

The contribution to the literature is threefold. First, this paper provides economic explanations for W-shapes through demonstrating how W-shapes arise rationally due to Bayesian learning rather than modelling the mixture weights as exogenous (Glasserman and Pirjol 2023) or simply documenting the phenomenon without providing a structural rationale (Alexiou *et al.* 2025). The model also provides both a structural and behavioural explanation for why the patterns occur at particular times (prior to an event and sharply after an event) and for why W-shapes contain significant economic content, as well as making cross-sectional predictions that are verified in the data. Second, structural estimation of filtered beliefs via a hidden Markov model creates a unifying explanation for how the shape of IV curves relates to investor beliefs, and ultimately to the variance risk premium, and estimates the critical parameter of uncertainty η based upon observables. Finally, direct distributional tests confirm that realized outcomes for W-shape events come from a heavy-tailed, hollowed-out distribution. Furthermore, the sharp drop-off in W-shape occurrences immediately following an event occurs as required by Bayes' rule and rational updating of beliefs by investors.

The remaining sections of this paper are organised as follows. Section 2 develops the theoretical foundation, including the regime-switching model, investor beliefs, and the main theorem describing when W-shapes can emerge. Section 3 describes the data and concavity detection method. Section 4 presents the empirical results. Section 5 concludes. Proofs are collected in Appendix A and robustness tests in Appendix C.

2. Theoretical Framework

Figure 1 illustrates the phenomenon that the model explains: a W-shaped IV curve (left panel) and the corresponding three-component risk-neutral density (right panel). Figure 2 compares an empirical W-shape observed before earnings with a standard convex curve.

2.1. *Model*

Time is continuous, $t \in [0, T]$. There is a single risky asset with price S_t and a risk-free asset with constant rate r. The economy is in one of two regimes, $s_t \in \{0, 1\}$. In regime 0, the calm regime, the

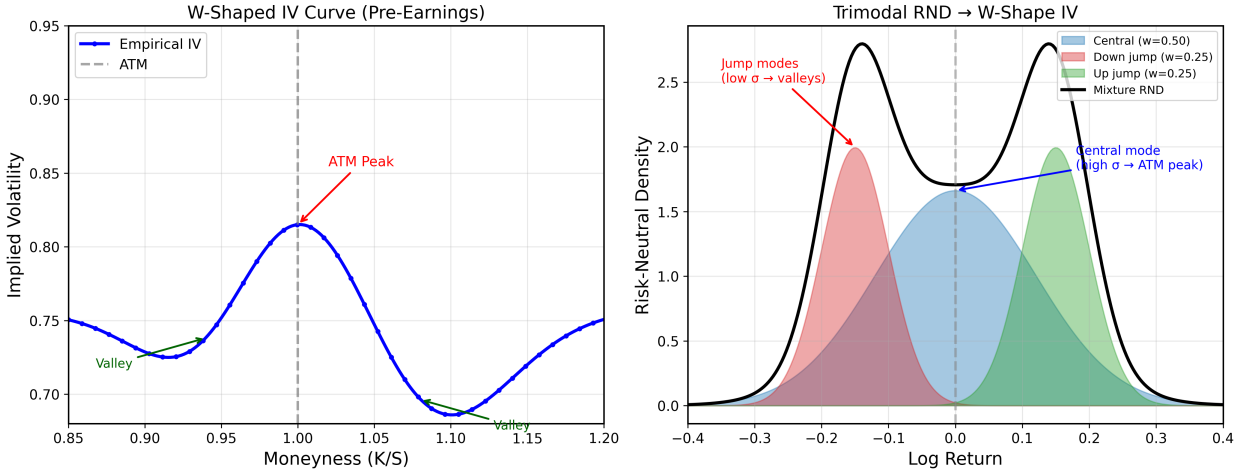

Figure 1. W-Shaped Implied Volatility and the Three-Component Mixture. Left panel shows a W-shaped IV curve with a local maximum near at-the-money (ATM), generated from the model with belief $\pi = 0.5$ and time to maturity $T = 0.1$. Right panel shows the corresponding risk-neutral density. The three-component mixture: a central "no jump" component (blue, weight 0.50) with higher volatility, and two wing components representing down jumps (red, weight 0.25) and up jumps (green, weight 0.25). A W-shape arises because the central component has higher volatility than the wings, inflating implied volatility near ATM to create a local maximum between the two valleys.

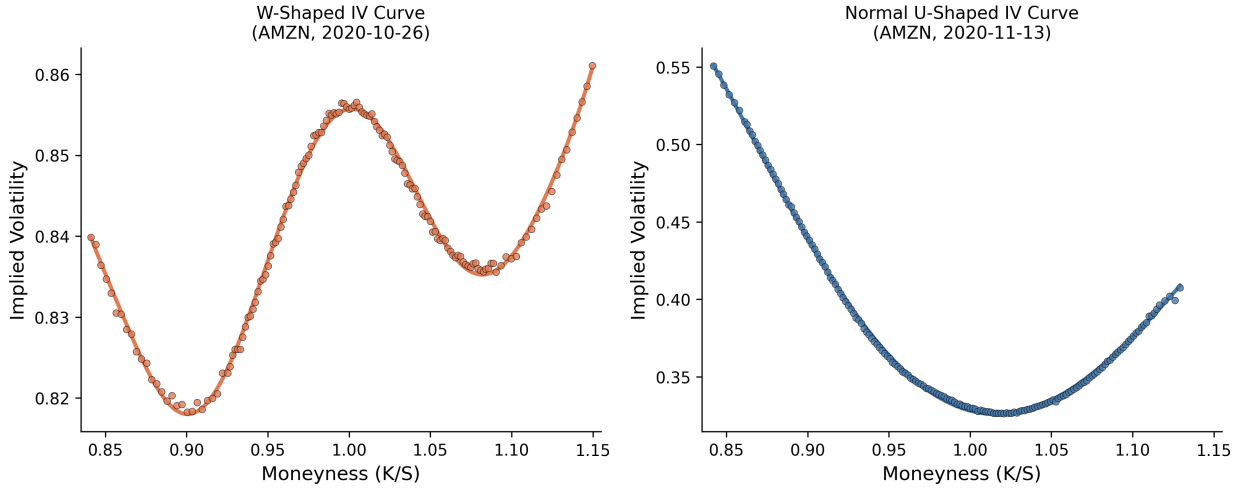

Figure 2. Empirical W-Shape vs. Normal IV Curve. Comparison of implied volatility curves for AMZN. Left panel shows a W-shaped curve on October 26, 2020 (expiry October 30, spanning the October 29 earnings announcement). Right panel shows the standard U-shaped curve on November 13, 2020 (expiry November 20), two weeks after the earnings announcement when event uncertainty has resolved.

asset follows pure diffusion. In regime 1, the event regime, jumps are possible. The regime follows a continuous-time Markov chain with generator

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 \\ \lambda_1 & -\lambda_1 \end{pmatrix} \tag{1}$$

where λ_0 is the rate of entering the event regime and λ_1 is the rate of leaving it. Before earnings, λ_0 may be elevated as the event becomes more likely. After the announcement, λ_1 is high as uncertainty resolves quickly.

Under the risk-neutral measure $\mathbb{Q}$, the asset price satisfies

$$\frac{\mathrm{d}S_t}{S_{t-}} = (r - q - s_t\lambda_J\xi)\mathrm{d}t + \sigma\mathrm{d}W_t^{\mathbb{Q}} + (e^{J_t} - 1)\mathrm{d}N_t \tag{2}$$

where $\xi = \mathbb{E}^{\mathbb{Q}}[e^J - 1]$ is the expected proportional jump size (ensuring the discounted price is a $\mathbb{Q}$-martingale), $W_t^{\mathbb{Q}}$ is a standard Brownian motion under $\mathbb{Q}$, N_t is a counting process with regime-dependent intensity $\lambda_N(t) = s_t \cdot \lambda_J$, and J_t is the log jump size with bidirectional distribution

$$J \sim \begin{cases} \mathcal{N}(\mu^+, \sigma_+^2) & \text{with probability } p \\ \mathcal{N}(\mu^-, \sigma_-^2) & \text{with probability } 1 - p \end{cases} \tag{3}$$

where $\mu^+ > 0 > \mu^-$, so that jumps can be either up or down. The bidirectional jump structure is crucial as it creates the three-component mixture (no jump, down jump, up jump) needed for W-shapes.

2.2. *Information Structure and Investor Beliefs*

Assumption 2.1 (Observable Prices, Hidden Regime) Investors observe the asset price path $\{S_u : u \leq t\}$ but not the regime s_t directly. They can infer information about the regime from price movements, especially jumps.

The investor's belief regarding whether the event regime is currently active at time t, given observed prices, is denoted $\pi_t = \mathbb{Q}(s_t = 1 \mid \mathcal{F}_t^S)$.

Proposition 2.2 (Wonham Filter) *The belief π_t satisfies*

$$\mathrm{d}\pi_t = [\lambda_0(1 - \pi_t) - \lambda_1\pi_t - \lambda_J\pi_t(1 - \pi_t)]\mathrm{d}t + (1 - \pi_{t-})\mathrm{d}N_t \tag{4}$$

The three drift terms represent: transition into the event regime (λ_0), transition out of the event regime (λ_1), and Bayesian learning from the absence of jumps (λ_J). The learning term is nonlinear in π because the informational content of "no jump" depends on the current belief. If π is close to 1, the absence of a jump is highly informative against the event regime. If π is near 0, it is uninformative. When jumps occur, beliefs jump to $\pi = 1$ because this observation reveals with certainty that the event regime is active.

Between jumps, the steady-state belief $\bar{\pi}$ is the positive root of

$$\lambda_J\pi^2 - (\lambda_0 + \lambda_1 + \lambda_J)\pi + \lambda_0 = 0 \tag{5}$$

which reduces to $\bar{\pi} = \lambda_0/(\lambda_0 + \lambda_1)$ when $\lambda_J = 0$ (no learning from jump absence).

Remark 2.3 (Observation through Prices) The filter in equation (4) uses the counting process N_t as the observation, omitting the diffusion component of the full price path. When jumps are large relative to diffusion volatility—the empirically relevant case for earnings, where single-day moves of 5–20% are common against daily volatility of 1–3%—the simplified filter closely approximates the full solution (Liptser and Shiryaev 2001).

Lemma 2.4 (Belief Dynamics Between Jumps) *If no jumps occur, the between-jump ODE*

$$\frac{d\pi}{dt} = \lambda_0(1 - \pi) - \lambda_1\pi - \lambda_J\pi(1 - \pi)$$

is a Riccati equation whose solution is a logistic function converging to $\bar{\pi}$ at a rate that depends on the separation of the two roots of the associated quadratic. For the parameter values relevant to the application ($\lambda_J \gg \lambda_0, \lambda_1$), the convergence is rapid and qualitatively similar to exponential mean-reversion toward $\bar{\pi}$.

2.3. *Risk-Neutral Density and Mixture Representation*

Given belief π_t, the investor's perceived RND of S_T is a mixture:

$$f_{S_T}(x \mid \pi_t) \approx w_0(\pi_t) \cdot f^{(0)}(x) + w_-(\pi_t) \cdot f^{(-)}(x) + w_+(\pi_t) \cdot f^{(+)}(x) \tag{6}$$

where $w_0(\pi_t) = \exp(-\tilde{\pi}\lambda_J \tau)$ is the no-jump weight, $w_-(\pi_t) = (1 - w_0)(1 - p)$ is the down-jump weight, $w_+(\pi_t) = (1 - w_0)p$ is the up-jump weight, and $\tilde{\pi}$ is the time-averaged belief over $[t, T]$.

Each component distribution is log-normal: conditional on being in component $k \in \{0, -, +\}$, the log terminal price is Gaussian with $\log S_T \mid k \sim \mathcal{N}(\mu_k, v_k^2 \tau)$. Specifically, the central (no-jump) component has volatility v_0, while the jump components have volatility v_J (assumed equal for tractability). The mixture representation assumes at most one jump over $[t, T]$, a reasonable approximation for the short-dated options (3–13 days) in the sample. This Gaussian mixture structure is the key requirement for applying the Glasserman and Pirjol (2023) characterization of W-shaped implied volatility.

