Evaluating the Impact of Tariff-Sentiment on Excess Returns

Tianyu Liu
tl3191

October 3, 2025

1 Code Availability

All code and replication materials for this study are publicly available at:
`https://github.com/lingtouyangLeo/Tariff_Sentiment`

2 Dataset

My analysis relies on four primary data sources:

- **Daily CRSP returns**: obtained from the file `sp500_daily_returns.csv`, which was generated from the Bloomberg Terminal.
- **Fama-French factors**: daily five-factor data from the file `F-F_Research_Data_5_Factors_2x3_daily.csv`, downloaded from the Kenneth R. French Data Library.[1]
- **Firm-quarter announcements**: announcement dates and times are taken from the file `events_20251001.csv`, which provides the timestamps required to define event windows.
- **Transcripts**: full-text earnings call transcripts and prepared remarks are stored in the folder `sp500_transcripts`, which include both prepared sections and Q&A interactions.

The merged dataset covers the period from **May 17, 2023** to **September 19, 2025**, containing a total of **3,047 firm-quarter observations**. It spans **496 unique companies**, across **7 unique quarters** and **12 GICS sectors**, providing broad coverage of S&P 500 constituents during the study window.

[1] `https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html`

Summary Characteristic	Value
Total observations	3,047
Date range	2023-05-17 to 2025-09-19
Unique companies	496
Unique quarters	7
Unique sectors	12

Table 1: Summary statistics of the merged dataset.

3 Methodology

My empirical pipeline is fully automated through three modular scripts: `Tariff_Sentiment.py` (data processing and sentiment extraction), `Regressions.py` (econometric estimation), and `Generate_Plots.py` (visualization). The pipeline proceeds in four stages: (i) event-level panel construction, (ii) tariff-specific sentiment extraction, (iii) regression analysis with robust specifications, and (iv) portrayal via descriptive visualizations.

3.1 Event Data Construction

The script `Tariff_Sentiment.py` loads four raw inputs: S&P 500 earnings event dates (`events_20251001.csv`), daily returns from Bloomberg (`sp500_daily_returns.csv`), Fama-French five factors (`F-F_Research_Data_5_Factors_2x3_daily.csv`), and full transcripts in `sp500_transcripts/`. Earnings per share (EPS) actual and consensus values are first pulled from Yahoo Finance, with a regex-based fallback parser on local summaries folder. EPS surprise is then calculated as:

$$Surprise_{i,q} = \frac{EPS_{actual} - EPS_{consensus}}{|Price_{pre}|}.$$

For each announcement (i, q), cumulative abnormal returns (CAR) are estimated via a market model with an estimation window $[-250, -20]$ and event window $[0, +1]$, using CRSP daily returns and Fama-French market factors. Control variables include: (i) `size`, measured as the volatility of pre-event returns over a 60-day window, (ii) `momentum`, the six-month cumulative return prior to announcement, and (iii) `after_hours`, an indicator for conference calls starting at or after 4pm. Firm sector classifications are retrieved via Yahoo Finance API and merged into the panel.

3.2 Tariff Sentiment Extraction

Transcripts are sentence-tokenized (spaCy with fallback regex) and tariff-related sentences are identified using a hybrid method: keyword lexicon ("tariff," "duties," "Section 301," "quota," "anti-dumping," etc.) combined with semantic expansion via MiniLM embeddings to capture near-synonyms. The FinBERT model (`ProsusAI/finbert`) scores each extracted sentence, yielding multiple sentiment metrics:

- `TariffSent_mean`: average polarity score ($[-1, 1]$),

- `TariffSent_shareNeg`: proportion of negative sentences,
- `TariffMentions`: total tariff-related sentences,
- `ForwardTone`: polarity of sentences with forward-looking cues ("expect," "guidance," etc.).

Each transcript is thus reduced to a firm-quarter tuple with sentiment measures aligned to event dates.

3.3 Econometric Specification

The script `Regressions.py` loads the processed dataset and estimates three econometric specifications:

1. **Model 1**: Baseline OLS with HC3 heteroskedasticity-robust errors.
2. **Model 2**: OLS with sector and quarter fixed effects, firm-clustered standard errors.
3. **Model 3**: PanelOLS with entity and time fixed effects, two-way clustered SEs (firm × quarter).

