A Sentiment-Based Approach to Convenience Yield Dynamics in Commodity Futures Markets

Massab Qayum and Joe Janzen

A Sentiment-Based Approach to Convenience Yield Dynamics in Commodity Futures Markets

Massab Qayum[1]
Joe Janzen[1]

Paper prepared for the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, 2025.

[1]Department of Agricultural and Applied Economics, University of Illinois Urbana-Champaign
[1]Corresponding Author: mqayum2@illinois.edu

A Sentiment-Based Approach to Convenience Yield Dynamics in Commodity Futures Markets

Abstract: *Convenience yield acts as a negative cost (i.e. benefit) of holding a physical commodity and causes inverted markets or backwardation. The theory of storage relies on convenience yield to explain these backwardations through the 'Working Curve,' which has been empirically documented for storable commodities using annual stocks data. However, the intra-year dynamics of convenience yield is difficult to explore using annual inventories data. This study introduces a novel approach to analyzing convenience yield dynamics in commodity markets by leveraging sentiment analysis, enabling a more granular understanding of convenience yield dynamics in corn and soybean markets. The research questions we are trying to answer are twofold: how forward-looking and efficient is the information captured from news sentiment for storable commodity prices, and can it help explain convenience yield dynamics in these markets? The Global News Index developed by the Cline Center for Advanced Social Research at the University of Illinois is used to track global news attention to commodity markets. Futures prices and implied volatilities for corn and soybean are extracted from Bloomberg. The results from structural VAR show significant, immediate, and consistent negative responses in the implied volatilities to news sentiment shocks. No significant, immediate, and consistent responses are observed in convenience yield. Implied volatility tends to respond negatively by 0.5 to 1.5 percentage points to shocks in news sentiment, specifically in financial and soybean-specific news shocks. In conclusion, news sentiment does not provide forward-looking information regarding convenience yield dynamics but provides some information regarding price uncertainty 1-3 months ahead.*

Key Words: convenience yield, commodity futures markets, news sentiment analysis

Introduction

"Prices of seasonally produced commodities are speculative." - T.A. Hieronymus (1977, pp 141).

Seasonal commodity markets are driven by expectations of forthcoming market conditions, but expectations are hard to observe and quantify. Storage links expected future prices to prices in the present. Commodity futures markets provide observable values for these expected future prices. The theory of storage explains part of this link through observable costs of storage and part through the concept of convenience yield, an unobserved benefit of holding physical inventories that explains why the spread between current and expected future prices does not cover the full cost of storage. Explaining what determines the convenience yield, the component of observed intertemporal price spreads that cannot be explained by storage costs, and how convenience yields change over time is a major challenge for the economic analysis of commodity markets.

One way to understand convenience yield is to relate observed inventories to convenience yields. This relationship, which forms the basis of the 'Working Curve,' has been empirically documented for storable commodities using physical inventory data and futures spreads (Working 1949, Telser 1958, Brennan 1958, Joseph, Irwin, and Garcia 2016). However, the existence of this relationship does not describe in detail the intra-year dynamics of these spreads since inventories are typically observed almost a once a year.

This analysis presents a novel attempt to explain intra-year convenience yield dynamics in commodity markets. To do so, this study leverages data on market sentiment derived from news articles to facilitate a more granular analysis of convenience yields and market expectations. The underlying hypothesis is that sentiment expressed in news articles is a forward-looking measure of market expectations, so that unexpected variation in news-based sentiment explains observed changes in convenience yield. Unlike inventory data, which is infrequent and location-specific, sentiment potentially provides a continuous and broad-based measure of market outlook. This approach builds on the growing body of literature where sentiment analysis has been applied to provide valuable insights across diverse fields, spanning finance, politics, sociology, and psychology.

Many studies have found that news sentiment has predictive/forecasting power in the case of certain economic indicators, however this does not imply a causal relationship between sentiment and those indicators in the absence of some natural experiment or an economic model. There exists both recent and older applications of sentiment analysis from news articles to analyze financial markets. Chang and Suk (1998) analyzed insider trading disclosures in the Wall Street Journal's "Insider Trading Spotlight" section and found abnormal returns emerged after media reporting on Securities and Exchange Commission (SEC)-filed insider transactions. Klibanoff et al. (1998) found that prominent New York Times coverage of international markets and closed-end country funds led to faster price adjustments to fundamental news, accompanied by increased trading activity and heightened return volatility during major news weeks. Engelberg (2008) demonstrated that market prices may systematically underreact to news content, particularly when accounting for variations in media tone. Tetlock et al. (2008)’s study reveals that quantitative measures of negative language in news stories predict lower future corporate earnings and short-term stock returns. Feng et al. (2022) find that news sentiment, particularly macroeconomic news, significantly influences volatility states in the Japanese stock market. Costola et al. (2023) used a financial market-adapted BERT model to extract sentiment from online articles on COVID-19 and find a statistically significant positive relationship between sentiment scores and S&P 500 returns, with news sentiment and categories from NYTimes.com showing distinct effects on market returns.

