Working Paper: Modeling the Pricing Impact of Airline Headlines: A Structured and Archetype-Driven Approach

Anirudh Krovi*

July 11, 2025

Abstract

This paper develops a cognitively structured and behaviorally interpretable framework for modeling the pricing impact of financial news, using Delta Airlines (DAL) as a case study. We begin by benchmarking FinBERT, a black-box sentiment model, and highlight its limitations—particularly in contexts where linguistic sentiment diverges from economic interpretation, such as safety-related headlines. To address this, we propose a modular, interpretable scoring model based on four headline attributes: actor clarity, lever relevance, magnitude, and novelty. We then extend this structure by introducing investor archetypes—Anxious, Patient, and Momentum-Seeking—each defined by a distinct temporal response kernel. By convolving structured scores with archetype-specific kernels, we capture heterogeneity in investor reactions across themes. Empirically, we find that archetypes with fast or deliberative response profiles (Anxious and Patient) dominate pricing impact for DAL, while momentum-based behavior plays a limited role. Our results demonstrate that understanding "who reacts, when, and how" provides explanatory depth beyond sentiment alone, offering a flexible alternative to black-box NLP models.

1 Introduction

In prior work[1], we developed a cognitively-informed scoring framework for financial news headlines and demonstrated its effectiveness on Apple Inc. (AAPL), a liquid, information-rich stock. There, we showed that a structured approach—grounded in actor clarity, lever relevance, magnitude, and novelty—outperformed generic sentiment models like FinBERT, particularly for nuanced themes such as supply chains and macroeconomic policy.

This paper extends that framework in two key directions. First, we test its robustness on a very different firm: Delta Airlines (DAL). Compared to AAPL, DAL presents distinct modeling challenges—its headline corpus is smaller, its language more operational, and its price drivers are heavily shaped by macro variables such as fuel costs and safety incidents. These features make it an ideal testbed for evaluating whether cognitively structured signals retain explanatory power in less tech-dense domains.

Second, and more importantly, we augment our scoring framework by embedding a behavioral dimension: we hypothesize that different types of investors process news with distinct temporal profiles. While foundational models in behavioral finance (Daniel et al., 1998; Barberis et al., 1998)

*PhD, Northwestern University; MBA, NYU Stern; formerly at McKinsey & Company; email: anirudh.krovi@stern.nyu.edu

[1]See Krovi (2025), "Cognitively Structured Scoring of Financial Headlines: A Case Study on AAPL" (working paper).

explain mispricing through stylized investor biases like overreaction or conservatism, they typically operate at the aggregate level. In contrast, we introduce a set of *investor archetypes*—Anxious, Patient, and Momentum-Seeking—each associated with a time-dependent kernel that redistributes a headline's signal over a short horizon. This allows us to simulate heterogeneity in news digestion and connect pricing behavior to behavioral traits like fear, deliberation, or trend-following.

We organize our empirical analysis around three dominant themes in DAL headlines: (i) fuel and oil prices, (ii) safety incidents, and (iii) company-level operational updates. For each theme, we construct daily pricing signals by convolving our structured scores with archetype-specific behavioral kernels, and regress excess returns on these time series. We find that behaviorally *Anxious* and *Patient* responses dominate, especially for ambiguous or emotionally salient news such as safety incidents or commodity shocks. In contrast, Momentum-driven responses appear weak or absent—suggesting that short-term trend-following may not play a strong role in DAL's news-based pricing.

Together, these results advance the core argument of this paper: that the value of news for asset pricing lies not only in its content, but in understanding *who* reacts, *when*, and *how*. Our combined framework—structured scoring plus behavioral kernels—offers a transparent, flexible alternative to black-box NLP models, and points toward richer models of investor cognition in financial markets.

The remainder of the paper is organized as follows. Section 2 presents two headline scoring models: a FinBERT-based baseline and our cognitively structured alternative. Section 3 embeds behavioral heterogeneity by introducing archetype-specific temporal kernels that simulate how different investor types respond to news over time. Section 4 compares the predictive performance of the two scoring models, while Section 5 explores theme-specific patterns in investor behavior using the archetype framework. Section 6 concludes with implications and outlines potential extensions for scaling this approach.