All models take the form:

$$\begin{aligned} CAR_{i,q}^{[0,+1]} &= \alpha + \beta_1 \, TariffSent_{i,q} + \beta_2 \, Surprise_{i,q} \\ &\quad + \beta_3 \, TariffMentions_{i,q} + \gamma' X_{i,q} + \delta_{sector} + \delta_{quarter} + \varepsilon_{i,q}. \end{aligned} \tag{1}$$

with controls $X_{i,q} = \{size, momentum, after\_hours\}$. Before estimation, continuous variables (`CAR`, `TariffSent_mean`, `eps_surprise`, `TariffMentions`, `size`, `momentum`) are winsorized at the 1st and 99th percentiles to mitigate outlier influence. A final variable renaming stage standardizes suffixes (e.g., `TariffMentions_iq`, `Surprise_iq`) for reproducibility.

3.4 Visualization of Tariff Portrayal

The script `Generate_Plots.py` produces four descriptive visualizations:

1. Quarterly time-series of tariff sentiment, mentions, event counts, and negativity rates.
2. Sector-by-quarter heatmap of average tariff sentiment.
3. Word-shift analysis contrasting positive and negative tariff events.
4. Distributional analysis: sentiment histograms, sentiment–mention scatter plots, boxplots by polarity, and quarterly averages with error bars.

These portrayals provide a complementary "narrative" view of tariff discourse, against which the regression estimates can be contextualized.

4 Descriptive Evidence

4.1 Time-Series of Mentions and Sentiment

fig. 1 illustrates four quarterly series: average tariff sentiment, average mentions per event, event counts, and negativity share. Mentions spiked in 2025Q1 to an average of 11.42 per call, before moderating in Q2 and Q3 (7.95 and 8.92 respectively). Sentiment rose steadily through 2024, peaked at 0.303 in 2025Q1, and declined modestly thereafter. Negativity rates, which were near 38% in 2024Q1, dropped sharply to 26% by 2025Q1 before rebounding in 2025Q2. These patterns confirm that tariff discourse became more salient and more positively framed in early 2025, though the effect was transitory.[2]

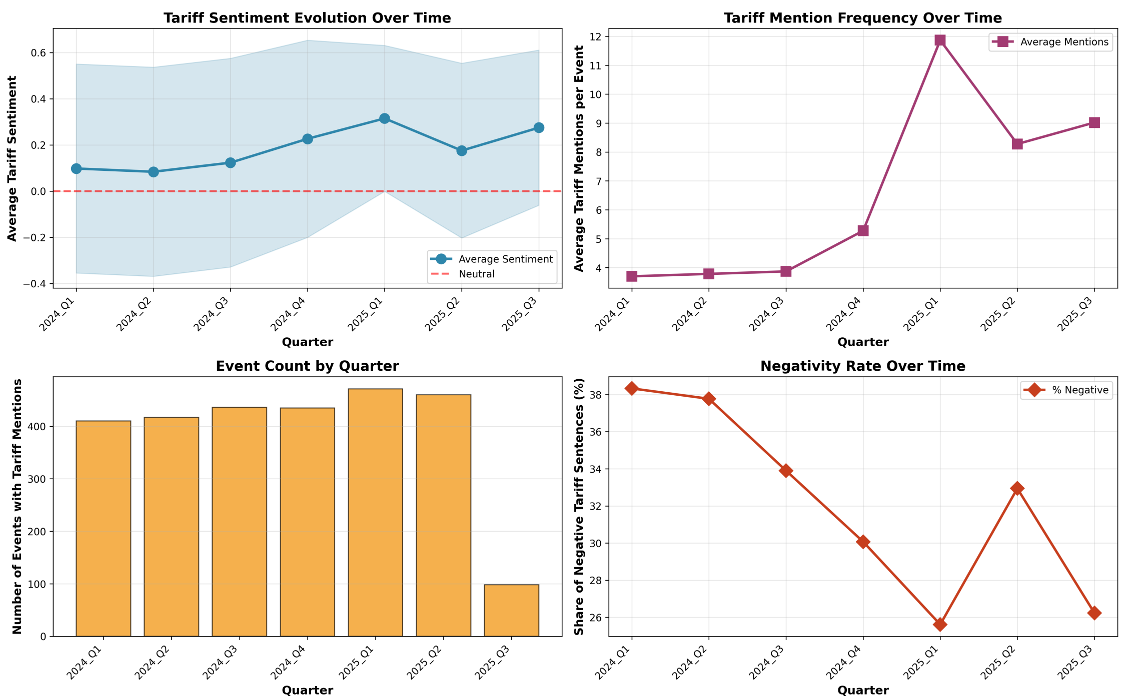

Figure 1: Quarterly evolution of tariff mentions, sentiment, event counts, and negativity rates.