Other studies have used sentiment to help measure specific economic parameters derived from economic theory. Shapiro and Wilson (2022) introduce a novel method for inferring the Federal Reserve's inflation target, an important economic policy parameter, by analyzing the language used in its internal meetings regarding inflation, unemployment, and other economic goals. They find that the Fed’s implicit average inflation target during 2000-2011 was approximately 1.5%, significantly below the widely assumed 2%.

This study is closer to this second approach to sentiment analysis by deriving measures of sentiment from news articles and studying how it relates to convenience yield dynamics in corn and soybean markets. The research questions we are trying to answer are twofold: how forward-looking and efficient is the information captured from news sentiment for storable commodity prices, and can it help explain convenience yield dynamics in these markets?

This study uses news article data relevant to corn and soybean markets to extract sentiment signals using two indicators: Bing Liu and General Inquirer. A systematic filtering process is used to capture market-relevant sentiment signals from a large news article database. We follow Carter, Rausser, and Smith (2017) to calculate convenience yield of the first-deferred and the 1 year ahead delivery contracts for corn and soybean. To assess the dynamic relationship between sentiment and convenience yield, we use Structural Vector Auto-Regression Model (SVAR). This model allows us to investigate whether news-based sentiment is forward-looking and whether it helps explain why changes in convenience yield in the first-deferred and the 1-year ahead time periods. Implied volatility indices for corn and soybean are used as a robustness check for whether sentiment is forward-looking in terms of price uncertainty. The results show some significant, immediate, and consistent responses in the implied volatilities to news sentiment shocks but not in convenience yield responses. This study calls into question the forward-looking nature of news sentiment and motivates further study news sentiment analysis in storable commodity markets.

Data and Methodology

Our analysis requires data on sentiment, soybean and corn market prices and implied volatility indices. We measure news-based sentiment relevant to commodity markets using data from the Global News Index developed by the Cline Center for Advanced Social Research at the University of Illinois Urbana Champaign is used to track global news attention to commodity markets. The Cline Center's database has over 170 million news articles and draws on two types of data resources: extreme-scale collections of unstructured textual data created by news content providers around the world (including Foreign Broadcast Information Service and BBC Summary of World Broadcasts), and structured datasets created by the Cline Center's Voyager web crawl system. Moreover, futures prices and 4-month ahead implied volatility indices for corn and soybean are extracted from Bloomberg.

Process of filtering relevant articles using article titles:

To capture market-relevant sentiment signals from the news article database, a systematic filtering process is implemented on the title of each article. Following Du and Dong (2023) and Baker et al. (2016), this process involves querying for specific words within article titles that are indicative of financial markets, macroeconomic trends, and commodity-specific developments. Certain words that could introduce noise unrelated to market sentiment are explicitly excluded.

The rationale for prioritizing article titles over full content analysis is supported by existing literature, which suggests that headlines significantly shape readers' perceptions and influence their retention of information. Lemarié, Lorch, and Péry-Woodley (2018) highlight that headlines can bias text comprehension by emphasizing particular aspects of an article's content. Moreover, Konnikova (2014) discusses how headlines frame the overall reading experience and influence subsequent recall and interpretation. While the content remains an essential component of an article, these studies indicate that the title plays a crucial role in shaping readers' understanding and perception, justifying its use as a filtering criterion in sentiment analysis.

There are four categories of articles made: 1. Financial articles, 2. commodity-specific financial articles, 3. corn-specific articles, and 4. soybean-specific articles.

Financial Articles:

To ensure that the retrieved articles are relevant to financial markets, the filtering process includes a predefined set of keywords associated with various financial sectors, including banking, investment, macroeconomics, and corporate trends. The selected keywords are listed in table 1:

Banking & Investment		**Macroeconomic Indicators**	
Bank	Mutual Fund	Monetary	Credit
Stock Market	Hedge Fund	Fiscal	Currency
Interest Rate	Capital Gains	Liquidity	Forex
Bond Yield	Asset Management	Deficit	Debt
Equity		Inflation	Loan
Portfolio		GDP	Mortgage
Market Conditions		**Trade & Commerce**	
Yield Curve	Retail Sales	Supply Chain	Tariff
Treasury	Prices	Futures Market	Subsidy
Federal Reserve	Income	Trade	
Recession		Export	
Market Conditions		Import	
Forecast			
Financial			
Consumer Confidence			
Corporate Finance		**Labor & Industry**	
Finance	Merger	Downturn	Manufacturing
Quantitative Easing	Acquisition	Deflation	Industrial Production
Austerity	Stimulus	Unemployment	PMI
IPO		Labor Market	Credit Crunch
Market Behavior		**Real Estate & Fintech**	
Bankruptcy	Risk Appetite	Housing Market	Property Values
Bailout	Earnings Report	Real Estate	Fintech
Capital	Revenue		
Derivatives	Investor		
Bull Market			
Bear Market			
Market Volatility			

Table 1: Title keywords used to identify finance and macroeconomic articles

It is important to note that while cryptocurrency and blockchain-related terms could be included, their emergence primarily in the early 2010s may disproportionately increase the number of articles in the latter part of the study period (2000-2025). To prevent potential biases in the sentiment analysis, cryptocurrency-related terms were excluded from the query.