2.4. *Effective Central Volatility*

The key to W-shape generation is the effective volatility of the central component.

Definition 2.5 (Effective Central Volatility)

$$v_{\text{eff}}^2(\pi) = v_0^2 + \eta \cdot \pi(1 - \pi) \tag{7}$$

where v_0 is base diffusion volatility and $\eta > 0$ captures how belief uncertainty translates to perceived volatility.

The $\pi(1 - \pi)$ term arises from the law of total variance: $\text{Var}(R) = \mathbb{E}[\text{Var}(R \mid s)] + \text{Var}(\mathbb{E}[R \mid s])$, where the second term equals $\pi(1 - \pi) \cdot (\mu_1 - \mu_0)^2$ for regime-specific expected returns μ_0, μ_1. This cross-regime uncertainty is maximized at $\pi = 0.5$ and vanishes as beliefs become extreme, translating to higher perceived volatility for the central ATM region.

Remark 2.6 (Gaussian Approximation) Strictly, the regime-averaged central component is a two-normal sub-mixture: conditional on no jump, the return distribution differs across regimes because the event regime has a jump-compensated drift ($\mu_1 \neq \mu_0$) even when no jump occurs. Averaging over the uncertain regime with belief π gives a distribution with variance $v_0^2\tau + \pi(1-\pi)(\mu_1 - \mu_0)^2\tau^2$, which is approximated as Gaussian with inflated variance $v_{\text{eff}}^2\tau$. This approximation is accurate when the drift difference is small relative to diffusion volatility—i.e., $(\mu_1 - \mu_0)^2\tau \ll v_0^2$—which holds for the short maturities ($\tau < 0.04$) and estimated parameter values ($\hat{\eta} \approx 0.86$, $v_0 \approx 0.40$) in the sample. The qualitative comparative statics in the theorem are robust to this approximation.

Remark 2.7 (Interpretation of η) The parameter η is a reduced-form variance-rate coefficient estimated from the data. Its structural origin is the law of total variance: conditional on no jump, the expected log return over $[t, T]$ differs across regimes because the event regime carries a jump compensator in the drift even when no jump occurs. The cross-regime variance of conditional means scales as $\pi(1 - \pi) \cdot \delta_\mu^2 \tau$. δ_μ is the annualized drift difference between regimes, and τ is the time to maturity. For the short maturities in the sample ($\tau < 0.04$), this maturity dependence is weak,

and η is approximately constant across option maturities. η is larger when the drift difference between regimes is greater. These are firms where the event regime implies more extreme expected outcomes—which generate the cross-sectional prediction tested in Section 4, estimating η directly from the observed relationship between belief proxies and ATM implied volatility.

2.5. *Main Theorem*

ASSUMPTION 2.8 (Jump Volatility Exceeds Base Volatility) The volatility of jump components exceeds the base diffusion volatility: $v_J > v_0$. This is natural when jumps represent large discrete moves that dominate the diffusive component.

THEOREM 2.9 (W-Shape from Learning) *Consider the regime-switching model with bidirectional jumps and investor learning under Assumption 2.8. Let v_J denote the volatility of jump components (assumed equal for up and down). Define*

$$\pi_{crit} = \frac{1}{2} - \frac{1}{2}\sqrt{1 - \frac{4(v_J^2 - v_0^2)}{\eta}} \tag{8}$$

Then the implied volatility curve can exhibit a W-shape only if

$$\pi \in (\pi_{crit}, 1 - \pi_{crit}) \tag{9}$$

provided that $\eta > 4(v_J^2 - v_0^2)$.

The proof is given in Appendix A. Intuitively, W-shapes require $v_{\text{eff}} > v_J$, which reduces to a quadratic inequality in π whose solution is the stated interval.

Remark 2.10 (Immediate Consequences) W-shape intensity, measured by $v_{\text{eff}}^2 - v_J^2$, is maximized at $\pi = 1/2$ (maximum uncertainty). When $\pi \approx 0$ or $\pi \approx 1$, the effective central volatility falls below the wing volatility and the IV curve exhibits a standard U-shape.

PROPOSITION 2.11 (W-Shape Intensity) *Define the W-shape intensity as $I(\pi) = \eta \cdot \pi(1-\pi) - (v_J^2 - v_0^2)$. Then:*

(i) $I(\pi)$ is maximized at $\pi = 1/2$, with maximum value $\eta/4 - (v_J^2 - v_0^2)$.
(ii) $I(\pi)$ is strictly increasing on $(\pi_{crit}, 1/2)$ and strictly decreasing on $(1/2, 1 - \pi_{crit})$.

The proof follows from $I'(\pi) = \eta(1 - 2\pi)$ and is given in Appendix A.

PROPOSITION 2.12 (Implied–Realized Variance Gap) *Under the belief-dependent mixture model, the gap between option-implied variance and expected realized variance satisfies*

$$\sigma_{ATM}^2(\pi) - \mathbb{E}^{\mathbb{P}}[RV] \approx \eta \cdot \pi(1-\pi) + IVG_0 \tag{10}$$

where IVG_0 is the baseline gap. Consequently, W-shapes are associated with an elevated wedge between implied and realized variance.

The proof is given in Appendix A. The key insight is that ATM implied volatility reflects the total predictive variance of the log return, which by the law of total variance includes the cross-regime mean uncertainty $\eta \cdot \pi(1-\pi)$. Realized variance depends on the actual regime path and does not contain this term. The wedge is therefore largest at intermediate beliefs, connecting W-shapes to elevated option premia.

PROPOSITION 2.13 (Cross-Sectional Variation) *The critical belief threshold π_{crit} satisfies:*

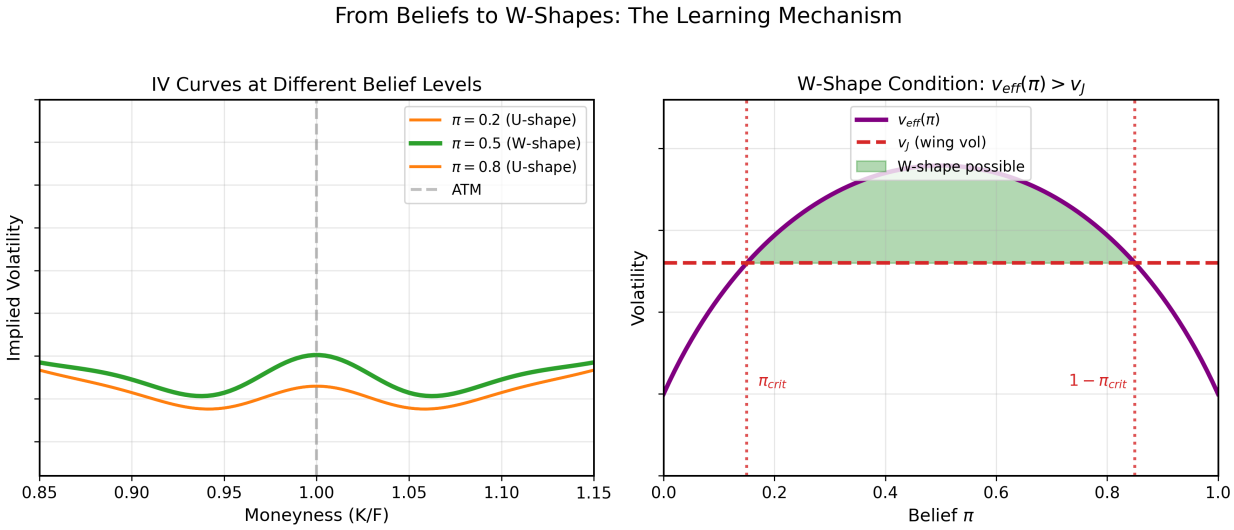

Figure 3. From Beliefs to W-Shapes: The Learning Mechanism. Left panel: IV curves for three belief levels—W-shapes emerge at intermediate π (green, $\pi = 0.5$) but not near the critical thresholds ($\pi = 0.2$ or $\pi = 0.8$, orange). Right panel: The threshold mechanism from Theorem 2.9. Effective central volatility $v_{\text{eff}}(\pi)$ (purple curve) exceeds wing volatility v_J (red dashed) only in the green shaded region $(\pi_{\text{crit}}, 1 - \pi_{\text{crit}})$, enabling W-shapes.

(i) $\partial\pi_{crit}/\partial\eta < 0$*: higher uncertainty premium* η *widens the W-shape region.*
(ii) $\partial\pi_{crit}/\partial v_J > 0$*: higher jump volatility* v_J *narrows the W-shape region.*

The proof follows by differentiating π_{crit} with respect to η and v_J and is given in Appendix A.

Remark 2.14 (Necessary vs. Sufficient Conditions) The condition $\pi \in (\pi_{\text{crit}}, 1 - \pi_{\text{crit}})$ is necessary but not sufficient for observable W-shapes. Even when $v_{\text{eff}} > v_J$, the jump components must have sufficient probability weight (controlled by $\tilde{\pi}\lambda_J\tau$) to create a detectable three-component structure.

2.6. *Dynamics Around Earnings*

Corollary 2.15 (Earnings Announcement Dynamics) *Consider an earnings announcement at time* T_E*. Before the announcement* ($t < T_E$)*, beliefs* π_t *rise as* λ_0 *increases, and if* π_t *enters* $(\pi_{crit}, 1 - \pi_{crit})$*, a W-shape appears and intensifies as* $\pi_t \to 1/2$*. After the announcement* ($t > T_E$)*, whether or not a jump occurred,* π_t *eventually exits the W-shape region and IV returns to a U-shape.*

3. Data

3.1. *Sample Construction*

The empirical study utilizes historical stock price data obtained from FirstRate Data. The sample contains 247 of the largest U.S. publicly traded firms whose common stock was included in the S&P 500 Index.[1]

These 247 firms represent six different industry categories including technology, financial services, healthcare, consumer products, energy, and industrial manufacturing. The period analyzed ranged

[1]The starting point is all 503 S&P 500 constituents as of February 2025. Each firm's option liquidity is checked by requiring at least 10 available option surfaces per quarter over at least two of three reference quarters (Q2-2020, Q1-2022, and Q1-2024), yielding 401 tickers. After filtering for data quality, DTE, moneyness, minimum strike prices, and other criteria, 247 stocks produce at least one valid snapshot.