2 Methodology: Headline Scoring Models

2.1 FinBERT Scoring

Our baseline model uses FinBERT, a domain-specific transformer model pretrained on financial texts, to compute sentiment scores for headlines. This follows the same approach as our prior analysis on AAPL,[2] where we treated FinBERT as a benchmark for off-the-shelf sentiment extraction.

We collected over 250 headlines related to Delta Airlines (DAL) by scraping Google Search results between January 1, 2019 and January 1, 2024[3]. These headlines were grouped into three key themes:

- **Company news** – headlines pertaining to earnings, partnerships, management, or operations
- **Fuel and oil prices** – headlines using keywords like "jet fuel prices," "oil prices," "crude prices," "Brent prices," "WTI performance," and "NYMEX performance"
- **Safety** – headlines involving terms such as "crash," "plane crash," "emergency landing," and other aviation incident indicators

[2]See Krovi (2025), "Cognitively Structured Scoring of Financial Headlines: A Case Study on AAPL" (working paper).

[3]While our primary focus is on theme and actor-level content, we note that headlines originate from a mix of financial news sources (e.g., CNBC, Reuters, WSJ). Differences in tone, depth, and intended audience may introduce subtle biases in how themes are framed. Controlling for such variation remains an open direction for refining the scoring model.

To measure price response, we used daily excess returns defined as DAL returns minus S&P 500 returns. This choice differs from our AAPL setup: while AAPL is a large index component and often moves with the market, DAL's idiosyncratic news effects are more likely to appear in its excess returns.

For each headline, we used FinBERT to extract positive and negative sentiment probabilities. We computed a simple sentiment score as:

$$\text{FinBERT Score} = \text{Positive Probability} - \text{Negative Probability}$$

Given the high volatility and occasional extreme values of this raw score, we applied a log transformation to stabilize its distribution. Importantly, we preserved the sign of the original score to retain sentiment directionality:

$$\text{Log FinBERT Score} = \log(|\text{FinBERT Score}| + \epsilon) \cdot \text{sign}(\text{FinBERT Score})$$

where ϵ is a small constant added to avoid singularities at zero (we use $\epsilon = 10^{-6}$). This transformation allows us to smooth out the influence of highly uncertain scores while preserving their polarity. We then regressed these log-transformed FinBERT scores against DAL's excess returns at t, $t+1$, and $t+2$ horizons.

As we show in Section 4, FinBERT performs reasonably for some themes (e.g., fuel prices) but fails to capture meaningful signals for safety-related headlines—an area where context and actor clarity are particularly crucial.

2.2 Structured Headline Scoring

To address the limitations of black-box sentiment models like FinBERT, we apply a structured, interpretable scoring approach first developed in our AAPL case study.[4] This method assigns a composite score to each headline based on four distinct dimensions:

- **Actor clarity** – Is the main economic actor (e.g., Delta itself, regulators, competitors) clearly identified?
- **Lever relevance** – Does the headline pertain to a financially meaningful driver of performance (e.g., fuel prices, demand, fleet management)?
- **Magnitude** – Is the scale of the event or change discernible or substantial?
- **Novelty** – Is the headline introducing new information, or reiterating known developments?

Each dimension is scored on a discrete scale reflecting both its presence and quality:

- **Actor and Novelty:** scored from 0.5 to 2.0 in 0.25 increments
- **Lever and Magnitude:** scored from 0.5 to 1.0 in 0.25 increments

The scoring system also explicitly tracks the directionality of the headline's implications. For example, in a headline about "fuel prices," the *lever* receives a neutral sentiment, while the *magnitude* ("falling" or "rising") determines the direction. A fall in jet fuel prices thus receives a positive overall effect (since falling cost is good for DAL), while a rise would yield a negative effect.