Exclusion of War-Related Articles:

A significant challenge in retrieving relevant financial articles was the frequent appearance of articles related to war and conflict. Although geopolitical events and military conflicts can have substantial economic implications relevant to commodity demand and trade flows, including these articles distorted sentiment scores by introducing a systematically negative bias. Therefore, terms associated with warfare and geopolitical conflicts were excluded to maintain the focus on market-specific sentiment. Excluded keywords are included in table 2:

war	conflict	military
combat	battle	army
navy	air force	soldier
troops	terrorism	militant
insurgent	rebels	warfare
genocide	occupation	conflict zone
armament	bombing	missile
nuclear	strike	battleground
peacekeeping	hostilities	counterinsurgency
terrorist	unrest	geopolitics
political instability	defense industry	
arms trade	weaponry	military-industrial complex
embargo	international security	foreign intervention
peace talks	bailout	casualty
casualties	killed	terror
murder	sanctions	

Table 2: Title keywords used to exclude articles related to war and conflict

Commodity-Specific Financial Articles:

A separate filtering approach was applied to identify articles related to commodities within the large set of financial articles. The commodity categories covered included agricultural products, energy resources, industrial metals, and textiles. The selected keywords are in table 3:

Category	Keywords
Agricultural Commodities	Corn, Soybean, Wheat, Barley, Oats, Rice, Sorghum, Rye, Canola, Cotton, Sugar, Coffee, Cocoa, Tea, Tobacco, Fruits, Vegetables
Livestock and Dairy	Cattle, Hogs, Poultry, Milk, Eggs, Wool, Leather
Energy Commodities	Crude Oil, Natural Gas, Gasoline, Diesel, Jet Fuel, Coal, Uranium, Electricity
Industrial and Precious Metals	Gold, Silver, Copper, Aluminum, Iron Ore, Zinc, Nickel, Lead, Platinum, Palladium, Tin, Lithium, Cobalt, Steel
Other Industrial Materials	Cement, Lumber, Rubber, Glass, Plastic, Sulfur, Phosphates, Potash, Ammonia, Silicon, Graphite, Neodymium, Dysprosium, Yttrium

Table 3: Title keywords used to identify commodity-specific articles

Corn- and Soybean- Specific Articles:

To get relevant information to corn and soybean markets, the following filtering was applied to isolate articles (table 4):

Crop	Keywords
Corn	Corn, Maize, Ethanol
Soybean	Soybean, Biodiesel

Table 4: Title keywords used to identify corn- and soybean-specific articles

Variants of the word "corn" such as "corner" and "corning" were excluded to avoid irrelevant articles. Similarly, the term "soy" was avoided to prevent the inclusion of articles unrelated to soybean markets, such as those related to cooking and dietary discussions. Meanwhile, the inclusion of "ethanol" and "biodiesel" ensured coverage of biofuel-related corn and soybean market discussions.

Sentiment Analysis:

This study uses two sentiment analysis techniques widely used in literature: the Bing Liu Opinion Lexicon and the General Inquirer. These indicators are used to analyze the content of articles and detect words (in their respective ways) related to sentiment according to different dictionaries or lexicons. These words are then assigned to each indicator's positive and negative categories. The count of positive and negative words per article is the output from each indicator.

The Bing Liu Opinion Lexicon, developed by Minqing Hu and Bing Liu in 2004, is a widely used lexicon for sentiment analysis. It contains a total of 6,786 words, with 2,005 classified as positive and 4,781 as negative. The lexicon operates on a binary classification system where each word is either positive or negative. Bing Liu's framework for opinion lexicon-based sentiment scoring provides a systematic approach to evaluate a news article's sentiment words by analyzing associated adjectives related to opinion in their specific context. The lexicon focuses on domain-specific adjectives and nouns extracted through iterative "double propagation" techniques. Each extracted adjective's polarity (positive/negative) is inferred directly from the lexicon word list or using WordNet's semantic relations (Williams and Anand 2009, Zhang and Liu 2011). Moreover, the opinion lexicon handles negative constructs in sentences. For example, "not completely awful" in a sentence would output a mildly positive sentiment rather than an extremely negative one.

The General Inquirer, developed at Harvard University (Stone et al. 1962), is a lexicon that assigns syntactic, semantic, and pragmatic information to words. This is a simpler technique compared to the Bing Liu. General Inquirer matches sentiment words against its curated dictionaries. Lexical categorization involves mapping words to semantic categories using techniques like polysemy resolution for contextual disambiguation (e.g., distinguishing "race" as competition versus ethnicity). The pre-defined General Inquirer dictionary contains 2,291 positive and 1,823 negative words. Unlike Bing Liu, General Inquirer does not use contextual sentiment modifiers and fails to recognize domain-specific sentiment of a word not found in its dictionary. Nor does it account for negative constructs in sentences.