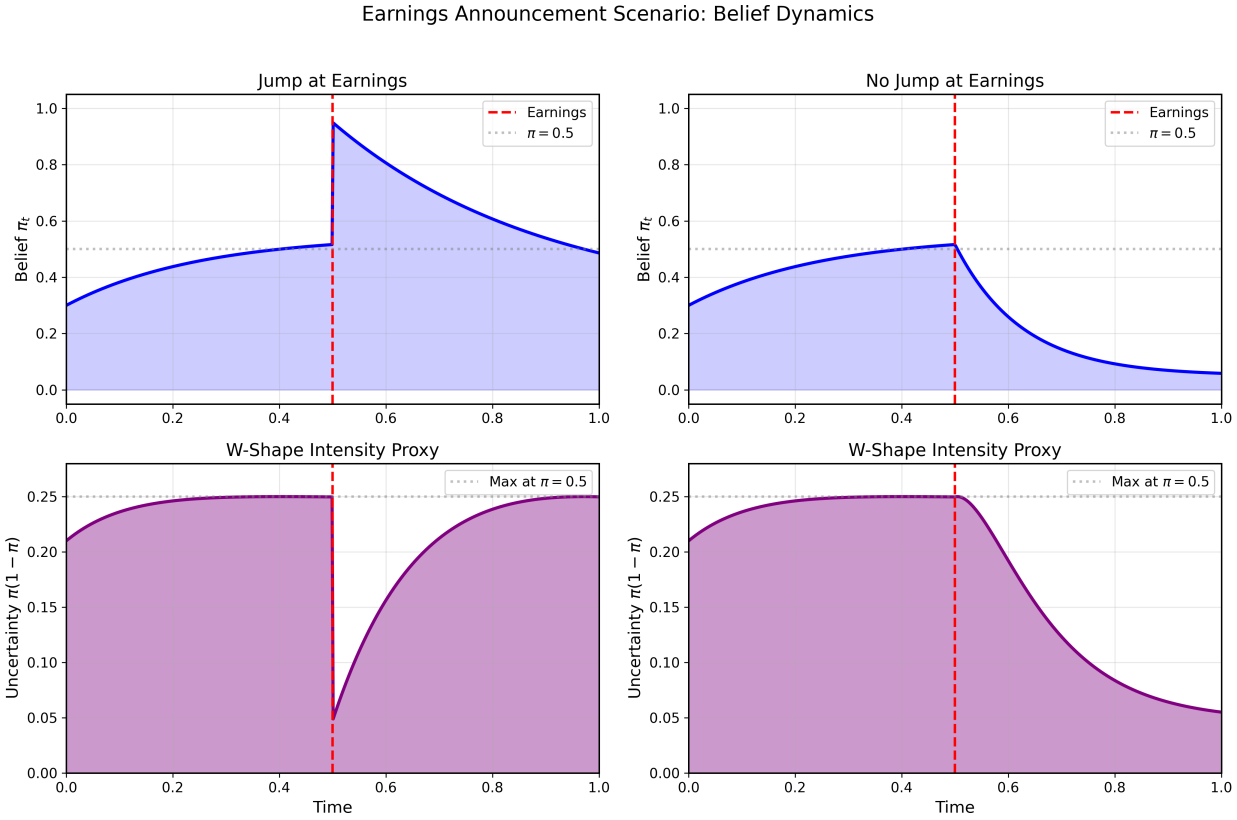

Figure 4. Belief Dynamics Around Earnings Announcements. Model-implied belief dynamics around an earnings announcement (dashed red line at $t = 0.5$). Left column: scenario where a jump occurs at earnings. Beliefs spike to $\pi = 1$, then decay as uncertainty resolves. Right column: scenario with no jump at earnings. Beliefs drop sharply as the absence of a jump reveals the calm regime. Bottom row shows $\pi(1 - \pi)$, the W-shape intensity proxy. In both scenarios, W-shape intensity peaks near the announcement when uncertainty is highest, then declines as beliefs move toward certainty.

from January 2017 to April 2025. For every trading day and ticker, the data provided includes option chains that include strike prices, expiration dates, bid/ask quotes, mid implied volatility, open interest, and volume. Spot prices were matched to options utilizing end-of-day closing prices.

The data utilized in the study focus on those options that expire within a relatively short time frame and thus provide insight into investor perceptions of event-driven risk. The options chosen expire between 3 and 13 days after the date of observation. This timeframe enables reasonable evaluations of event-driven risk. Any option with fewer than six viable strike prices has been removed, to ensure there is enough data to fit a smooth function to estimate the true surface. Three additional filters are applied at each snapshot level to eliminate surfaces that might include unreliable option pricing data: (i) mean open interest among all strike prices must exceed 50 contracts; (ii) the mean bid-ask IV spread must be less than 5 percentage points; and (iii) the maximum implied volatility among all strikes must be less than 150%.[1] Summary statistics for the sample are shown in Table 1 and a detailed description of how the sample was constructed is provided in Table 2.

Earnings announcement dates are collected from LSEG earnings calendars. An option is classified as "EAD-spanning" if its expiration date falls after the next scheduled earnings announcement; 188 of the 247 sample tickers have at least one such observation.

[1]These quality filters are applied during preprocessing to remove low-liquidity or mispriced surfaces that can produce apparent W-shapes driven entirely by data quality artifacts rather than genuine event risk. The subsequent spline fitting achieves a 99.4% success rate.

Table 1. Sample Summary Statistics

	Value
Sample Composition	
Total snapshots	122,689
Unique tickers	247
Sample period	Jan 2017–Apr 2025
Trading days covered	1,813
Option Characteristics	
Days to expiry: mean (median)	8.1 (9.0)
Days to expiry: range	3–13
Strikes per snapshot: mean (median)	13.2 (11.0)
Strikes per snapshot: range	6–25
Earnings Coverage	
Earnings events in sample	2,528
Tickers with EAD-spanning options	188
Snapshots spanning earnings	7,254 (5.9%)
Snapshots not spanning earnings	114,748 (94.1%)

Table 2. Sample Construction

Stage	Snapshots	Tickers	Filter applied
Option surfaces	122,689	247	DTE $\in [3, 13]$, $\geq$6 strikes, quality filters
Successful spline fits	122,002	247	Spline convergence
of which:			
EAD-spanning	7,254	188	Expiry after next earnings
Non-EAD	114,748	247	

Notes: The sample applies filters on days to expiry (3–13 calendar days), minimum strikes ($\geq$6), and snapshot-level quality screens (OI $\geq$ 50, IV spread $<$ 5pp, IV $<$ 150%). Spline fitting fails for 687 snapshots (0.6%) due to non-convergence or implausible RND modes.

3.2. *Concavity Detection*

Following the methodology in Alexiou *et al.* (2025), an IV curve is classified as concave (W-shaped) using a rule-based algorithm. First, for each snapshot, put and call implied volatilities are blended within ±2% of at-the-money moneyness, and out-of-the-money options are selected (puts for $K < S$, calls for $K \geq S$) to produce a single IV per moneyness level. A quintic smoothing spline is then fit to the IV curve as a function of moneyness $m = K/S$, minimizing

$$\sum_{i=1}^{n} (\mathrm{IV}_i - \hat{\mathrm{IV}}(m_i))^2 + \rho \int \left(\frac{d^3 \hat{\mathrm{IV}}}{dm^3} \right)^2 dm \tag{11}$$

where ρ is a smoothing parameter.[1] The procedure starts with $\rho = 0.01$ and iteratively increases the smoothing parameter, accepting the fit when the RMSE falls below 0.5 percentage points of IV and the fitted curve is everywhere positive.[2] Fits that produce negative implied volatilities or risk-

[1]The penalized objective is implemented via `scipy.interpolate.splrep`, which solves the equivalent constrained formulation where ρ controls the permitted sum of squared residuals.

[2]Alexiou *et al.* (2025) use a stricter RMSE target of 0.01%. This is relaxed to 0.5 percentage points because the data, with an average of 13 strikes per snapshot and heterogeneous liquidity across 247 tickers, requires more smoothing to avoid oscillatory

Table 3. Concavity Detection Results

	N	Rate
Overall (successful fits)		
Not concave	119,516	98.0%
Concave (W-shaped)	2,486	2.0%
By Earnings Proximity		
EAD-spanning snapshots	7,254	—
Concave	653	9.0%
Non-EAD snapshots	114,748	—
Concave	1,833	1.6%

neutral densities with more than two modes are excluded. Second, the second derivative $d^2\text{IV}/dm^2$ is computed numerically from the fitted spline on a fine grid of 1,001 equally-spaced moneyness points, and regions of negative curvature are identified. Third, an IV curve is classified as concave if it exhibits negative curvature ($d^2\text{IV}/dm^2 < 0$) over a continuous region spanning at least 3% of moneyness, with a stationary point (local maximum in IV) in the interior of the strike range, excluding the lowest and highest strikes. This last condition ensures that the detected concavity is not an artifact of extrapolation at the edges of the strike range. The spline fitting procedure achieves a 99.4% success rate (122,002 of 122,689 snapshots).

Table 3 reports concavity statistics for the sample.

3.3. *Additional Variables*

For each snapshot, a variety of other variables are also calculated. Realized volatility is the annualized standard deviation of the logarithm of returns during the lifetime of the option based on daily closing prices. This provides an ex-post estimate of actual price movement that can be compared to an ex ante estimate provided by implied volatility. ATM implied volatility represents the midpoint IV at the strike closest to the spot price, interpolating if necessary. Days to earnings is the number of calendar days from observation to the next scheduled earnings release. The curvature score is the greatest magnitude of a negative second derivative (absolute value) among the points classified as concave. This presents an intensity measure for the W-shapes, where higher scores indicate a more pronounced local maximum in the IV curve.

4. Empirical Analysis

4.1. *W-Shapes and Earnings Proximity*

The model predicts that the frequency of W-shaped patterns is greater around earnings announcement dates when investors' beliefs regarding the likelihood of an earnings surprise are high. As reported in Table 4, the proportion of options with concave IV curves is much larger among those options spanning an earnings announcement (9.0%) compared to those that do not (1.6%), with an odds ratio of 6.1.

fits. In practice, the median achieved RMSE is well below this threshold.

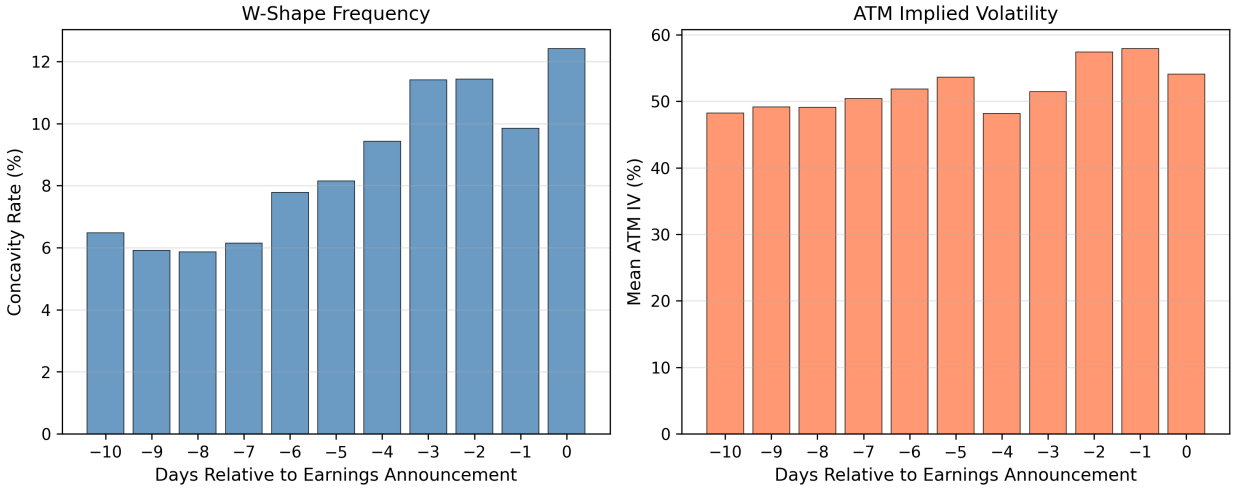

Figure 5. W-Shape Frequency and ATM IV Around Earnings Announcements. Left panel shows the concavity rate among EAD-spanning options by days relative to the earnings announcement (0 = announcement day). The rate increases monotonically as the event approaches. Right panel shows mean ATM implied volatility among EAD-spanning options on the same scale. Both build up as the event draws nearer, consistent with the theoretical prediction that W-shapes and elevated IV reflect event anticipation.

Table 4. Concavity Rate by Earnings Proximity

Category	Concavity Rate	N (Concave)	N (Total)
Non-EAD	1.6%	1,833	114,748
EAD-spanning	9.0%	653	7,254
Odds Ratio	6.09		
Chi-square	$\chi^2 = 1870.23$, $p < 10^{-300}$		

In addition to illustrating how the level of option price curvature varies depending on the proximity to earnings announcement dates, the data also illustrate how the degree of curvature changes over time in relation to those same events. As illustrated in Table 5, the proportion of options with concave IV curves declines smoothly over time as the distance to the event date increases. For example, while nearly 13 percent of options were observed to have concave IV curves during the last few days prior to their respective announcement dates, only roughly 6 percent of options exhibited such curvature at longer horizons. Furthermore, as shown in Figure 5, which illustrates the dynamic relationship between concavity of IV curves and both the timing of earnings announcement dates and the corresponding levels of ATM implied volatility, there is a smooth monotonic increase in the proportion of concave IV curves as the event date draws nearer. A linear probability model that includes both a measure of investor beliefs (proxied by $\hat{\pi} = 1/(1+d)$) and ticker fixed effects results in a statistically significant positive estimate ($\hat{\beta} = 0.064$, $p < 0.001$), confirming that the treatment effect is continuous rather than binary.