To ensure robustness in headline coverage, we augment each dictionary programmatically using natural language processing tools. This augmentation includes:

[4] See Krovi (2025), "Cognitively Structured Scoring of Financial Headlines: A Case Study on AAPL" (working paper).

- Tense variations (e.g., “increase” vs. “increased”),
- Singular/plural forms (e.g., “deal” vs. “deals”),
- Synonym families (e.g., “surge,” “jump,” “spike”).

This ensures that our dictionaries are semantically broad and linguistically flexible, while still grounded in curated domain expertise.

The four components are aggregated as follows:

- **Actor:** aggregated across entities via a product of individual scores
- **Novelty:** max score across redundant headlines or time-proximate stories
- **Lever and Magnitude:** aggregated via a normalized softmax to emphasize stronger drivers while preserving numerical stability

The final score is then computed as the product of all four dimensions:

$$\text{Final Score} = \text{Actor} \times \text{Novelty} \times \text{Lever}_{\text{softmax}} \times \text{Magnitude}_{\text{softmax}}$$

As in our FinBERT analysis, we apply a signed log transform to stabilize the distribution while retaining sentiment polarity:

$$\text{Log Final Score} = \log(|\text{Final Score}| + \epsilon) \cdot \text{sign}(\text{Final Score})$$

where $\epsilon = 10^{-6}$.

To illustrate how this model diverges from FinBERT, consider the following example:

“Jet-fuel prices plummet as coronavirus feeds travel fears” (Barron’s, March 6, 2020)

This headline was tagged as **Fuel and Oil Prices**, and was labeled *Unexpected* and *Positive* from Delta’s perspective, as falling jet fuel prices lower operational costs. The corresponding market data and scores are summarized in Table 1.

DAL Return	1.96%
S&P 500 Return	-1.71%
Excess Return	+3.66%
FinBERT Sentiment Score	-0.99999
Actor Score	1.5
Lever Score	0.226
Magnitude Score	0.762
Novelty Score	0.0
Final Structured Score	0.129

Table 1: Example headline and scoring comparison. FinBERT interprets the headline as strongly negative due to linguistic cues, while our structured model captures its economically positive implications for DAL.

As the example shows, FinBERT assigns a strongly negative sentiment score—likely due to the presence of fear-related language (“feeds travel fears”)—despite the headline containing favorable economic news for an airline. Our structured model correctly isolates the relevant lever (fuel prices) and recognizes the magnitude (“plummet”) as positive from Delta’s cost perspective.

For each theme—**Company news**, **Fuel and Oil Prices**, and **Safety**—we regress excess returns on the log of the structured final score. As we will show in Section 4, our model consistently outperforms FinBERT for themes with clearer economic levers. However, it also struggles with safety-related headlines—motivating the need for a richer, behaviorally-informed model of investor reactions, which we present next.

3 Archetypes and Behavioral Kernels

While foundational models in behavioral finance—such as those by Daniel et al. (1998) and Barberis et al. (1998)—emphasize sentiment-driven deviations from rational pricing, they typically operate at the aggregate level, often assuming representative agents or single-factor behavioral updates. More recent work (e.g., Barberis, 2013) highlights the layered nature of investor psychology, but implementations often abstract away from how these behaviors unfold in response to specific information events.

In contrast, we develop a fine-grained framework that embeds *temporal heterogeneity* directly into headline-level pricing signals. Our goal is not just to capture whether news matters, but to model *who reacts*, and *when*—aligning with emerging interest in heterogeneous agent models, learning dynamics, and market microstructure interpretations of news digestion (e.g., Tetlock, 2007; Kelly et al., 2016; Bordalo et al., 2020).

Motivation and Setup

We hypothesize that different types of market participants react to the same news headline in distinct temporal patterns. To capture this, we define a set of archetypes, each associated with a behavioral kernel that governs how the signal from a given headline diffuses over time. This approach extends our structured scoring model by embedding time-dependent cognitive filters, allowing us to simulate stylized investor reactions without needing to identify specific agents.