For example, the sentence: *"Stocks plummeted amid chaotic trading, but analysts praised the Fed's decisive intervention*" will have '*plummeted', 'chaotic', 'praised',* and '*decisive'* detected as sentiment words by the Bing Liu indicator. *'Plummeted'* will be paired with *'stocks'* and inferred via WordNet as a negative sentiment word because it is synonymous with the words 'collapse' and 'crash'. '*Chaotic'* will be inferred as negative sentiment because it is a direct match from the Bing Liu's lexicon list. The word *'trading'* comes after the word *'chaotic'* and thus will be modified forming a negative feature-opinion pair. *'Praised'* will be a positive sentiment word according to the Bing Liu lexicon and attributed to *'analysts',* the subject of the sentence. *"Decisive"* is seen as a contextually positive term in financial domains and paired with *'intervention'.* The overall output would be 2 negative words (*plummeted* and *chaotic*) and 2 positive words (*praised* and *decisive*) according to the Bing Liu indicator. Meanwhile, the General Inquirer will classify the words '*plummeted'* (positive)*, 'chaotic'* (positive)*,* and '*praised'* (negative) as sentiment words but will not detect the word '*decisive'.* The overall output would be 2 negative words (*plummeted* and *chaotic*) and 1 positive word (*praised*) according to the General Inquirer indicator.

The Bing Liu Opinion Lexicon excels at detailed article-specific sentiment analysis, while the General Inquirer offers a broader framework for sentiment. Together, Bing Liu Opinion Lexicon and General Inquirer provide robust methodologies for capturing market sentiments and track shifts over time.

Sentiment Score Calculation:

To quantify sentiment captured from a news article, the normalized differences approach is used to calculate a sentiment score using positive and negative words identified by the Bing

Liu and General Inquirer. The normalized difference method is a conventional technique in sentiment analysis for quantifying textual sentiment polarity. Its core formula is defined as:

$$Polarity\ Score = \frac{P - N}{P + N}$$

where P and N represent the counts of positive and negative sentiment-bearing words per article, respectively. This approach yields a bounded score ranging from −1 (maximally negative) to +1 (maximally positive), with a neutral baseline at 0. For instance, a text containing 1 positive and 2 negative terms generates a score of $\frac{1-2}{1+2} = -0.33$, indicating mild negativity.

The formula's denominator $P + N$ inherently normalizes scores relative to the total sentiment-bearing words, enabling length-invariant comparisons between different types of news articles. For example, contrasting a 100-word news report to a 1,000-word opinion article. However, texts lacking both positive and negative terms ($P + N = 0$) produce undefined scores. To mitigate this, a pseudo count adjustment is made to ensure computational stability while minimally influencing the ratio:

$$Polarity\ Score = \frac{P - N}{P + N + \alpha}$$

where $\alpha = 0.00001$

The primary strengths for using normalized differences are the method's intuitive interpretability and standardized measurement scales (from -1 to 1), a critical advantage when analyzing large corpora with diverse document sizes. Moreover, the normalized output enables comparisons across different sentiment analysis techniques. The polarity score indicates the overall direction of the sentiment in an article. Figure 1 shows the monthly averaged time series of the polarity score for all categories of the article sets from 2000 to 2025. They show polarity score trends in each category from leaning towards positivity (>0) or negativity (<0).