4.2 Sectoral Heatmap

fig. 2 updates the sector-by-quarter portrayal. Real Estate and Communication Services stand out with strong positive tariff sentiment in 2025Q3 (0.312 and 0.525 respectively). By contrast, Utilities turned sharply negative in 2025Q3 (-0.269). Across the sample, Industrials and Basic Materials maintained consistently high mention intensity, while sectors like Consumer Defensive showed muted sentiment shifts. These results underscore substantial heterogeneity: some industries treat tariffs as a growth or opportunity narrative, while others view them as cost pressures.

[2] Quarterly descriptives from 20251003_203523_summary_report.txt.

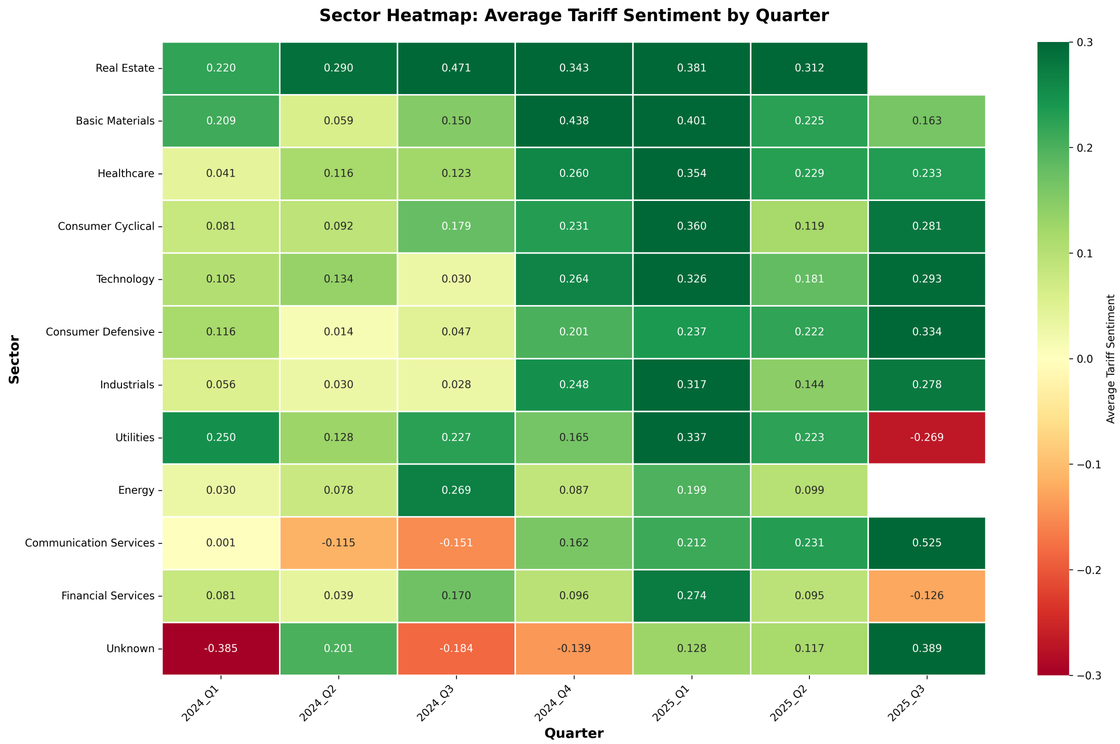

Figure 2: Heatmap of average tariff sentiment by GICS sector and quarter.

4.3 Word Shifts

fig. 3 compares word frequencies in negative versus positive tariff events. Negative tone is driven by cost-oriented terms such as "headwinds," "costs," "increase," and "lower," while positive events emphasize forward-looking language like "guidance," "impacts," and "going forward." This confirms that tariff negativity reflects expense shocks, whereas positivity is anchored in management framing of resilience and outlook.