Each headline already has a log-transformed final score s_i derived from our dictionary model. We interpret this as the latent signal strength of a news event. Rather than assuming all impact occurs on the day of publication, we allow the signal to diffuse across a short window using archetype-specific kernels.

Investor Archetypes and Kernels

We define three stylized archetypes:

- **Anxious**: Reacts strongly and immediately, but fades quickly. Suitable for investors highly sensitive to downside risk (e.g., retail panic or safety-related concerns).
- **Patient**: Updates more gradually, reflecting deliberative behavior by long-term investors or institutions.
- **Momentum-Seeking**: Shows increasing responsiveness over time before reverting, simulating trend-following behavior with bounded confidence.

Each archetype is associated with a fixed kernel over a 5-day window (centered at the publication date and extending ± 2 trading days). The kernels are defined as:

Anxious (Spike Kernel):

$$k_{\text{spike}}(t) = \begin{cases} 1.0 & \text{if } t = 0 \\ 0.1 & \text{if } t = 1 \\ 0.0 & \text{otherwise} \end{cases}$$

Patient (Gaussian Kernel):

$$k_{\text{gaussian}}(t) = \exp\left(-\frac{1}{2}\left(\frac{t-1.0}{1.2}\right)^2\right)$$

Momentum-Seeking (Reversion Kernel):

$$k_{\text{reversion}}(t) = \left(e^{-0.2t} - e^{-1.0t}\right) \cdot \mathbb{I}_{\{t \geq 0\}}$$

All kernels are normalized to sum to 1, ensuring that the total weight of a headline's signal is preserved:

$$\tilde{k}_a(t) = \frac{k_a(t)}{\sum_{u=-2}^{2} k_a(u)}$$

Signal Construction

For each theme (Company, Fuel and Oil Prices, Safety), we construct daily time series signals per archetype by summing the contributions from all relevant headlines:

$$\text{Signal}^{(a)}(t) = \sum_{i \in \mathcal{H}} s_i \cdot \tilde{k}_a(t - t_i)$$

Here, $\mathcal{H}$ denotes the set of headlines tagged with the given theme, s_i is the log final score of headline i, t_i is its publication date, and $\tilde{k}_a$ is the normalized kernel for archetype a.

If multiple headlines fall within the same window, their signals are summed—mimicking overlapping news impact and how investors may integrate related signals over a brief horizon.

Theme-Specific Regressions

We regress daily excess returns on these archetype-specific signals, separately for each theme:

$$\text{Excess Return}_t = \beta_0 + \beta_1 \cdot \text{Signal}^{\text{anx}}(t) + \beta_2 \cdot \text{Signal}^{\text{pat}}(t) + \beta_3 \cdot \text{Signal}^{\text{mom}}(t) + \varepsilon_t$$

This setup allows us to evaluate which archetype dominates pricing behavior for each type of news. As we show in Section 4, anxious and patient behaviors are most influential in the case of DAL—particularly for ambiguous or high-salience themes like safety—while momentum-based behavior plays a relatively minor role.

Implementation Notes

- Signal construction uses a rolling kernel over a ± 2 day window, aligned with a cleaned trading calendar.
- Archetype signals are initialized as zero vectors and updated additively.
- Theme-specific regressions help isolate heterogeneity in investor response by economic driver.

Relation to Prior Literature

Our kernel-based approach connects to three strands of literature:

1. Behavioral models of underreaction and overreaction (Daniel et al., 1998; Barberis et al., 1998), which motivate the distinction between fast and slow pricing forces.
2. Attention and salience models (Bordalo et al., 2020), which suggest that certain news themes may disproportionately affect some investor types.
3. News-based prediction and learning (Tetlock, 2007; Kelly et al., 2016), which emphasize how textual signals interact with price formation, often in aggregated ways.

By modeling distinct archetypes with temporal kernels, we bridge these areas—offering an interpretable and testable framework for understanding how different investor behaviors shape the pricing of specific news events.