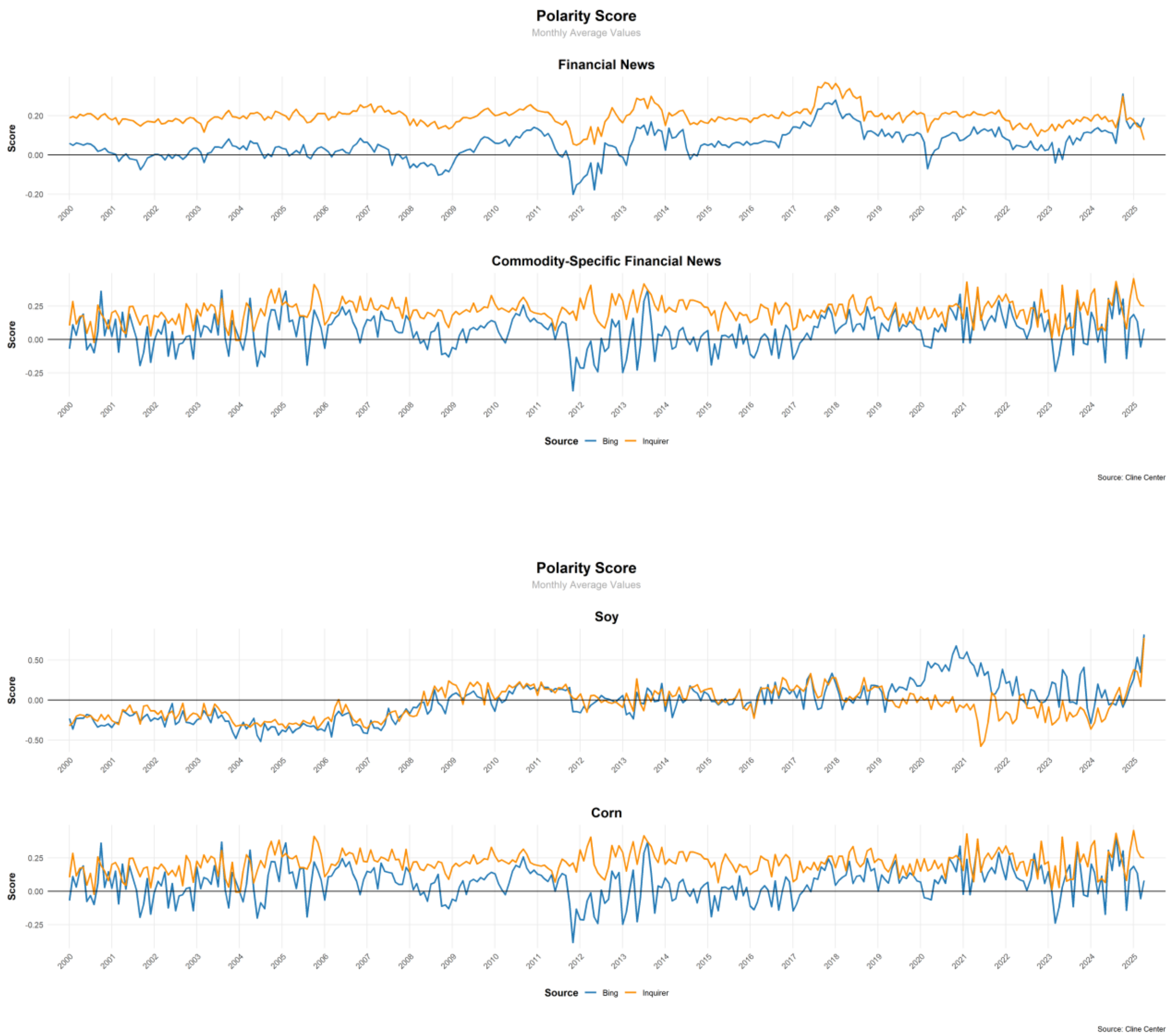

Figure 1: Polarity score in all four categories of article sets (2000-2025)

However, to what extent does a net positive/negative sentiment reflect future market condition expectations in the corn and soybean markets? Polarity (positivity/negativity) might not always align with future price expectations relative to present. For example, an article explaining how "corn markets rallied on expectations of robust ethanol demand and bullish export prospects" will result in a net positive sentiment both from the Bing Liu and General Inquirer indicators. However, there will be a mix of price and spread expectations from different market participants from this news. Corn producers' expectations for the future will rise but it will be the opposite for natural short for corn (e.g. ethanol plants). Similarly, an article about how the "market reacted negatively to USDA's forecast of record corn yields this season" will result in a net negative sentiment from both indicators. But there will be different expectation reactions from market participants (if there is a reaction). Price expectations for corn producers will not be as favorable as compared to corn processing plants. Therefore, we argue that the strength of sentiment is a more relevant indicator for market expectations in the future—not necessarily directionally "good" or "bad" indicators. Thus, we take the absolute value of the polarity score to extract the strength of the net sentiment:

$$Absolute\ Polarity\ Score = \left|\frac{P - N}{P + N + \alpha}\right|$$

However, the normalized difference has a notable limitation regarding word magnitude. Identical ratios (e.g. $\frac{3}{1}\ vs\ \frac{30}{10}$) yield identical scores despite differing absolute counts, potentially masking word length-driven nuances. To address this issue, an *intensity* score is calculated as a robustness check:

$$Intensity\ Score = \frac{P + N}{W}$$

where W is the total word count in an article. The intensity score considers the length of a news article and acts as a robustness check for any variation in sentiment caused by word counts in each news article. Figure 2 shows the monthly averaged time series of the absolute polarity and intensity scores for each category from 2000 to 2025.