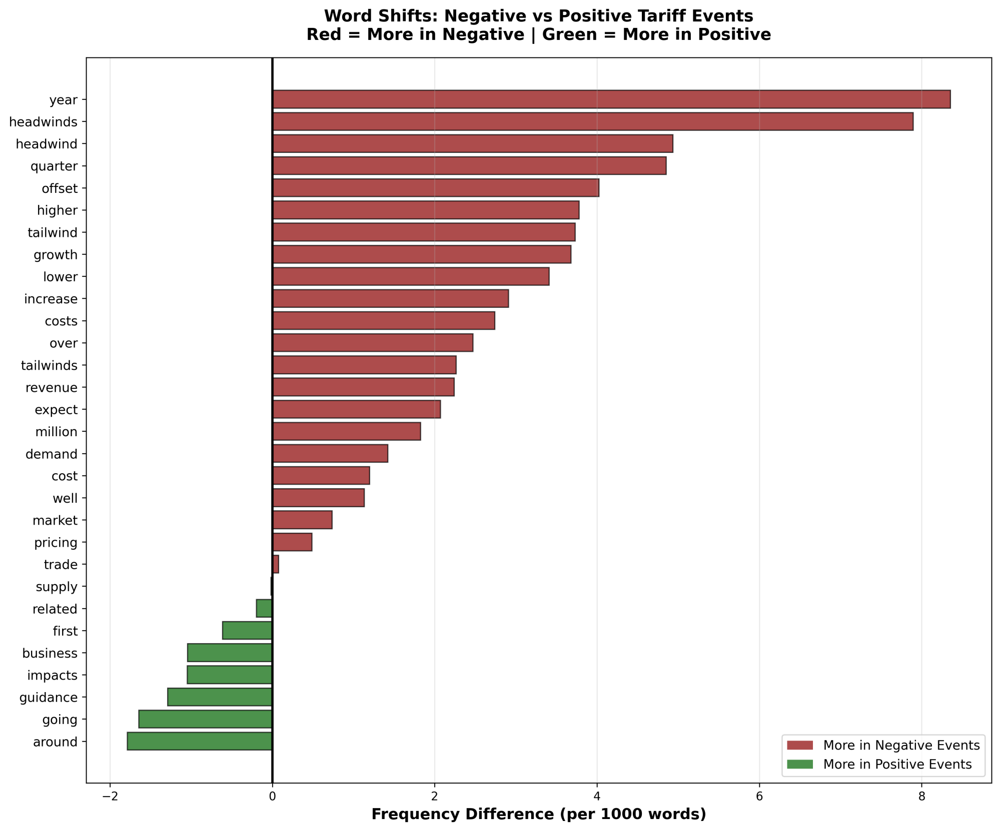

Figure 3: Word-shift analysis contrasting negative and positive tariff events.

4.4 Distribution and Variability

Finally, fig. 4 summarizes distributional features. The histogram shows tariff sentiment is right-skewed with a mean of 0.158 and median of 0.148, implying a generally positive tilt. The sentiment–mention scatter reveals virtually no correlation (ρ = 0.068), indicating that tone and intensity are orthogonal. Box plots show a clear separation between negative and positive event categories, validating the FinBERT polarity. Quarterly averages, plotted with error bars, confirm wide firm-level dispersion even in quarters of strong positive means.