4 Empirical Results

We evaluate the predictive performance of both FinBERT and our structured scoring model by regressing excess returns on their respective log-transformed scores. For each headline theme—Fuel and Oil Prices, and Company—we run regressions at three return horizons: day t, day $t+1$, and the sum of day $t+1$ and $t+2$. This allows us to examine both immediate and slightly delayed price impacts.

Fuel and Oil Prices

Table 2 presents regression results for fuel and oil-related headlines. Our structured score exhibits stronger and more interpretable relationships to excess returns, particularly on the event day.

Model	Horizon	$\mathbf{R^2}$	p-value	Coefficient
Structured Score	t	0.0513	0.0448	-0.0130
	$t+1$	0.0047	0.5507	0.0039
	$t+1+2$	0.0331	0.1131	-0.0105
FinBERT Score	t	0.0428	0.0675	0.0135
	$t+1$	0.0374	0.0896	0.0127
	$t+1+2$	0.0004	0.8614	0.0013

Table 2: Regression results for Fuel and Oil Prices headlines. Structured scores show stronger explanatory power at t and $t+1+2$ horizons.

Our model yields statistically significant results on day t ($R^2 = 0.0513$, p $= 0.0448$), capturing the immediate effect of changes in input costs such as jet fuel. Interestingly, the coefficient is negative, reflecting the fact that lower fuel prices (which receive positive structured scores) are beneficial for Delta and hence associated with positive excess returns. FinBERT also performs reasonably on the event day, but its coefficients are positive across the board—indicating it likely misinterprets falling prices as bad news due to its textual emphasis on phrases like "plummet" or "collapse."

At the $t+1$ and $t+2$ horizon, FinBERT's performance deteriorates sharply ($R^2 = 0.0004$, p $= 0.86$), whereas the structured model retains modest signal ($R^2 = 0.0331$), suggesting some lagged pricing effects that are economically plausible.

Company News

Regression results for company-related headlines are shown in Table 3. Again, the structured model dominates on event day performance.

Model	Horizon	R^2	p-value	Coefficient
Structured Score	t	0.0508	0.0220	0.0120
	$t+1$	0.0192	0.1653	0.0074
	$t+1+2$	0.0163	0.2038	0.0066
FinBERT Score	t	0.0021	0.6454	0.0028
	$t+1$	0.0196	0.1606	-0.0084
	$t+1+2$	0.0131	0.2537	0.0068

Table 3: Regression results for Company news headlines. Structured scores show stronger signal at t; FinBERT performance is weak overall.

Here, our structured score achieves an R^2 of 0.0508 ($p = 0.022$) at the event day with a positive coefficient, indicating that company news is priced quickly and cleanly by the market. FinBERT, in contrast, fails to show any significant explanatory power at t ($R^2 = 0.0021$, $p = 0.6454$), and its coefficients are inconsistent across horizons.

Interpretation

Across both themes, our model demonstrates consistently higher explanatory power and more coherent economic interpretations. This is particularly true at the t horizon, where structured scores show alignment with plausible market logic (e.g., falling input prices $\rightarrow$ positive returns).

While predictive power decays modestly beyond the event day, this does not imply a failure of the model. Rather, it suggests that different themes may be digested over different time horizons by different types of market participants. This motivates the next step in our framework: incorporating behavioral heterogeneity through archetype modeling.

Instead of asking whether a score "works," we ask a deeper question: *who in the market is reacting to the headline, how do they react over time, and how does this vary across themes like safety, fuel pricing, or company operations?* The archetype model is thus not a corrective layer, but an explanatory one—designed to uncover the diversity of investor behavior that underlies observed return dynamics.

5 Archetype-Based Analysis

We investigate how different types of investors—modeled as behavioral archetypes—respond to news across three themes: Company News, Fuel and Oil Prices, and Safety. Rather than ask if news predicts returns in the aggregate, we ask: *Who* reacts, *when*, and *how*?

Each archetype represents a stylized response profile:

- **Anxious:** immediate and short-lived reactions.
- **Patient:** gradual assimilation of information.
- **Momentum:** delayed reaction aligned with trend-following.