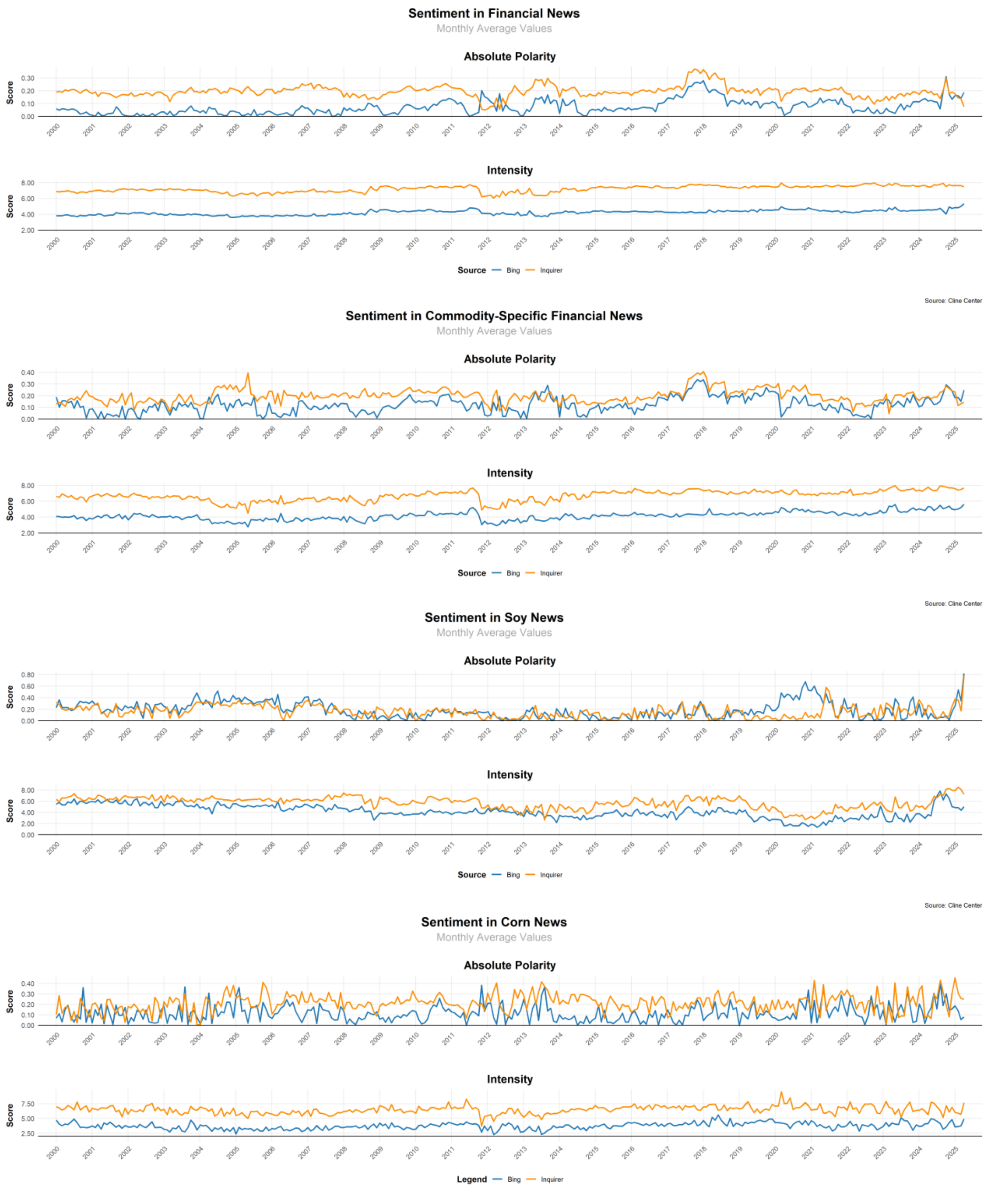

Figure 2: Absolute polarity and intensity scores in all four categories of article sets (2000-2025)

	Soybean News	Corn News	Commodity-Specific Financial News	Financial News
Polarity Score (Bing)				
Obs	29,258	26,009	496,437	11,411,042
Mean	-0.040	0.075	0.131	0.072
Std. dev.	0.530	0.491	0.528	0.512
Min	-1	-1	-1	-1
Max	1	1	1	1
Polarity Score (Inquirer)				
Obs	29,258	26,009	496,437	11,411,042
Mean	-0.046	0.211	0.210	0.204
Std. dev.	0.470	0.346	0.401	0.390
Min	-1	-1	-1	-1
Max	1	1	1	1
Intensity Score (Bing)				
Obs	29,072	25,461	479,226	10,960,372
Mean	4.298	3.830	4.305	4.332
Std. dev.	2.324	1.804	2.358	2.217
Min	0	0	0	0
Max	16.667	18.519	28	60
Intensity Score (Inquirer)				
Obs	29,072	25,461	479,226	10,960,372
Mean	5.940	6.630	6.907	7.343
Std. dev.	2.489	2.384	2.829	2.774
Min	0	0	0	0
Max	22.807	23	50	100

Table 5: Summary statistics for sentiment scores by news category

Spread (Convenience Yield) Calculations:

We consider the convenience yield, "the percentage by which the futures price falls below the value implied by full carrying costs" (Carter, Rausser, and Smith, 2017) as our main measure of the state of a commodity market. We follow the financial literature and use first-nearby futures contract prices to represent the spot market value of corn and soybeans. Relative to available spot market price data, futures prices provide liquid and transparent price benchmarks. The standardized delivery terms and timing for a given futures market make clear the storage horizon between any two particular futures delivery dates.

According to the theory of storage, the cost-of-carry rule denotes that the distant futures price must be less than or equal to the nearby futures prices plus the cost of carry. Given the condition of no arbitrage opportunities, the deferred futures price follow the full-carry model as follows:

$$F_{0,T} = F_{0,t} + C_{t,T}$$

$C_{t,T}$ is the cost of carry including interest rates and fixed storage costs, represented in the following equation:

$$C_{t,T} = F_{0,t}(r_{t,T}) + w_{t,T}$$

where $F_{0,T}$ is the futures price at time 0 for delivery at deferred time T, $F_{0,t}$ is the futures price at time 0 for delivery at nearby time t, $r_{t,T}$ denotes the cost of capital while $w_{t,T}$ represents a fixed cost of physical storage between time t and T.

The deferred futures price is defined as follows:

$$F_{0,T} = F_{0,t}(1 + r_{t,T}) + w_{t,T} + y_{t,T}$$

where $y_{t,T}$ is the convenience yield from nearby time t to deferred time T.