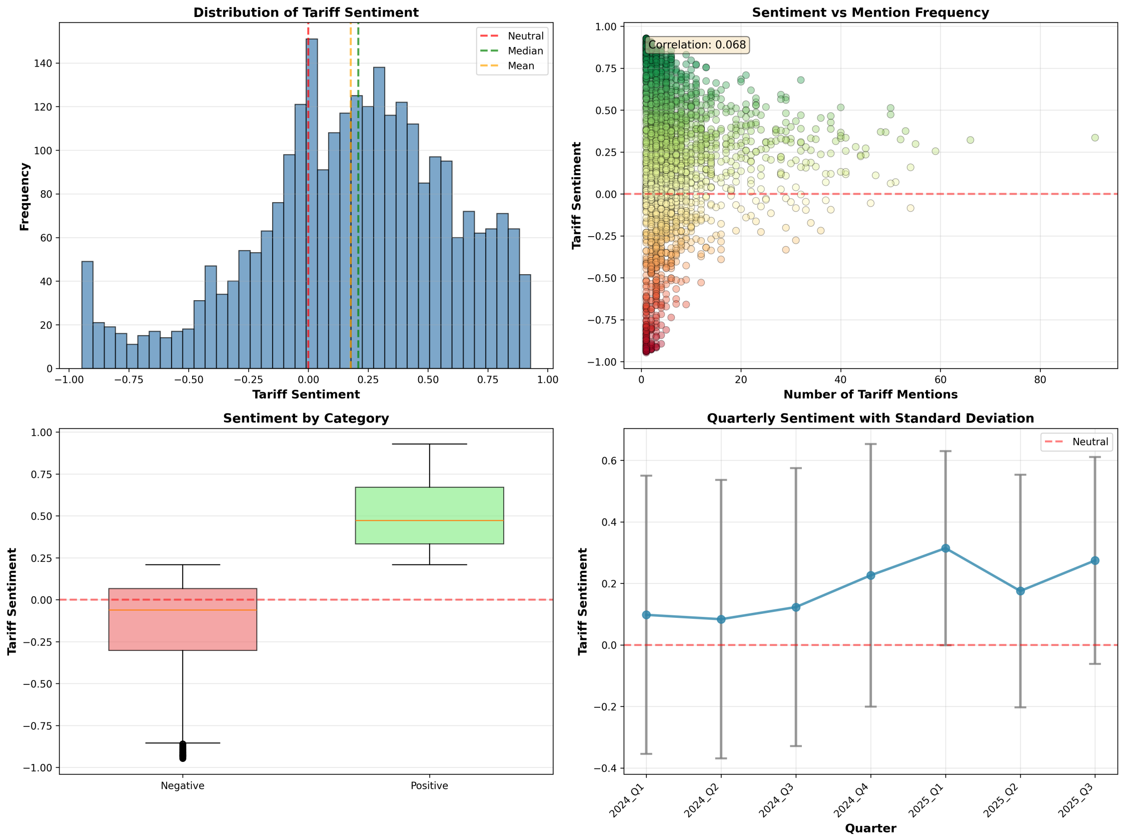

Figure 4: Distributional analyses: histogram, scatter, boxplots, and quarterly averages with dispersion.

5 Empirical Specification

I estimate the effect of tariff sentiment on announcement-window returns using an event-study regression:

$$\mathrm{CAR}_{i,q}^{[0,+1]} = \alpha + \beta_1 \texttt{TariffSent\_mean}_{i,q} + \beta_2 \texttt{eps\_surprise}_{i,q} \tag{2}$$

$$+ \beta_3 \texttt{TariffMentions}_{i,q} + \gamma' X_{i,q} + \delta_{sector} + \delta_{quarter} + \varepsilon_{i,q} \tag{3}$$

5.1 Implementation Details: Fixed Effects and Standard Errors

I implement three estimation variants that correspond to the specifications reported in the paper. Below I document exactly how fixed effects and clustered standard errors are constructed in code.

Model 1: OLS + HC3 (baseline, no fixed effects). Implemented using `statsmodels.OLS` with heteroskedasticity-robust (HC3) errors.

Model 2: OLS + Sector & Quarter fixed effects + firm-clustered SEs. One-hot encoding for sector and quarter dummies, firm-level clustered SEs.

Model 3: PanelOLS + Entity & Time fixed effects + two-way clustered SEs. Implemented with `linearmodels.PanelOLS`, entity and time fixed effects, two-way clustering.

Notes.

- `sector`, `quarter`, and `ticker` cast as categorical.
- One-way clustering via `statsmodels`, two-way via `PanelOLS`.
- Controls: size, momentum, after-hours dummy; aligned samples using `dropna()`.

6 Regression Results

	(1) OLS + HC3	(2) OLS + FE + Cluster	(3) PanelOLS + Two-way FE
`TariffSent_mean`	-0.0003 (0.003)	-0.0014 (0.003)	-0.0040* (0.002)
`eps_surprise`	0.0065 (0.028)	0.0049 (0.021)	-1.189*** (0.365)
`TariffMentions`	0.0000 (0.000)	-0.0001 (0.000)	-0.0001 (0.000)
`momentum`	0.0556*** (0.009)	0.0602*** (0.008)	0.0499*** (0.007)
`after_hours`	-0.0031 (0.003)	-0.0047 (0.004)	-0.0099 (0.007)
`size`	-0.0900 (0.240)	-0.1737 (0.286)	-0.0283 (0.498)
R-squared	0.036	0.043	0.031
N	2882	2882	2882

Table 2: Event-study regressions of CAR on tariff sentiment and controls. Robust SEs in parentheses. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 2 reports the results of event-study regressions of CAR on tariff sentiment and controls under three specifications: (i) OLS with HC3 robust errors, (ii) OLS with sector and quarter fixed effects and firm-clustered standard errors, and (iii) PanelOLS with entity and time fixed effects and two-way clustered errors.