We regress excess returns on archetype-specific signal series. Table 4 summarizes key results.

Behavior Across Themes

Fuel and Oil Prices ($R^2 = 0.075$): Anxious investors show a marginally significant negative coefficient ($p = 0.051$), consistent with short-term concerns over volatile input costs. Patient investors exhibit a moderate positive coefficient, suggesting slower adjustment. Momentum response is negative—indicating lack of follow-through, possibly due to rapid mean reversion.

Company News ($R^2 = 0.042$): The signal is weakest here. Anxious investors show a small positive coefficient (0.012), while Patient and Momentum responses are negligible. High correlation between archetypes (especially Patient and Momentum: 0.82) suggests similar response windows or multicollinearity.

Safety News ($R^2 = 0.029$): No statistically significant effects emerge, though Anxious responses are negative and Patient responses slightly positive. This likely reflects high emotional salience but mixed market interpretation—perhaps due to attribution ambiguity or quick PR containment.

Interpretation

These results suggest theme-dependent investor dynamics:

- **Fuel and Oil:** Dominated by fast-acting, anxious behavior; some patient adjustment.
- **Company:** Mostly short-horizon trading interest; weak overall response.
- **Safety:** Spiky but inconsistent reactions, likely due to heterogeneity in investor interpretation.

Rather than seeking a single "best" archetype, this framework reveals *conditional behavior*: who responds and when depends critically on the type of news. It offers more explanatory nuance than static sentiment models.

Theme	Archetype	Coefficient	p-value
Fuel and Oil	Anxious	-0.0358	0.051
	Patient	0.0732	0.149
	Momentum	-0.0468	0.130
Company	Anxious	0.0119	0.259
	Patient	-0.0063	0.849
	Momentum	0.0056	0.788
Safety	Anxious	-0.0216	0.272
	Patient	0.0311	0.590
	Momentum	-0.0264	0.461

Table 4: Archetype regression results across themes. While no coefficients are statistically significant at conventional levels, their magnitudes and directions provide behavioral insight into theme-specific market reactions.

Sensitivity: Sharpening the Anxious Kernel

We test a pure one-day spike kernel for the Anxious archetype ($k(t) = 1$ for $t = 0$, 0 otherwise), simulating near-instantaneous reaction. For Fuel and Oil Prices, this raises R^2 to 0.084 and makes the Anxious coefficient statistically significant ($p = 0.034$). Company News results remain largely unchanged, indicating less dependence on timing.

The sharper kernel also reduces correlation with other archetypes (from 0.65 to ¡0.47), reinforcing that some themes are best captured by immediate response profiles.

Theme	**Archetype**	**Baseline Spike Kernel**		**Sharp Spike Kernel**	
		Coef.	p-value	Coef.	p-value
Fuel and Oil	Anxious	-0.0358	0.051	**-0.0320**	**0.034**
	Patient	0.0732	0.149	0.0689	0.130
	Momentum	-0.0468	0.130	-0.0471	0.107
Company	Anxious	0.0119	0.259	0.0093	0.282
	Patient	-0.0063	0.849	-0.0008	0.978
	Momentum	0.0056	0.788	0.0032	0.871

Table 5: Comparison of archetype regression coefficients under baseline vs. sharp spike kernel. Sharpening the Anxious kernel improves significance and interpretability for Fuel and Oil Prices, suggesting immediate market reaction is key in that theme.

Alternate Kernels: Tapered Spike and Smoothed Gaussian

Using alternate shapes—a tapered Anxious spike (1.0, 0.1) and smoother Gaussian for Patient—yields new insights. Notably, Safety News now shows a marginally significant negative Anxious coefficient (p = 0.084), suggesting that fear-based responses may be short-lived and intense.