Table 5. Concavity Rate by Days to Earnings

Days to EAD	Concavity Rate	N
-3 to 0 (imminent)	12.4%	838
0 to 3	10.7%	2,637
3 to 7	7.6%	2,421
> 7	6.0%	1,358

Among concave curves, mean curvature is highest at 0–3 days before earnings and declines at longer horizons (Table B1 in the Appendix).

4.2. *Economic Significance*

A primary issue for practitioners is determining if W-shaped patterns can be interpreted as predictions of real event risk or simply as a pricing phenomenon. Table 6 compares the realized volatility to the implied volatility of all options spanning EAD by W-shape type.

Table 6. Economic Significance of W-Shaped Implied Volatility

	W-Shape	Normal	p-value
Panel A: Volatility comparison			
Realized volatility	30.5%	29.9%	0.344
Implied volatility (ATM)	69.8%	51.3%	< 0.001
IV premium (IV − RV)	39.3%	21.4%	< 0.001
IV ratio (IV / RV)	2.29×	1.72×	< 0.001
Panel B: Earnings-day returns			
Mean $\lvert R_{\text{EAD}}\rvert$	5.90%	4.30%	0.000***
Median $\lvert R_{\text{EAD}}\rvert$	4.17%	3.02%	0.000***
Fraction $\lvert R\rvert > 2\%$	77.0%	65.8%	
Fraction $\lvert R\rvert > 5\%$	45.0%	30.0%	
N (events)	278	714	
Panel C: Implementable straddle returns			
Mean return (mid entry)	2.5%	−0.3%	0.447
Mean return (ask entry)	1.1%	−1.8%	0.425
Median return (mid entry)	−23.3%	−20.0%	0.839
Win rate (mid entry)	41.0%	41.1%	

Notes: Panel A compares realized and implied volatility for EAD-spanning options. Panel B reports absolute earnings-day stock returns; each observation is an earnings event, with "W-Shape" meaning at least one EAD-spanning snapshot exhibited a concave IV curve. Panel C reports actual straddle P&L: buy ATM call + put at mid (or ask) price, hold to expiry. p-values from Welch t-test (means) and Mann-Whitney U (medians). N = 652 (W-shape), 6584 (normal) for Panels A and C; N = 278 (W-shape), 714 (normal) for Panel B.

As demonstrated in the table, realized volatility is essentially equivalent regardless of W-shape existence (30.5% vs. 29.9%). Conversely, there is a large difference between implied volatility based on W-shape existence (69.8% vs. 51.3%, $p < 0.001$), yielding a considerable IV premium gap.

Analysing returns directly during the earnings announcement day also confirms that W-shapes are indicative of real event risk. As shown in Panel B of Table 6, earnings announcements preceded by W-shapes generated substantially larger absolute returns (5.9% vs. 4.3%, $p < 0.001$), and 45% of W-shape events produced total return movements greater than 5% while only 30% of non-W-shape events experienced such large moves.

Panel C of Table 6 contains implementable ATM straddle returns based upon actual option prices. For every snapshot, a straddle is constructed by purchasing the call and put nearest to moneyness 1.0 at the midpoint (or ask) price and holding through expiration.[1] All W-shape and normal straddles exhibited negative median returns (−23.3% vs. −20.0%), indicating that both types of straddles suffer from a substantial negative variance risk premium around earnings. W-shape straddles yield marginally greater average returns (+2.5% vs. −0.3%); however, this difference is not statistically significant ($p = 0.447$). Table B2 in the Appendix presents further details including ask-entry returns and bid-ask spreads.

[1]Payoff at expiry is $\max(S_T - K_c, 0) + \max(K_p - S_T, 0)$, where S_T is the stock closing price on the expiration date, consistent with the PM settlement convention for standard U.S. equity options. When the expiration date falls on a weekend or holiday, the most recent available close is used.

The insignificant straddle return differential, along with significantly greater realised moves, demonstrates that the market appropriately priced in the increased event risk: the uncertainty premium η is included in option costs, thereby eliminating the possibility of exploiting W-shapes as mispriced opportunities. Section 4.5 evaluates the entire distribution of realised outcomes and verifies that the "hollowed-out centre" pattern—more probability density in the tail regions and less near zero—is consistent with the three-component mixture predicted by the model.

These findings provide refinement to the theoretical interpretation. The presence of W-shapes signifies that investors perceive high event risk and consequently price in the risk (via the IV premium gap) into option premiums. The substantial difference between implied and realised volatilities supports Proposition 2.12, which posits that the implied–realised variance wedge increases as investors' belief about event occurrence becomes intermediate. Also, because straddle returns for both W-shape and normal options are statistically indistinguishable, it follows that the market correctly prices the increased event risk into option premiums—the uncertainty premium $\eta \cdot \pi(1-\pi)$ is embedded in option costs rather than representing exploitable alpha.

4.3. *Structural Estimation of Filtered Beliefs*

A challenge with applying the theory for testing is that investor beliefs π cannot be directly observed. Until now, we used an ad hoc proxy for beliefs, being the number of trading days until the next earnings announcement. In this section, we will use a more structured method. We estimate a hidden Markov model (HMM), using the ATM implied volatility series for each stock. We can use these models to generate filtered estimates $\hat{\pi}$ of the probability that the market is in the high-volatility (event) regime.

This method may seem circular, as we're estimating $\hat{\pi}$ based upon ATM IV levels, and then using these levels to predict IV curvature. However, this circularity concern is mitigated due to the difference between IV *level* (a scalar) and IV *curvature* (a second derivative characteristic across strike prices). High ATM IVs typically result in conventional convex smiles; therefore, we do not automatically conclude concavity when we see high ATM IV levels.

We assume that our HMM has two underlying latent regimes—a calm regime, and an event regime—and that ATM IV follows state-dependent Gaussian distributions. More formally, let $z_t \in \{0, 1\}$ represent the latent state. The emission model is:

$$\text{ATM IV}_t \mid z_t = k \sim \mathcal{N}(\mu_k, \sigma_k^2), \quad k \in \{0, 1\}$$

where the mean IV for state 0 (the calm state) is less than that of state 1 (the event state) ($\mu_0 < \mu_1$). The transition matrix A captures how likely it is to remain in a particular state. The parameters $\{\mu_0, \mu_1, \sigma_0, \sigma_1, A\}$ are determined independently for each stock using the Baum-Welch (EM) algorithm. Each observation produces a $\hat{\pi}_t$—the posterior probability of being in the event regime at time t—based upon a forward-backward pass over the sequence of IV values leading up to time t. The filtered belief $\hat{\pi}$ is analogous to the Wonham-filtered belief in the theory.

Table 7. Hidden Markov Model: Filtered Beliefs and Predictive Power

Panel A: Filtered beliefs $\hat{\pi}$ by W-shape status

	W-Shape	Normal	Difference
Mean $\hat{\pi}$	0.847	0.431	+0.416
Mean $\hat{\pi}(1-\hat{\pi})$	0.013	0.023	−0.009

Panel B: Predictive power

Model	In-Sample AUC	Out-of-Sample AUC
Naive (spans_ead only)	0.604	0.647
$\hat{\pi}$ only	0.790	—
Full ($\hat{\pi}$ + uncertainty + EAD)	0.795	0.758

These results show that W-shapes are associated with high filtered beliefs: $\hat{\pi}$ averages 0.847 for W-shape observations versus 0.431 for normal curves. Given that π_{crit} is close to zero (see Section 4.7), the required condition for observing a W-shape, namely that $\hat{\pi}$ be in $(\pi_{\text{crit}}, 1-\pi_{\text{crit}})$, is easily satisfied for W-shape observations. That W-shapes cluster at higher levels of $\hat{\pi}$ rather than at $\pi = 0.5$ illustrates a composition effect. Both conditions need to be met simultaneously for an observation to be classified as a W-shape: first, that the event regime is active, and second, that there exists enough probability mass associated with the jump components (which depends on $\tilde{\pi}\lambda_J\tau$); both increase with $\hat{\pi}$. The intensity prediction—namely that $\pi(1-\pi)$ reaches its peak value when $\pi = 0.5$—applies to variation within the W-shape region for a given firm, not to the cross-sectional comparison in Panel A of Table 7.

As shown in Panel B of Table 7, predictions of W-shapes can be improved by using the filtered belief $\hat{\pi}$. When used alone, the in-sample AUC increases from 0.604 (naive) to 0.790 ($\hat{\pi}$ only), an improvement of 0.186. The out-of-sample AUC also increases from 0.647 (naive) to 0.758 (full model). Thus, structurally estimating beliefs adds predictive ability beyond simple event-based indicators.

To provide further evidence that the findings are not due to look-ahead bias when estimating the HMM, the means and variances of regimes derived from full-sample data and rolling 252-day windows are found to be virtually indistinguishable (correlation > 0.95 for filtered beliefs). This suggests that the HMM identifies regime characteristics that persist through time rather than simply fitting specific events. The out-of-sample AUC reported above is calculated using a temporal split such that prediction models are trained only on pre-test-period data.

Estimation of η. The primary model parameter η determines how Bayesian uncertainty manifests as increased volatility. Estimation of η follows directly from the relationship $v_{\text{eff}}^2(\pi) = v_0^2 + \eta \cdot \pi(1-\pi)$, where observable surrogates for π are built using the timing of earnings announcements relative to the current date. In order to avoid circularity—estimating π from IV and then regressing IV against π—all methods use calendar-time surrogates rather than IV-derived measures.

Table 8. Estimation of Uncertainty Premium η

Method	$\hat{\eta}$	$\hat{v}_0^2$
EAD contrast	1.215	0.1637
Days to EAD	0.632	0.2243
Imminent vs. distant	0.738	0.2809
Average	0.862	

Using three distinct methodologies yields estimates ranging from 0.64 to 1.21, with an average of $\hat{\eta} \approx 0.86$. Method A utilises the distinction between EAD-spanning and non-EAD observations,

assuming that $\pi \approx 0.5$ at times proximate to an earnings announcement (maximum uncertainty) and that $\pi \approx 0$ elsewhere, thereby producing an upper bound of $\hat{\eta} = 4(\bar{\sigma}^2_{\text{EAD}} - \bar{\sigma}^2_{\text{non-EAD}}) = 1.21$. This upper bound reflects the strong assumption that all EAD-spanning observations have $\pi = 0.5$. Method B fits $\sigma^2_{\text{ATM}} = v_0^2 + \eta \cdot \pi(1-\pi)$ on EAD-spanning observations using the days-to-earnings proxy $\hat{\pi} = 1/(1+d)$, yielding $\hat{\eta} = 0.64$.[1] Method C compares imminent ($d \leq 3$) to distant ($d > 7$) earnings observations, giving $\hat{\eta} = 0.74$. At maximum uncertainty ($\pi = 0.5$), this implies additional variance of $\eta/4 \approx 0.22$ and additional volatility of approximately 46%.

4.4. *Negative Controls and Alternative Mechanisms*

A key prediction of the model is that W-shapes should form around *firm-specific* events, which can have very good or very bad outcomes for an individual firm, resulting in a bimodal risk-neutral distribution. In contrast, *systematic* events, which cause every stock to move in the same direction, should not result in W-shapes.

Federal Open Market Committee (FOMC) announcements provide a natural negative control. Although FOMC decisions are scheduled, anticipated, and cause large moves in the market, they represent systematic rather than firm-specific risk. The interest rate decision causes all stocks to move similarly; therefore, no individual firm faces a bimodal outcome distribution.

Table 9. Concavity Rates: Earnings vs. FOMC

Event Type	Concavity Rate	N	Odds Ratio vs. Neither
Neither EAD nor FOMC	1.51%	88,626	1.00
FOMC only (no EAD)	1.89%	26,122	1.26
EAD only (no FOMC)	9.37%	4,642	6.74
Both EAD and FOMC	8.35%	2,612	5.94

The results confirm the prediction made by the model. Options spanning only FOMC announcements (without an accompanying earnings event) exhibit a concavity rate of 1.9%, comparable to the baseline rate for options spanning neither event (1.5%, odds ratio 1.26). In contrast, options spanning only earnings announcements exhibit significantly higher concavity rates (9.4%, odds ratio 6.7). When both events coincide, the earnings effect dominates (8.4%).