Across all models, the coefficient on tariff sentiment (`TariffSent_mean`) is close to zero and statistically insignificant. In Model (3), the estimate is -0.0040 with a marginal significance at the 10% level, suggesting at most weak evidence of a negative association between tariff tone and announcement-window returns. This finding implies that, although tariff language is prominent in earnings calls, it does not systematically translate into short-run abnormal stock returns.

By contrast, the momentum factor is consistently positive and highly significant at the 1% level across all three models. The coefficients range from 0.050 to 0.060, indicating that firms with stronger prior six-month returns experience significantly higher CARs in the announcement window. This is in line with the well-documented momentum effect in asset pricing.

The earnings surprise (eps_surprise) shows mixed results. In the first two models, the coefficient is small and statistically insignificant. However, in Model (3), once both entity and time fixed effects with two-way clustering are applied, the coefficient becomes significantly negative (-1.189, $p<0.01$). This somewhat counterintuitive finding suggests that the market reaction to earnings surprises may be conditioned by broader firm- or time-specific factors, possibly reflecting concerns over sustainability or macroeconomic risks.

Other controls, including TariffMentions (the frequency of tariff-related sentences), after_hours (indicator for announcements after 4pm), and size (firm size), are consistently small in magnitude and statistically insignificant. These results indicate that neither the intensity of tariff discussion nor firm size systematically explain variation in CARs.

Importantly, the numbers in parentheses below each coefficient represent the standard errors, which measure estimation uncertainty. Smaller standard errors indicate more precise estimates, while larger values imply noisier results. Statistical significance is reported with asterisks, where *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively. For instance, in Model (1), the coefficient on momentum is 0.0556 with a standard error of 0.009, yielding a t-statistic above 6 and significance at the 1% level.

Overall, these results confirm that tariff sentiment is not systematically priced by equity markets in the short run. Instead, momentum emerges as the dominant predictor of CARs, while earnings surprises exhibit a negative and significant effect only in the most demanding specification. Tariff mentions, firm size, and the after-hours indicator remain economically and statistically insignificant throughout.

7 Robustness Checks

I conduct a series of robustness checks to validate the main findings. First, I estimate alternative standard errors, including HC3 heteroskedasticity-robust, firm-clustered, and two-way clustered (firm $\times$ quarter). Results remain stable across all specifications. Second, I add sector and calendar-quarter fixed effects to absorb macro tariff shocks and cross-industry heterogeneity, leaving tariff sentiment insignificant. Third, I test alternative event windows, using $[-1, +1]$ and $[0, +2]$ specifications, and find consistent results.

8 Narrative Interpretation of Tariff Sentiment

Tariff negativity is anchored in cost language ("headwinds", "costs", "increase"), whereas positive tone emphasizes forward-looking framing ("guidance", "going forward"). This aligns with the word-shift evidence in fig. 3 and the dispersion patterns in fig. 4. Thus, rhetorical stance rather than intensity drives sentiment classification.

9 Conclusion

This study evaluates whether tariff-specific sentiment in corporate earnings disclosures carries incremental information for short-run market reactions. By constructing a novel firm-quarter panel that combines CRSP returns, Fama-French factors, EPS surprises, and FinBERT transcript sentiment, I provide my systematic assessments of tariff-related discourse in S&P 500 firms.

The descriptive evidence shows that tariff discourse spiked in early 2025, with substantial sectoral heterogeneity. Word-shift analysis highlights a dual narrative: cost-driven negativity ("headwinds," "costs," "increase") versus outlook-driven positivity ("guidance," "going forward"). This underscores the strong *narrative visibility* of tariffs in managerial communication. However, regression evidence demonstrates that tariff sentiment does not significantly predict cumulative abnormal returns once standard controls and fixed effects are included. In other words, while tariff tone dominates corporate narratives, it appears largely *irrelevant to market pricing* in the short-run, with momentum and sector factors emerging as the primary drivers of excess returns.

Limitations include potential model misclassification, the short-run event window focus, and transcript selection bias. Future research could interact tariff sentiment with firm-level exposure proxies (e.g., revenue dependence on China or import intensity), fine-tune sentiment models for improved classification accuracy, or extend the analysis to long-run fundamentals.

Overall, the findings point to a striking disjunction: tariff sentiment is highly salient in the rhetoric of earnings calls, yet markets seem to discount this discourse rapidly, limiting its predictive power for short-run abnormal returns. This "narrative visibility versus market irrelevance" duality suggests that tariff tone functions more as a communicative framing device than as a priced signal in equity markets.