Theme	**Archetype**	**Coefficient**	**p-value**
Fuel and Oil	Anxious	-0.0227	0.115
	Patient	0.0210	0.369
	Momentum	-0.0088	0.384
Company	Anxious	0.0123	0.162
	Patient	-0.0052	0.757
	Momentum	0.0031	0.707
Safety	Anxious	**-0.0300**	**0.084**
	Patient	0.0435	0.213
	Momentum	-0.0156	0.296

Table 6: Regression results with alternate kernel specification: smoother Gaussian, tapered spike. Spiky behavior becomes more explanatory in the Safety theme, suggesting sharp fear-driven reaction on the day of the event.

Kernel Tuning for Fuel and Oil Price Themes

We further explore response dynamics by tuning the Anxious spike kernel (via decay α) and the Patient Gaussian kernel (via width σ and delay μ).

Anxious Kernel Tuning. Figure 1 shows the coefficient is strongest and most significant at $\alpha = 0$, consistent with immediate reaction. As decay increases, explanatory power declines.

Patient Kernel Tuning. Figures 2a and 2b reveal best performance when $\sigma \approx 1.0$ and μ is near-zero—suggesting even “patient” investors respond within 1–2 days.

Conclusion: Archetypes as Behavioral Hypotheses

Market reaction is shaped not only by *what* the news says, but by *how* it’s processed. Archetypes—defined via temporal kernels—reveal investor cognition in motion. Their tuning enables interpretable dis-

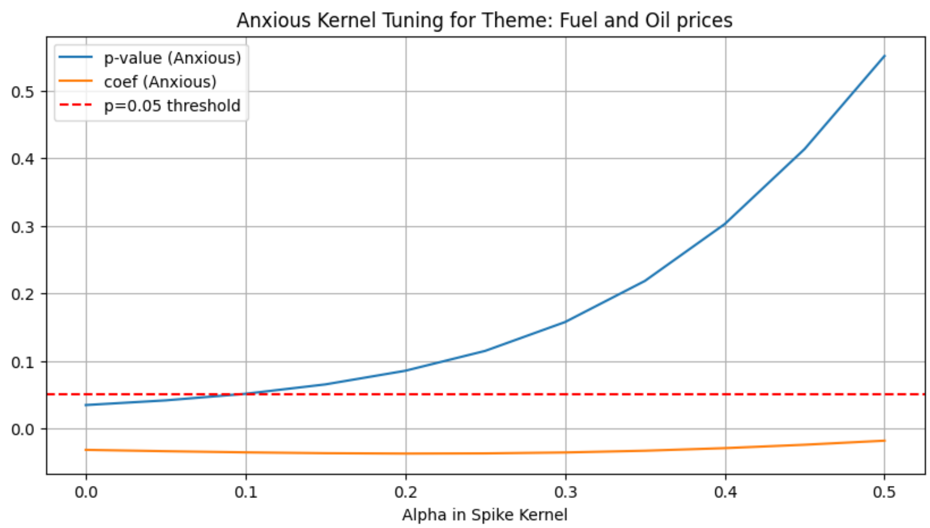
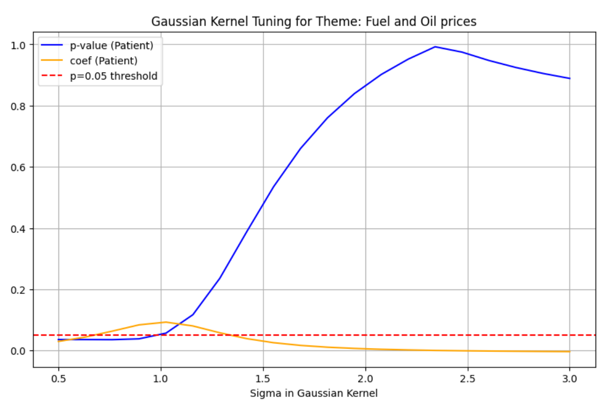
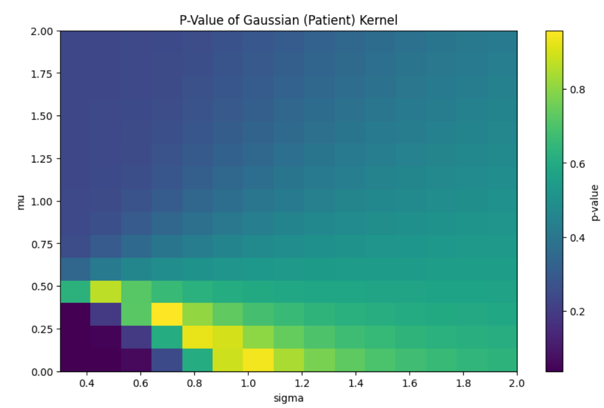