This observed behaviour is consistent with the learning model. Earnings create firm-specific uncertainty regarding whether the information announced will be favourable or unfavourable *for that firm*, generating the bimodal distribution that produces W-shapes. FOMC decisions, while uncertain, cause all firms to move in roughly the same direction—there is no single-name "up jump" versus "down jump" component. Since the W-shape mechanism requires a *single-name* bimodal distribution, systematic shocks cannot create the three-component mixture in any individual firm's risk-neutral density. Supporting this distinction, Jacobs *et al.* (2025) demonstrate that FOMC-related event risk manifests in option-implied *tail* moments (skewness and kurtosis) of the aggregate market, not in the IV curve shape of individual stocks—confirming that systematic and firm-specific event risk affect different parts of the risk-neutral distribution.

Investor disagreement. An alternative to the learning explanation is that W-shapes arise from *investor disagreement* rather than common uncertainty. Under the disagreement hypothesis, optimistic investors who expect positive earnings increase demand for out-of-the-money calls while pessimistic investors increase demand for out-of-the-money puts, inflating both wings of the IV curve and creating a W-shape regardless of any uncertainty about the regime. If this were true, the

[1]When $d = 0$, this proxy gives $\hat{\pi} = 1$, implying $\pi(1-\pi) = 0$. This is consistent with the model: when the event is imminent, the market believes with near certainty that the event regime is active ($\pi \to 1$) and therefore faces no regime uncertainty. The high IV observed at $d = 0$ reflects the jump component weights rather than the $\pi(1-\pi)$ term. The estimation of η relies on variation at $d \geq 1$.

cross-sectional predictor of W-shapes should be the degree of disagreement among investors rather than the level of uncertainty facing a representative agent.

This hypothesis is tested using analyst EPS forecast data from Refinitiv/LSEG. For each of the 687 earnings events in the sample (28 tickers with analyst coverage), consensus estimates snapshotted seven days before the announcement are collected and two variables are constructed. The *disagreement* proxy $\text{DISP}_{it} = \sigma(\text{EPS}_{it})/|\bar{\text{EPS}}_{it}|$ is the coefficient of variation of analyst forecasts, measuring how much analysts disagree with each other about forthcoming earnings.[1] The *uncertainty* proxy UNCERT_{it} is the rolling standard deviation of historical earnings surprises, measuring how difficult this firm's earnings are to predict for *all* analysts—a proxy for fundamental uncertainty rather than disagreement. If disagreement drives W-shapes, DISP_{it} should be a positive predictor of W-shape occurrence.

Although DISP_{it} and UNCERT_{it} are modestly correlated (Pearson $r = 0.13$), they are empirically distinguishable. Table 10 presents linear probability models for W-shape occurrence. Regardless of specification, DISP_{it} is either statistically insignificant or carries a *negative* coefficient—contradicting the disagreement prediction. In the horse-race specification (column 3), both DISP_{it} ($\hat{\beta} = -0.001$) and UNCERT_{it} ($\hat{\beta} = -0.001$) are statistically insignificant, and both effects are small relative to the dominant predictors: earnings proximity (Spans EAD) and ATM implied volatility level. With full controls and time fixed effects (column 4), neither analyst-based variable retains significance, consistent with IV level subsuming both effects.

Table 10. Analyst Disagreement and W-Shape Probability

	(1) DISP only	(2) UNCERT only	(3) Horse race	(4) + Controls
DISP_z	−0.0007		−0.0006	−0.0041
	(0.0022)		(0.0022)	(0.0031)
UNCERT_z		−0.0012	−0.0012	−0.0077*
		(0.0019)	(0.0018)	(0.0040)
Spans EAD	0.0983***	0.0984***	0.0984***	0.0561***
	(0.0213)	(0.0213)	(0.0214)	(0.0132)
ATM IV_z				0.0427***
				(0.0066)
N strikes				−0.0016***
				(0.0003)
Ticker FE	Yes	Yes	Yes	Yes
Year-quarter FE	No	No	No	Yes
R^2	0.059	0.059	0.059	0.105
N	39,749	39,749	39,749	39,749

Notes: Dependent variable is a binary indicator for W-shaped IV. Linear probability model with OLS. DISP is the coefficient of variation of analyst EPS forecasts (std/—mean—), measuring cross-analyst disagreement. UNCERT is the rolling standard deviation of past earnings surprises, measuring fundamental earnings unpredictability. Both are standardized to unit variance. Standard errors (in parentheses) clustered by ticker. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

These results indicate that W-shapes are not driven by analyst disagreement. DISP_{it} fails as a predictor—contrary to the disagreement hypothesis—and UNCERT_{it} is also statistically insignificant, suggesting that analyst-based measures of fundamental uncertainty are dominated by option-market-based proxies (ATM IV, earnings proximity) used in the preceding analysis. The lack of predictive power of DISP_{it}, combined with the FOMC negative control (ruling out systematic

[1]This measure is standard in the disagreement literature; see Diether *et al.* (2002).

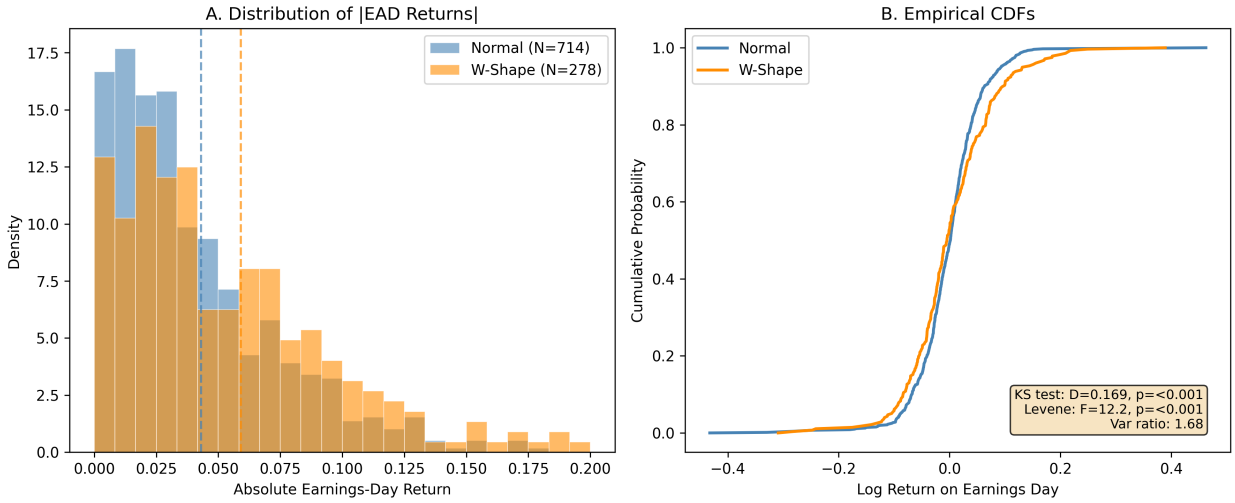

Figure 6. Bimodality of Realised Outcomes: W-Shape vs. Normal Events. Left panel: histogram of absolute earnings-day returns for events preceded by W-shaped IV (orange) vs. normal IV (blue). W-shape events exhibit heavier tails. Right panel: empirical CDFs of signed log returns. The Kolmogorov-Smirnov test rejects distributional equality ($p < 0.001$).

events) and the monotonic buildup pattern documented in Figure 5 (consistent with Bayesian belief updating), supports the representative-agent learning interpretation over the heterogeneous-beliefs alternative.

4.5. *Direct Tests of the Mixture Model*

The theoretical model predicts that the risk-neutral density underlying W-shaped IV curves is a three-component mixture, implying that realisations should be drawn from a distribution with heavier tails and a thinner centre than normal curves. This prediction is tested directly.

Bimodality of realised outcomes. If W-shapes represent a genuine three-component mixture in the RND—a central "no-jump" component flanked by "up-jump" and "down-jump" components—then earnings events preceded by W-shapes should exhibit more extreme and more dispersed realised returns. Figure 6 compares the distribution of absolute earnings-day returns conditional on W-shape status. Among W-shape events, 45% of earnings-day returns exceed 5% in absolute value, compared to 30% for non-W-shape events. Conversely, only 23% of W-shape events have returns below 2% in absolute value, compared to 34% for non-W-shape events. The Kolmogorov-Smirnov test rejects distributional equality ($p < 0.001$), and Levene's test confirms significantly higher variance for W-shape events ($p < 0.001$), with a variance ratio of 1.68. This "hollowed-out centre" pattern—more mass in the tails, less near zero—is the empirical signature of a bimodal or multi-modal distribution, consistent with the three-component mixture that generates the W-shape. Importantly, this departure from normality is concentrated in the *body* of the distribution rather than the tails: it is the depletion of probability mass near zero—where a Gaussian distribution places the most weight—that distinguishes W-shape events, not heavier extreme tails per se.

These distributional distinctions are consistent with the three-component mixture model; however, it should be noted that heavier tails in absolute returns do not uniquely identify a three-component structure—they could equally well arise from elevated variance or a one-sided jump mechanism. A distinct feature of the model is the "hollowed-out" centre (the depletion of probability mass around zero), which is difficult to obtain through an increased variance component alone. When combined with the monotonic pre-earnings buildup documented in Figure 5—and further supported by the FOMC negative control—collectively these pieces of evidence provide substantial support for the mixture-model mechanism. While no single piece of evidence is definitive, it is

the collective evidence from timing, cross-sectional, distributional, and negative-control tests that points toward the mixture-model mechanism.

4.6. *Cross-Sectional Variation*

Proposition 2.13 states that since firms with larger earnings effects (larger values of η) will have a wider belief region $(\pi_{\text{crit}}, 1 - \pi_{\text{crit}})$ in which W-shapes can occur, such firms should therefore experience W-shapes more often. To examine this, tickers are grouped into five quintiles based on their historical mean absolute earnings-day return (which serves as a proxy for η), and Table 11 reports the average W-shape frequency during EAD-spanning periods per quintile.

Table 11. Cross-Sectional Variation: Earnings Volatility and W-Shape Frequency

Quintile	$\lvert R_{\text{EAD}}\rvert$ range	Tickers	Mean $\lvert R_{\text{EAD}}\rvert$	Mean W-Shape Rate	N (EAD)
Q1	1.2–2.7%	31	2.3%	2.2%	1387
Q2	2.8–3.7%	30	3.2%	4.2%	1681
Q3	3.7–5.0%	30	4.3%	6.5%	1277
Q4	5.1–6.4%	30	5.8%	8.1%	1222
Q5	6.4–21.1%	31	9.1%	15.9%	1583
Spearman ρ	0.465 ($p < 0.001^{***}$), $N = 152$ tickers				

Notes: Tickers are sorted by mean absolute earnings-day return $|R_{\text{EAD}}|$ and grouped into quintiles. Mean W-Shape Rate is the equal-weighted average across tickers of the fraction of EAD-spanning snapshots classified as concave. Spearman rank correlation is computed at the ticker level. Figure 7 shows the underlying scatter.

The Spearman rank correlation between historical earnings volatility and W-shape occurrence frequency is $\rho = 0.465$ ($p < 0.001$) across 152 tickers with sufficient data. The relationship is illustrated in Figure 7. Firms experiencing large earnings changes tend to exhibit W-shapes more frequently during EAD-spanning events, whereas firms experiencing small earnings changes seldom exhibit W-shapes. The observed cross-sectional pattern lends considerable support to the comparative-static prediction of the model, and the large sample of 152 tickers substantially strengthens the statistical credibility of this test.

To test this prediction using all observations rather than ticker-level averages, Table B3 in the Appendix reports pooled panel regressions with an interaction between the EAD-spanning indicator and each ticker's historical earnings volatility (standardised). The main effect of spanning earnings remains significant across all specifications.

4.7. *Panel Regression and Falsification*

The comparisons above are unconditional. Table 12 provides linear probability regression results for concavity with progressively richer specifications.[1]

[1]The sample size in Table 12 (120,290) is slightly smaller than the full 122,002 successful fits because observations with missing control variables (ATM IV or strike count) are dropped across all columns to maintain a consistent sample.

