

# News Signal Modeling for Apple Inc.

One-Page Summary

## Key Questions Addressed

Theme	Question
Structured News Scoring	Can a cognitively motivated, theme-aware scoring system outperform standard sentiment models like FinBERT in predicting AAPL returns?
Behavioral Lag Modeling	Can structured scores capture delayed investor responses across themes such as Chip Supply and Competition, beyond immediate sentiment effects?
Practical Applicability	Is the model transparent, modular, and scalable for real-time signal construction in financial NLP workflows?

## Conceptual Ideas Proposed

- A **cognitively structured scoring framework** rates news headlines along four axes: Actor Clarity, Lever Relevance, Magnitude, and Novelty, using a custom dictionary for each dimension.
- Each article is assigned a **theme tag** (e.g., Chip Supply, Competition) and scored by multiplying the four axes, reflecting both economic salience and behavioral response likelihood.
- **Manual keyword curation and NLP-based expansion** ensure linguistic coverage across tense, synonyms, and phrase variations while preserving interpretability.
- The framework is designed for **transparent, modular deployment** alongside econometric or ML models, supporting signal decomposition and cross-theme attribution.

## Key Results

- For the **Chip Supply theme**, the model shows its strongest signal at T+3 with  $R^2 = 0.080$  and  $p = 0.100$ , outperforming FinBERT at longer lags where behavioral effects dominate.
- For the **Competition theme**, the model achieves a statistically significant signal at T+3 with  $p = 0.045$  and  $R^2 = 0.024$ , capturing delayed investor responses not detected by FinBERT.
- **Same-day (T+0) performance is limited**, as expected, aligning with the model’s focus on structured, lag-aware inference rather than immediate sentiment capture.
- The framework provides a **transparent, behaviorally grounded alternative** to black-box sentiment models, suitable for scalable signal construction across firms and asset classes.

## Illustrative Tables

Theme	Horizon	$n$	Coef.	$t$	$p$	$R^2$
Analyst	3d	36	-0.0019	-0.37	0.716	0.004
Analyst	4d	36	0.0037	0.51	0.613	0.008
Chip Supply	3d	35	0.0112	1.69	0.100	0.080
Chip Supply	4d	35	-0.0029	-0.49	0.627	0.007
Company	3d	27	-0.0009	-0.06	0.956	0.000
Company	4d	27	0.0147	0.89	0.383	0.031
Competition	3d	166	<b>-0.0115</b>	<b>-2.02</b>	<b>0.045</b>	<b>0.024</b>
Competition	4d	166	-0.0048	-0.84	0.402	0.004

Table 1: **Regression results using structured scoring.** Strongest signals appear for Chip Supply and Competition themes at T+3 and T+4.

Theme	Horizon	$n$	Coef.	$t$	$p$	$R^2$
Analyst	3d	36	-0.0023	-0.67	0.510	0.013
Analyst	4d	36	-0.0004	-0.08	0.933	0.000
Chip Supply	3d	35	-0.0072	-1.57	0.125	0.070
Chip Supply	4d	35	-0.0019	-0.49	0.631	0.007
Company	3d	27	0.0077	1.20	0.242	0.054
Company	4d	27	0.0101	1.58	0.126	0.091
Competition	3d	166	0.0028	1.43	0.154	0.012
Competition	4d	166	-0.0003	-0.13	0.893	0.000

Table 2: **Regression results using FinBERT scoring.** Structured model shows clearer theme-specific effects and stronger statistical significance.