The full-carry model, like Carter, Rausser, and Smith (2017), outputs the convenience yield as the spread between the first-nearby and deferred prices, adjusted for interest rates and storage costs:

$$y_{t,T} = F_{0,T} - [F_{0,t}(1 + r_{t,T}) + w_{t,T}]$$

To facilitate comparisons of convenience yield over time, we calculate convenience yield, the storage-cost-adjusted spread between deferred and nearby futures prices, as a percentage of full cost of carry ($S_{t,T}$). This measure of price of storage captures the component of supply independent of price levels:

$$S_{t,T} = \frac{y_{t,T}}{C_{t,T}} \times 100\%$$

Futures prices for corn and soybean come from Bloomberg's Chicago Board of Trade (CBOT) data. We measure the 1-Year Constant Maturity Treasury rates are used as interest rates adjusted for the number of months between futures maturity dates t and T. Following Carter, Rausser, and Smith (2017), the monthly price of storage is defined as a fixed \$0.05 per bushel adjusted for inflation.

To compare between short-term and long-term spreads, the deferred delivery period consists of two types of contracts: first-deferred delivery contract and the 1 year ahead delivery contract month. For example, the first-nearby contract for corn in January 2025 is the March'25 contract, the first-deferred contract is the May'25 contract, and the 1-year ahead contract is the March'26 contract. To avoid bias related to the maturity of futures contracts (Irwin, Sanders and Yan 2022), the first 15 days of a delivery month is rolled to the next delivery month. Moreover, to distinguish between old and new crop, delivery months close to harvest (September and August for soybean while September for corn) are never considered in the spread calculations. Figure 3 shows the monthly averaged time series of variable $S_{t,T}$ for corn and soybean from 2000 to 2025. The vertical black line denotes the shift between contango (above the line) and backwardation (below the line) in the respective markets.

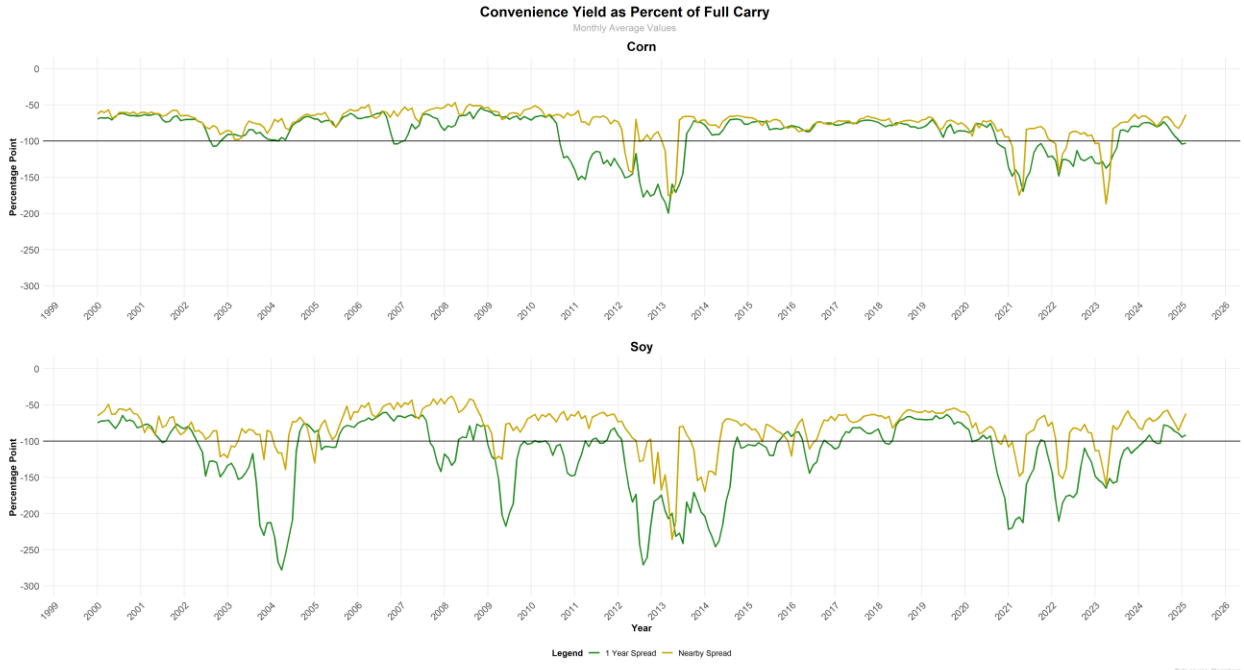

Figure 3: Corn and soybean convenience yield as a percent of full carry (2000-2025)

Implied volatilities for corn and soybeans represent market-driven expectations of future price fluctuations in their respective futures contracts, derived from options pricing. These metrics quantify the anticipated magnitude of price risk over a specified horizon. Like convenience yield, they provide a measure of price uncertainty which captures some forward-looking aspect of expectations. To test the forward-looking nature of news sentiment, implied volatility acts as a robustness check for convenience yield. Historical data for corn and soybean implied volatilities are available from 2012 to 2018, with methodologies capturing four-month deferred contract volatilities to avoid front-month expiration distortions. Figure 4 shows the monthly averaged time series of implied volatilities for corn and soybean sourced from Bloomberg.