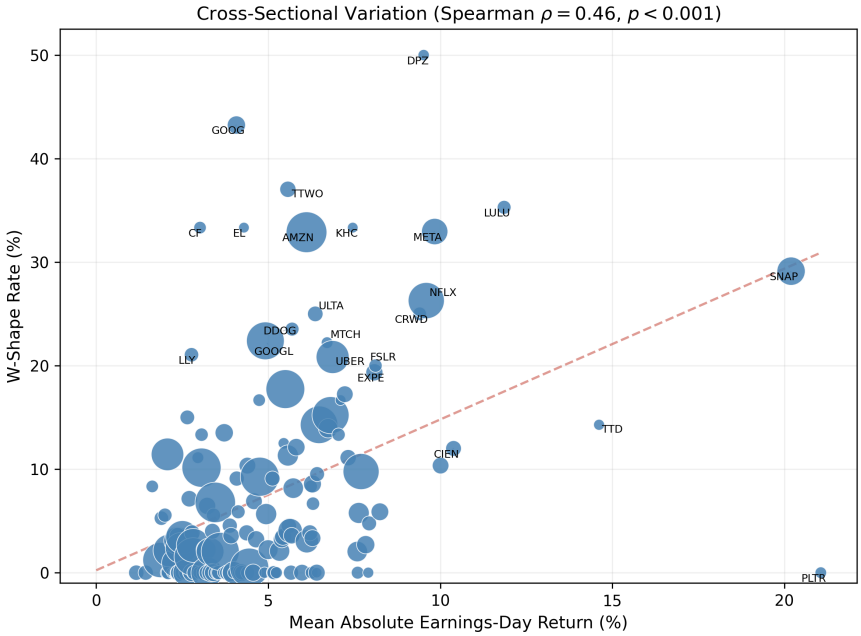

Figure 7. Cross-Sectional Variation: Earnings Volatility and W-Shape Frequency. Each point represents a ticker. The dashed line shows the fitted linear relationship. Spearman $\rho = 0.465$ ($p < 0.001$), $N = 152$ tickers.

Table 12. Panel Regressions: W-Shape Probability

	(1) Ticker FE	(2) + Time FE	(3) + Controls
Spans EAD	0.0767***	0.0748***	0.0354***
	(0.0120)	(0.0117)	(0.0072)
ATM IV			0.2459***
			(0.0223)
N strikes			−0.0015***
			(0.0002)
Ticker FE	Yes	Yes	Yes
Year-quarter FE	No	Yes	Yes
R^2	0.0375	0.0456	0.0842
N	120,290	120,290	120,290

Notes: Dependent variable is a binary indicator for W-shaped IV (concavity detected). Linear probability model with OLS. Standard errors (in parentheses) clustered by ticker. Column (1) includes ticker fixed effects only. Column (2) adds year-quarter fixed effects. Column (3) adds ATM implied volatility and number of strikes as controls. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Column (1) includes only ticker fixed effects: spanning an earnings announcement raises the probability of concavity by 7.7 percentage points ($p < 0.001$), closely matching the unconditional difference. Column (2) includes year-quarter fixed effects to capture potential time-dependent factors in markets—such as elevated volatility during COVID or changes in option market microstructure—that could independently affect concavity rates. The Spans EAD coefficient decreases by only 0.2 percentage points to 7.5 percentage points ($p < 0.001$), indicating that the earnings–concavity association is not driven by time-dependent aggregate conditions. Column (3) includes ATM implied volatility and strike count as additional controls. The Spans EAD coefficient attenuates further to 3.5 percentage points ($p < 0.001$) and ATM IV is strongly significant ($\hat{\beta} = 0.246$, $p < 0.001$). This attenuation is expected and informative: the model predicts that elevated beliefs (higher π) raise *both* ATM IV (via the $v_{\text{eff}}^2 = v_0^2 + \eta \cdot \pi(1-\pi)$ mechanism) and concavity probability. Therefore, ATM

IV acts as a mediator rather than a confounder of the earnings effect. The residual 3.5 percentage point effect of spanning earnings after controlling for IV level suggests that the shape of the curve (second-derivative concavity) contains information beyond the level.

With 247 ticker clusters, asymptotic cluster-robust standard errors are reliable. Table B4 in the Appendix reports a wild cluster bootstrap under the null $\beta_{\text{SpansEAD}} = 0$, using Rademacher weights at the ticker level with 999 replications. The bootstrap p-value is < 0.001, confirming that the result is robust to finite-cluster inference.

Falsification tests. Any event-study design may face concerns regarding whether the observed association between concavity and earnings reflects a confound—for example, if options written on particular expiry cycles mechanically possess different curve shapes. Table 13 reports two falsification tests.

Table 13. Placebo Tests: Shifted Earnings Dates

Shift (days)	N spanning	N concave	Concavity rate	Odds ratio
-30	4,457	23	0.005	0.24
-15	4,987	34	0.007	0.32
+0	7,254	653	0.090	6.09
+15	4,857	42	0.009	0.41
+30	4,088	24	0.006	0.28
Permutation test (1,000 iterations, shuffle within ticker)				
Actual OR			6.09	
Permutation mean			1.03	
Permutation 95th pctile			1.17	
Permutation p-value			< 0.001	

Notes: Upper panel shifts all earnings dates by the indicated number of days and recomputes the concavity–earnings odds ratio. The actual earnings dates (shift = 0, bold) produce the highest odds ratio. Lower panel reports results from a permutation test that randomly shuffles the EAD-spanning indicator within each ticker 1,000 times. The actual odds ratio exceeds all permuted values.

First, a permutation test randomly shuffles the EAD-spanning indicator within each ticker 1,000 times, maintaining the marginal distribution of both concavity and event proximity within each stock. All 1,000 permuted values fall below the observed odds ratio of 6.1 (permutation $p < 0.001$), with a mean of 1.03 and a 95th percentile of 1.17. This rules out the possibility that the association arises from the joint distribution of concavity and EAD-spanning status within tickers.

Second, shifting all earnings dates by ± 15 or ± 30 calendar days causes the odds ratio to collapse to 0.24–0.41, well below 1. The concavity–earnings association is sharply localised to actual earnings dates and cannot be explained by any seasonality or cyclicality present in the data. Together, these tests provide evidence that the relationship between W-shapes and earnings is genuine rather than an artifact of the option sample structure.

Interpreting the threshold condition. The estimated parameters yield an approximate value of zero for π_{crit}. Thus, the required condition of Theorem 2.9 is fulfilled in nearly all of the observations. However, W-shapes are relatively rare (2.0%). While the required condition allows for W-shape emergence under most positive beliefs, the required condition lacks discriminatory power in the data. Therefore, the threshold is necessary, but not sufficient. The jump components must also have enough probability mass to create a distinguishable three-component pattern, and meet the requirements imposed by the detection algorithm (Remark 2.14).

In this way, the model receives its empirical strength from the *comparative statics* (Propositions 2.11 and 2.13), and not from the threshold itself. W-shapes are more likely when π is large (especially near events), there is greater intensity at points of greatest uncertainty, and the width of the firm's belief area is larger for firms whose η values are larger. As such, these comparative-static predictions are exactly what the preceding tests validated.

5. Conclusion

The paper provides an economic framework to describe how investor learning generates W-shaped implied volatility (IV) curves. Investor learning leads to the creation of W-shapes endogenously by increasing belief uncertainty as it produces a higher effective volatility in the middle portion of the risk-neutral density; this will produce a three-component mixture that will map to a W-shaped IV curve. The main proposition outlines the necessary condition for the emergence of W-shapes, while subsequent propositions provide evidence that both the likelihood of occurrence and size of W-shape peaks increase monotonically with belief uncertainty. One of the largest contributions of the model is its ability to empirically test comparative statics. As such, the model predicts that W-shapes are most likely to occur near events, that they should decrease as one moves further away from events, and that they should completely disappear once the event has been resolved. Further, the model predicts that W-shapes should be more common than expected among companies whose stock price reacts more dramatically to events.

Empirical evidence shows that W-shapes cluster around earnings announcements, building up as the event approaches and collapsing sharply immediately after. Firms with more variable earnings reactions exhibit W-shapes more frequently, while systematic events such as FOMC announcements do not produce W-shapes (concavity rate of 1.9%, comparable to baseline). Realised returns surrounding W-shape events follow the bimodal distribution predicted by the three-component mixture. Panel regressions with ticker and time fixed effects, validated by wild cluster bootstrap and falsification tests, confirm that the association is robust and sharply localised to actual announcement dates.

Economically, the presence of W-shapes indicates the pricing of increased bidirectional jump risk. However, although implied volatility is much greater when W-shapes exist, straddle returns for W-shape and normal options do not differ statistically; this suggests the event risk priced into options is correct. The uncertainty premium exists in the cost of options rather than creating exploitable alpha opportunities. Therefore, the relationship of W-shapes to a realised–implied variance gap near events illustrates that the premium has been captured in option costs rather than being available for exploitation. The structural estimation of filtered beliefs based upon a hidden Markov model demonstrates that W-shapes occur primarily at high values of $\hat{\pi}$. This corresponds to the required condition under the theorem and provides additional evidence that the filtered belief offers predictive information regarding the likelihood of W-shapes occurring over and above simple measures of event proximity.

As the latent belief underlying the model is unobservable, the structural estimation uses calendar-time proxies instead of direct measures of investor uncertainty. The HMM recovers regime identification (i.e., whether the event regime is active); however, it cannot separate outcome uncertainty within the event regime. Developing proxy variables to identify these dimensions could create a more precise framework for structural testing. An alternative area of research is expanding the sample to include other firm-specific discrete events such as FDA approvals, merger votes, or product launches; the W-shape mechanism can then be tested across multiple types of firm-specific events to determine if it extends beyond earnings-related events.

Acknowledgments

The author thanks Robert Almgren for guidance on related work during the author's undergraduate thesis at Princeton. The author used Claude (Opus 4.6, Anthropic) to assist with proofreading and language refinement of the manuscript. The author reviewed all outputs and takes full responsibility for the content.

Disclosure statement

The author reports no competing interests.

Funding

This research received no specific grant from any funding agency.

Data availability statement

The option data that support the findings of this study are available from FirstRate Data. Restrictions apply to the availability of these data, which were used under licence for this study. Data are available from the author with the permission of FirstRate Data. Analyst forecast data were obtained from Refinitiv/LSEG under an institutional licence and are subject to the same restrictions.

References

Alexiou, L., Goyal, A., Kostakis, A. and Rompolis, L., Pricing Event Risk: Evidence from Concave Implied Volatility Curves. *Review of Finance*, 2025, **29**, 963–1007.

Bakshi, G., Cao, C. and Chen, Z., Empirical Performance of Alternative Option Pricing Models. *Journal of Finance*, 1997, **52**, 2003–2049.

Bates, D.S., The Crash of '87: Was It Expected? The Evidence from Options Markets. *Journal of Finance*, 1991, **46**, 1009–1044.

Bates, D.S., Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options. *Review of Financial Studies*, 1996, **9**, 69–107.

Bollen, N.P.B., Valuing Options in Regime-Switching Models. *Journal of Derivatives*, 1998, **6**, 38–49.

Breeden, D.T. and Litzenberger, R.H., Prices of State-Contingent Claims Implicit in Option Prices. *Journal of Business*, 1978, **51**, 621–651.

David, A., Fluctuating Confidence in Stock Markets: Implications for Returns and Volatility. *Journal of Financial and Quantitative Analysis*, 1997, **32**, 427–462.

Diether, K.B., Malloy, C.J. and Scherbina, A., Differences of Opinion and the Cross Section of Stock Returns. *Journal of Finance*, 2002, **57**, 2113–2141.

Duffie, D., Pan, J. and Singleton, K., Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica*, 2000, **68**, 1343–1376.

Ederington, L.H. and Lee, J.H., The Creation and Resolution of Market Uncertainty: The Impact of Information Releases on Implied Volatility. *Journal of Financial and Quantitative Analysis*, 1996, **31**, 513–539.

Glasserman, P. and Pirjol, D., W-Shaped Implied Volatility Curves and the Gaussian Mixture Model. *Quantitative Finance*, 2023, **23**, 557–577.