Figure 1: Effect of spike kernel decay (α) on Anxious coefficient and p-value for Fuel and Oil Prices. Highly localized (spiky) reactions dominate.



(a) 1D tuning of Patient (Gaussian) kernel for Fuel and Oil Prices. Peak signal strength occurs when $\sigma \approx 1.0$.



(b) 2D p-value heatmap for Patient kernel tuning across μ (delay) and σ (spread). Bottom left region reveals localized reaction behavior.

covery of who reacts, when, and how—and why some signals price in while others fade.

6 Conclusion and Future Work

This paper develops a cognitively structured and behaviorally interpretable approach to modeling the pricing impact of financial news headlines. Building on our prior work with Apple Inc. (AAPL), we apply this framework to Delta Airlines (DAL), a firm with different economic drivers, textual structure, and information salience. We show that our structured scoring model outperforms black-box sentiment tools like FinBERT, particularly for macro-sensitive themes such as fuel and oil prices.

Our central contribution lies in augmenting this scoring system with behavioral structure. By introducing investor archetypes—each defined by a kernel that distributes signal over time—we move beyond the question of "what does the news say?" to "who reacts, when, and how?" Across

DAL headlines, we find that Anxious and Patient archetypes explain more variation in excess returns than momentum-based behavior, especially for emotionally charged or ambiguous themes such as safety.

Importantly, the archetypes used here are intentionally simple—serving as conceptual scaffolding for prototyping. As this framework scales, we envision defining archetypes as *linear combinations of basis kernels*, or learning them directly from empirical response patterns. This opens the door to personalization, empirical grounding, and more flexible modeling, while maintaining the core interpretability that distinguishes this approach from opaque machine learning models.

Future Directions. Several extensions follow naturally:

- **Cross-asset generalization:** Applying the framework to other industries (e.g., financials, energy, retail) to explore how theme structure and archetype dominance vary by domain.
- **Learned archetypes:** Moving beyond fixed kernels by estimating response functions from data or modeling archetypes as mixtures over kernel bases (e.g., via functional regression or NMF).
- **Forecasting applications:** Integrating archetype-weighted signals into predictive models or trading strategies to assess practical value beyond explanatory regression.
- **Systemic and multi-firm modeling:** Extending the analysis to portfolios or correlated firms to explore how investor types differentially respond to idiosyncratic vs. systemic news.
- **Behaviorally constrained LLMs:** Using the scoring and kernel logic as structural priors for fine-tuning or prompting large language models in financial NLP tasks.

In sum, this paper contributes a modular and interpretable pipeline for financial news analysis—bridging cognitive structure, behavioral dynamics, and empirical rigor. As headlines increasingly shape asset pricing in real time, such transparency may become not just useful, but essential. While more complex or black-box models may offer marginal gains in predictive accuracy, our approach prioritizes interpretability and modularity—making it especially suitable for high-stakes domains where transparency, auditability, and behavioral insight matter as much as raw performance.

References

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, 49(3):307–343.

Barberis, N. C. (2013). Thirty years of prospect theory in economics: A review and assessment. *Journal of Economic Perspectives*, 27(1):173–196.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Memory, attention and choice. *The Quarterly Journal of Economics*, 135(3):1399–1442.

Daniel, K., Hirshleifer, D., and Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *The Journal of Finance*, 53(6):1839–1885.

Kelly, B., Pastor, L., and Veronesi, P. (2016). The price of political uncertainty: Theory and evidence from the option market. *The Journal of Finance*, 71(5):2417–2480.

Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*, 62(3):1139–1168.