Hamilton, J.D., A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica*, 1989, **57**, 357–384.

Heston, S.L., A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 1993, **6**, 327–343.

Hu, G.X., Pan, J., Wang, J. and Zhu, H., Premium for Heightened Uncertainty: Explaining Pre-Announcement Market Returns. *Journal of Financial Economics*, 2022, **145**, 909–936.

Hull, J. and White, A., The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance*, 1987, **42**, 281–300.

Hull, J.C., *Options, Futures, and Other Derivatives*, 7th , 2009, Pearson Prentice Hall.

Jacobs, K., Ke, S. and Pan, X.N., Tail Risk Around FOMC Announcements. *Journal of Financial and Quantitative Analysis*, 2025, pp. 1–33.

Kou, S.G., A Jump-Diffusion Model for Option Pricing. *Management Science*, 2002, **48**, 1086–1101.

Liptser, R.S. and Shiryaev, A.N., *Statistics of Random Processes II: Applications*, 2nd , 2001, Springer.
Melick, W.R. and Thomas, C.P., Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil During the Gulf Crisis. *Journal of Financial and Quantitative Analysis*, 1997, **32**, 91–115.
Merton, R.C., Option Pricing When Underlying Stock Returns Are Discontinuous. *Journal of Financial Economics*, 1976, **3**, 125–144.
Naik, V., Option Valuation and Hedging Strategies with Jumps in the Volatility of Asset Returns. *Journal of Finance*, 1993, **48**, 1969–1984.
Pástor, v. and Veronesi, P., Stock Valuation and Learning About Profitability. *Journal of Finance*, 2003, **58**, 1749–1789.
Patell, J.M. and Wolfson, M.A., Anticipated Information Releases Reflected in Call Option Prices. *Journal of Accounting and Economics*, 1979, **1**, 117–140.
Rubinstein, M., Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Option Classes. *Journal of Finance*, 1985, **40**, 455–480.
Rubinstein, M., Implied Binomial Trees. *Journal of Finance*, 1994, **49**, 771–818.
Veronesi, P., Stock Market Overreaction to Bad News in Good Times: A Rational Expectations Equilibrium Model. *Review of Financial Studies*, 1999, **12**, 975–1007.
Wonham, W.M., Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering. *SIAM Journal on Control*, 1965, **2**, 347–369.

Appendix A: Proofs

A.1. *Proof of Proposition 2.2*

The Wonham filter for a two-state Markov chain observed through a point process follows from standard filtering theory (Wonham 1965, Liptser and Shiryaev 2001). A self-contained derivation for the specific setup is provided below.

Let $q_t^{(i)}$ for $i \in \{0, 1\}$ denote the unnormalized filter weights (Zakai equation). The observation process is the counting process N_t with intensity $\lambda_N(t) = s_t \cdot \lambda_J$ (jumps only occur in regime 1).

Step 1: Unnormalized filter between jumps. When no jump occurs over $[t, t + dt)$, the unnormalized weights evolve as:

$$dq^{(0)}/dt = -\lambda_0 q^{(0)} + \lambda_1 q^{(1)} \tag{A1}$$

$$dq^{(1)}/dt = \lambda_0 q^{(0)} - \lambda_1 q^{(1)} - \lambda_J q^{(1)} \tag{A2}$$

The first two terms in each equation reflect the Markov chain generator. The $-\lambda_J q^{(1)}$ term in (A2) penalizes regime 1 for the non-observation of jumps.

Step 1b: Normalized belief. Setting $S = q^{(0)} + q^{(1)}$ and $\pi = q^{(1)}/S$, the quotient rule gives:

$$\frac{d\pi}{dt} = \frac{1}{S}\frac{dq^{(1)}}{dt} - \pi \cdot \frac{1}{S}\frac{dS}{dt}$$

Since $dS/dt = -\lambda_J q^{(1)} = -\lambda_J \pi S$, substitution yields:

$$\frac{d\pi}{dt} = \lambda_0(1 - \pi) - \lambda_1 \pi - \lambda_J \pi(1 - \pi)$$

The learning term $-\lambda_J \pi(1 - \pi)$ is nonlinear: it captures Bayesian updating from the absence of jumps, with the $(1 - \pi)$ factor arising from the normalization.

Step 2: Update at jump times. When a jump occurs at time τ, Bayes' rule gives:

$$\pi_\tau = \frac{\mathbb{Q}(\text{jump} \mid s_{\tau-} = 1) \cdot \pi_{\tau-}}{\mathbb{Q}(\text{jump} \mid s_{\tau-} = 1) \cdot \pi_{\tau-} + \mathbb{Q}(\text{jump} \mid s_{\tau-} = 0) \cdot (1 - \pi_{\tau-})}$$

Since jumps only occur in regime 1, $\mathbb{Q}(\text{jump} \mid s = 0) = 0$ and $\mathbb{Q}(\text{jump} \mid s = 1) = \lambda_J\, dt > 0$, yielding:

$$\pi_\tau = \frac{\lambda_J \cdot \pi_{\tau-}}{\lambda_J \cdot \pi_{\tau-} + 0 \cdot (1 - \pi_{\tau-})} = 1$$

Observing a jump reveals with certainty that the event regime is active.

Step 3: Combine into SDE. Writing the dynamics in stochastic differential form:

$$d\pi_t = [\lambda_0(1 - \pi_t) - \lambda_1 \pi_t - \lambda_J \pi_t(1 - \pi_t)]\, dt + (1 - \pi_{t-})\, dN_t$$

The jump term $(1 - \pi_{t-})\, dN_t$ adds $(1 - \pi_{t-})$ to π at each jump, pushing beliefs from π_{t-} to 1. The between-jump ODE is a Riccati equation; its solution is a logistic function converging to the steady-state $\bar{\pi}$, the positive root of $\lambda_J \pi^2 - (\lambda_0 + \lambda_1 + \lambda_J)\pi + \lambda_0 = 0$. For the parameter values relevant to the application, the convergence is rapid and qualitatively similar to exponential mean-reversion. □

A.2. *Proof of Theorem 2.9*

The Glasserman and Pirjol (2023) characterization of W-shaped implied volatility is applied to the belief-dependent mixture model.

Step 1: Glasserman-Pirjol condition. For a Gaussian mixture with $N = 3$ components, Glasserman and Pirjol (2023) show that a necessary condition for the implied volatility curve to exhibit a W-shape (local maximum at ATM, valleys on either side) is that the central component volatility exceeds the wing component volatilities. In the notation of this paper: $v_{\text{eff}} > v_J$.

Step 2: Apply to belief-dependent volatility. From Definition 1, the effective central volatility is $v_{\text{eff}}^2(\pi) = v_0^2 + \eta \cdot \pi(1 - \pi)$. The W-shape condition becomes:

$$v_0^2 + \eta \cdot \pi(1 - \pi) > v_J^2$$

Rearranging:

$$\pi(1 - \pi) > \frac{v_J^2 - v_0^2}{\eta} \equiv \Delta$$

Step 3: Solve the quadratic inequality. Under Assumption 2.8, $v_J > v_0$ implies $\Delta > 0$. The function $g(\pi) = \pi(1 - \pi)$ is a downward parabola with maximum $1/4$ at $\pi = 1/2$. The inequality $g(\pi) > \Delta$ has solutions when $\Delta < 1/4$, i.e., when $\eta > 4(v_J^2 - v_0^2)$.

When this condition holds, $g(\pi) = \Delta$ has two roots:

$$\pi = \frac{1 \pm \sqrt{1 - 4\Delta}}{2}$$

The W-shape region is $\pi \in (\pi_{\text{crit}}, 1 - \pi_{\text{crit}})$ where $\pi_{\text{crit}} = (1 - \sqrt{1 - 4\Delta})/2$.

Step 4: Symmetry and maximum. The region is symmetric around $\pi = 1/2$ because $g(\pi) = g(1 - \pi)$. W-shape intensity $v_{\text{eff}}^2 - v_J^2 = \eta \cdot \pi(1 - \pi) - (v_J^2 - v_0^2)$ is maximized at $\pi = 1/2$. The condition $\pi \in (\pi_{\text{crit}}, 1 - \pi_{\text{crit}})$ is necessary: even when $v_{\text{eff}} > v_J$, the jump components must carry sufficient probability weight (controlled by $\tilde{\pi}\lambda_J \tau$) to create a detectable three-component mixture, so the condition is not sufficient. □

A.3. *Proof of Proposition 2.11*

$I(\pi) = \eta \cdot \pi(1-\pi) - (v_J^2 - v_0^2)$. The derivative is $I'(\pi) = \eta(1-2\pi)$, which is positive for $\pi < 1/2$ and negative for $\pi > 1/2$. Hence I is unimodal with maximum at $\pi = 1/2$, where $I(1/2) = \eta/4 - (v_J^2 - v_0^2)$. □

A.4. *Proof of Proposition 2.12*

For ATM options with short maturity τ, Black-Scholes implied variance satisfies $\sigma^2_{\text{ATM}} \approx \text{Var}^{\mathbb{Q}}(\log S_T)/\tau$. By the law of total variance over the uncertain regime:

$$\text{Var}^{\mathbb{Q}}(\log S_T) = \underbrace{\mathbb{E}^{\mathbb{Q}}[\text{Var}(\log S_T \mid s)]}_{\text{within-regime, linear in } \pi} + \underbrace{\text{Var}^{\mathbb{Q}}(\mathbb{E}[\log S_T \mid s])}_{\pi(1-\pi)\cdot\delta_\mu^2\tau^2}$$

The second term is the cross-regime mean uncertainty, which contributes $\eta \cdot \pi(1-\pi)$ to the implied variance rate. Under $\mathbb{P}$, the realized variance over $[t,T]$ depends on the actual regime path $\{s_u\}_{u\in[t,T]}$: conditional on the true regime, the quadratic variation does not contain $\pi(1-\pi)$ because the regime is deterministic given the path. Taking expectations, $\mathbb{E}^{\mathbb{P}}[\text{RV}]$ is a linear function of π. The gap $\sigma^2_{\text{ATM}} - \mathbb{E}^{\mathbb{P}}[\text{RV}]$ therefore contains $\eta \cdot \pi(1-\pi)$ plus a baseline component IVG_0. Since W-shapes require $\eta \cdot \pi(1-\pi) > v_J^2 - v_0^2 > 0$, they are associated with an elevated implied–realized gap. This connects to the "premium for heightened uncertainty" documented by Hu *et al.* (2022), who show that uncertainty about the magnitude of impending announcements generates a distinct risk premium reflected in the variance swap market. □

A.5. *Proof of Proposition 2.13*

Recall $\pi_{\text{crit}} = \frac{1}{2}(1 - \sqrt{1-4\Delta})$ where $\Delta = (v_J^2 - v_0^2)/\eta$. Since $\partial\Delta/\partial\eta = -(v_J^2 - v_0^2)/\eta^2 < 0$ and $\partial\pi_{\text{crit}}/\partial\Delta = 1/\sqrt{1-4\Delta} > 0$, the chain rule gives $\partial\pi_{\text{crit}}/\partial\eta < 0$. Similarly, $\partial\Delta/\partial v_J = 2v_J/\eta > 0$ yields $\partial\pi_{\text{crit}}/\partial v_J > 0$. □

Appendix B: Supplementary Tables

Table B1. Curvature Score by Days to EAD (Concave Curves Only)

Days to EAD	Mean Curvature	Std Dev	N
0–3 days	20.60	26.32	387
3–7 days	12.83	14.81	184
7–14 days	16.47	22.52	82

Table B2. ATM Straddle Returns by W-Shape Status (EAD-Spanning Options)

	W-Shape	Normal	p-value
Panel A: Mid-price entry			
Mean return	2.5%	−0.3%	0.447
Median return	−23.3%	−20.0%	0.839
Win rate (payoff > entry)	41.0%	41.1%	
N	652	6584	
Panel B: Ask-price entry (conservative)			
Mean return	1.1%	−1.8%	0.425
Median return	−24.1%	−21.2%	0.821
Win rate (payoff > entry)	40.2%	40.5%	
Panel C: Spread			
Avg bid-ask spread (% of mid)	2.9%	3.2%	

Notes: ATM straddle constructed by buying the call and put closest to moneyness = 1. Entry at mid-price (Panel A) or ask-price (Panel B). Payoff computed from stock close on expiration date. p-values from Welch t-test (means) and Mann-Whitney U test (medians). ***, **, * denote significance at 1%, 5%, 10% levels.

Table B3. Cross-Sectional Interaction: Earnings Volatility and W-Shape Probability

	(1)	(2)	(3)
Spans EAD	0.0762***	0.0775***	0.0374***
	(0.0121)	(0.0098)	(0.0070)
Spans EAD × Earnings Vol		0.0484***	0.0271***
		(0.0086)	(0.0067)
ATM IV			0.2432***
			(0.0213)
N strikes			−0.0013***
			(0.0002)
Ticker FE	Yes	Yes	Yes
Year-quarter FE	Yes	Yes	Yes
R^2	0.0456	0.0525	0.0878
N	121,855	121,855	121,855

Notes: Dependent variable is a binary indicator for W-shaped IV. Linear probability model. Earnings Vol is the ticker's mean absolute earnings-day return, standardized (mean zero, unit variance). The interaction term tests whether the earnings-concavity association is stronger for firms with historically larger earnings moves (Proposition 2.13). Standard errors clustered by ticker. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B4. Inference Comparison: Asymptotic vs. Wild Cluster Bootstrap

	Asymptotic	Wild Cluster Bootstrap
Coefficient (Spans EAD)	0.0761	
Standard error	0.0121	0.0166
p-value	0.0000^{***}	0.0000^{***}
95% null bounds	—	$[-0.0326, 0.0323]$
N clusters	247	247
Bootstrap replications	—	999

Notes: Comparison of asymptotic cluster-robust standard errors and wild cluster bootstrap inference for the main LPM specification (concavity on Spans EAD with ticker and year-quarter fixed effects). The bootstrap uses Rademacher weights at the ticker level under the null hypothesis $\beta_{\text{SpansEAD}} = 0$, with 999 replications. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Appendix C: Robustness Tests

Several robustness tests are conducted to validate the stability of the empirical findings.

C.1. *Subsample Stability*

Table C1 reports the key concavity–earnings relationship across subsamples.

Table C1. Subsample Stability of Concavity–Earnings Relationship

	N	Concavity rate	EAD rate	Non-EAD rate	Odds ratio	χ^2 p-value
Panel A: By time period						
Full sample	122,002	2.04%	9.00%	1.60%	6.09	< 0.001
2017–2020	61,260	2.12%	9.03%	1.75%	5.58	< 0.001
2021–2024	60,742	1.96%	8.98%	1.44%	6.74	< 0.001
Panel B: By ticker volatility tercile						
Low vol tercile	28,024	0.88%	3.24%	0.72%	4.63	< 0.001
Medium vol tercile	47,662	1.82%	9.72%	1.23%	8.67	< 0.001
High vol tercile	46,169	2.96%	12.93%	2.49%	5.82	< 0.001
Panel C: Excluding extreme curvature						
Baseline	122,002	2.04%	9.00%	1.60%	6.09	< 0.001
Excl. top 5% curvature	122,002	1.94%	8.84%	1.50%	6.37	< 0.001

Notes: Panel A splits the sample at year-end 2020. Panel B classifies tickers into terciles by median realized volatility. Panel C excludes the 125 observations with curvature scores above the 95th percentile (89.5), which may reflect spline fitting artifacts. Odds ratio is the ratio of EAD-to-non-EAD odds of concavity. χ^2 p-value tests independence between spanning status and concavity.

Time periods (Panel A). Splitting the sample at year-end 2020, the EAD concavity rate is stable across periods (9.0% in both halves). The overall concavity rate is similar across periods (2.1% vs. 2.0%), with similar odds ratios (5.6 vs. 6.7). Both are highly significant, affirming that the earnings effect is not isolated to a single time period.

Ticker volatility (Panel B). Grouping tickers into terciles where tickers in the same tercile have similar median realized volatility, the concavity–earnings association holds across all groups. High-volatility tickers have the highest overall concavity rate (3.0%), but comparable odds ratio (5.8) since their baseline non-EAD rate is also high. The low volatility tickers have lower overall rates and an odds ratio of 4.6. The χ^2s are significant for every tercile.

Excluding extreme curvature (Panel C). Excluding 125 observations with curvature scores above the 95th percentile (curvature > 89.5), lowers the overall rate from 2.04% to 1.94%, and raises odds ratio marginally from 6.1 to 6.4. EAD rate is roughly unchanged at 8.8%.

C.2. *Out-of-Sample Validation*

To verify that the concavity–earnings relationship is not an artifact of in-sample overfitting, the data is split at December 31, 2021 into a training period (2017–2021; 85,903 observations) and a test period (2022–2025; 36,099 observations). Table C2 reports key statistics computed independently in each period.

Table C2. Out-of-Sample Validation: Key Statistics by Period

	Train (2017–2021)	Test (2022–2025)	p-value (test)
Panel A: Concavity rates and odds ratio			
EAD-spanning concavity rate	8.4%	10.1%	
Non-EAD concavity rate	1.7%	1.3%	
Odds ratio (EAD vs non-EAD)	5.31	8.23	< 0.001
Panel B: LPM with ticker FE (clustered SE)			
β_{SpansEAD}	0.0366^{***}	0.0596^{***}	< 0.001
	(0.0095)	(0.0134)	

Notes: Train period: 2017–2021. Test period: 2022–2025. Odds ratios compare concavity rates between EAD-spanning and non-EAD options via Fisher's exact test. LPM (linear probability model) includes ticker fixed effects and ATM IV as control; standard errors clustered by ticker. ***, **, * denote significance at 1%, 5%, 10% levels.

The results are reassuring. The odds ratio for EAD-spanning concavity actually *increases* from 5.31 in the training period to 8.23 in the test period, driven by a decline in the non-EAD baseline rate (1.7% to 1.3%) while the EAD rate remains stable (8.4% to 10.1%). A linear probability model with ticker fixed effects, ATM IV control, and standard errors clustered by ticker yields qualitatively similar coefficients across periods ($\beta = 0.037$ in-sample vs. $\beta = 0.060$ out-of-sample), both significant at the 0.1% level. The concavity–earnings association is thus temporally stable and, if anything, strengthens in the more recent period.

C.3. *Alternative Concavity Detection Methods*

The baseline uses quintic spline fitting with adaptive smoothing parameter ρ following Alexiou *et al.* (2025). To assess sensitivity to the detection methodology, Table C3 reports results using progressively stricter curvature magnitude thresholds.

Table C3. Sensitivity to Curvature Magnitude Cutoffs

Curvature threshold	Concavity rate	EAD rate	Non-EAD rate	Odds ratio	N concave
> 0.4 (P0)	2.04%	9.00%	1.60%	6.10	2,485
> 4.3 (P10)	1.83%	7.54%	1.47%	5.46	2,237
> 8.1 (P25)	1.53%	5.56%	1.27%	4.56	1,864
> 16.0 (P50)	1.02%	3.10%	0.89%	3.58	1,243
> 31.8 (P75)	0.51%	1.27%	0.46%	2.77	622

Notes: Each row applies a stricter curvature threshold, counting a curve as concave only if its curvature score exceeds the stated value. Thresholds correspond to percentiles (P0, P10, etc.) of the curvature distribution among detected W-shapes. The baseline (P0, threshold > 0) reproduces the main results. Higher thresholds reduce the number of detected W-shapes but preserve the qualitative EAD–concavity relationship.

Table C3 shows that the odds ratio is robust to detection stringency. Requiring curvature above the 25th percentile reduces the number of detected W-shapes but preserves a large odds ratio. Even at the median threshold, the odds ratio remains substantially above 1. The EAD–concavity relationship is thus qualitatively robust to detection stringency, though the effect size attenuates as fewer marginal W-shapes are included.

C.4. *Detection Parameter Sensitivity*

The concavity detection algorithm depends on two key preprocessing parameters: the maximum number of strikes per surface ("strike cap") and the moneyness range filter. To verify that results are not driven by specific parameter choices, the detection pipeline is re-run on a stratified pilot sample of 1,000 snapshots (500 EAD-spanning, 500 non-EAD) using the full raw option chain data before preprocessing caps are applied.

Table C4 reports concavity detection rates and odds ratios across parameter values. Panel A varies the strike cap while holding the moneyness range fixed at $[0.85, 1.15]$. Panel B varies the moneyness range while holding the strike cap at 25.

Table C4. Detection Parameter Sensitivity

	Strikes used	EAD rate	Non-EAD rate	Odds ratio	RMSE (%)	Success (%)
Panel A: Strike cap (moneyness range $= [0.85, 1.15]$*)*						
10	10.0	0.084	0.010	8.9	0.117	97.8
15	14.9	0.070	0.008	9.1	0.167	97.0
20	19.4	0.078	0.009	9.5	0.195	91.8
25	**23.1**	**0.068**	**0.005**	**15.2**	**0.216**	**86.5**
30	25.6	0.065	0.005	14.1	0.227	83.8
50	28.2	0.068	0.000	∞	0.233	80.3
Uncapped	29.7	0.064	0.000	∞	0.234	79.5
Panel B: Moneyness range (strike cap $= 25$*)*						
$[0.80, 1.20]$	24.0	0.065	0.013	5.4	0.240	80.2
$[0.82, 1.18]$	23.7	0.064	0.013	5.3	0.230	82.0
$[\mathbf{0.85, 1.15}]$	**23.1**	**0.068**	**0.005**	**15.2**	**0.216**	**86.5**
$[0.88, 1.12]$	21.6	0.074	0.013	6.0	0.196	93.1
$[0.90, 1.10]$	19.9	0.081	0.017	5.1	0.176	95.1

Notes: Detection parameter sensitivity from a stratified pilot sample of 1,000 snapshots (500 EAD-spanning, 500 non-EAD). Each row re-runs the full detection pipeline (spline fitting, RND extraction, concavity detection) on raw option chain data with the indicated parameter value. Baseline parameters in bold. "Strikes used" is the mean number of strikes entering the spline fit. "RMSE" is the mean spline fit root-mean-squared error. "Success" is the fraction of snapshots where the spline fit converges and passes all validation checks. The odds ratio is computed as (EAD rate/(1 − EAD rate))/(non-EAD rate/(1 − non-EAD rate)).

The results confirm that the baseline parameters are well-calibrated. In Panel A, the EAD concavity rate is stable at 6–8% across all strike cap levels, while the non-EAD rate generally declines from 1.0% (cap = 10) to 0% (cap $\geq$ 50). The odds ratio peaks near the baseline cap of 25. Notably, even with uncapped strikes, the average number of strikes used is only 30 (vs. 23 at the baseline) because the moneyness range—not the cap—is the binding constraint.

In Panel B, the moneyness range has a stronger effect on discriminating power. The odds ratio peaks sharply at the baseline $[0.85, 1.15]$ (OR = 15.2), where the non-EAD false positive rate reaches its minimum (0.5%). Wider ranges admit noisy deep out-of-the-money options that create spurious concavity; narrower ranges miss the W-shape's characteristic wing valleys. The spline success rate also varies substantially—from 80% at the widest range to 95% at the narrowest—reflecting the increased difficulty of fitting noisy wing data.

Across all 12 specifications in both panels, the odds ratio exceeds 5.1 and the EAD concavity rate remains in the range 6–8%. The qualitative finding—that W-shapes are substantially more likely near earnings—is robust to preprocessing choices, even though the precise magnitude of discrimination varies.

C.5. *Additional Robustness*

The belief proxy π is based on days-to-earnings. Using a simple spans-EAD binary indicator produces similar directional findings, though the continuous proxy captures more variation. Exponential decay specifications $\pi = e^{-\tau/\kappa}$ with various κ values produce similar η estimates within estimation uncertainty. The key economic predictions are also robust: using 5-day, 10-day, or 30-day realized volatility windows yields qualitatively similar results, and both additive ($IV - RV$) and multiplicative (IV/RV) premium measures show W-shapes associated with higher premia.