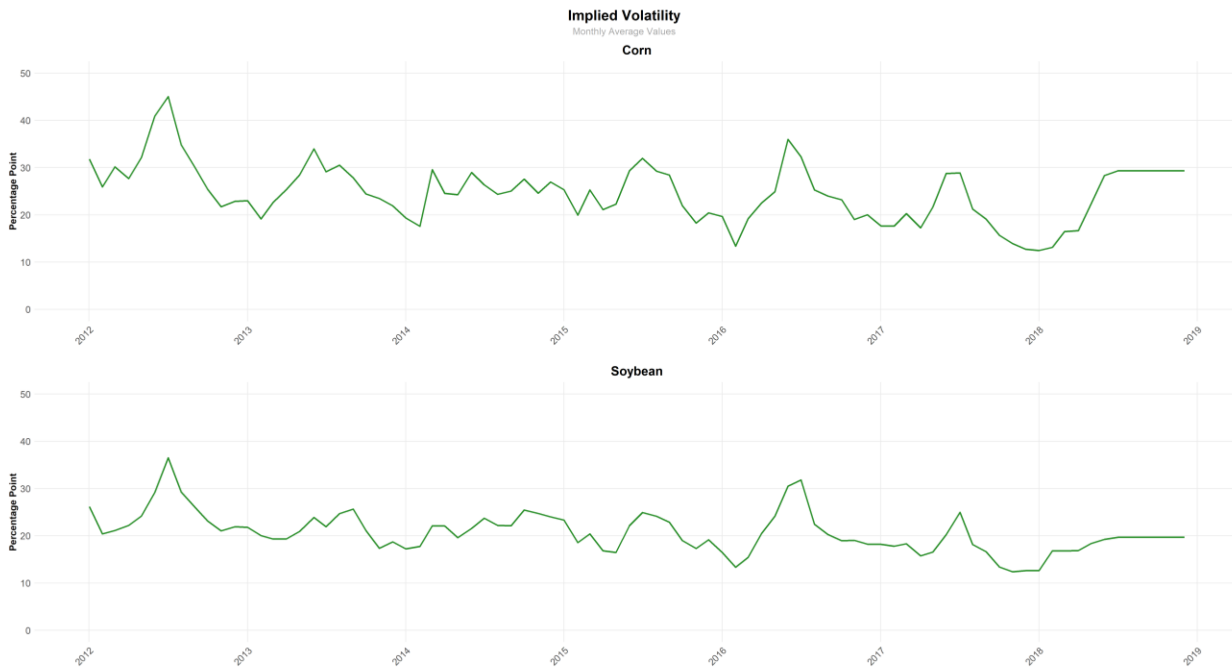

Figure 4: Corn and soybean implied volatility (2012-2018)

Variable	Obs	Mean	Std. Dev.	Min	Max
Soybean Convenience Yield (Nearby)	6,565	-81.559	32.168	-480.305	133.026
Soybean Convenience Yield (1-Year)	6,565	-120.114	50.803	-315.234	-52.436
Corn Convenience Yield (Nearby)	6,565	-75.372	24.351	-340.193	-38.336
Corn Convenience Yield (1-Year)	6,565	-90.453	29.428	-212.746	-51.939
Corn Implied Volatilty	3,620	27.291	5.628	10.400	52.080
Soybean Implied Volatilty	3,620	20.394	3.472	7.730	48.230

Table 6: Summary statistics for convenience yields and implied volatility indices

Contemporaneous Correlations:

Figures 5-10 explore the contemporaneous relationships between convenience yield and the sentiment scores for all four categories of news articles. The vertical black line denotes the shift between contango (above the line) and backwardation (below the line) in the respective markets. The correlations do not provide a clear sense of the relationship between convenience yield and the sentiment scores. But one general trend that can be observed in the nearby convenience yields is that as each sentiment score increases in value, the market tends to be in contango or close to it. This trend is mostly observed in the nearby spreads but not consistently observed in the 1-year convenience yield. Therefore, not much can be concluded from simple correlations alone.

Estimation

To assess the dynamic relationship between sentiment and convenience yield, we use Structural Vector Auto-Regression Model (SVAR). Using impulse response functions from the SVAR estimates allows us to understand the time path of convenience yield responses from isolated news sentiment shocks. This allows us to investigate whether news-based sentiment is forward-looking and whether it helps explain why changes in convenience yield in the first-deferred and the 1-year ahead time periods. Implied volatility indices for corn and soybean are used as a robustness check for whether sentiment is forward-looking in terms of price uncertainty.

Time Series Properties:

The time series properties for each series are checked for stationarity and temporal dependence using auto-correlation functions and augmented Dickey-Fuller tests. If a series has very strong temporal dependence it is hard to assess whether it is related with another series. The figures 11-15 show the auto-correlation functions for each time series. They show that, at the monthly level, the convenience yield series and some sentiment scores contain some seasonality. Generally, all series show a gradual decay in the monthly lags indicating auto-regressive processes and stationarity. However, that is not the case in daily and weekly lag levels. The augmented Dickey-Fuller unit root tests in tables 7 and 8 show that each series follows a stationary process.

Structural Vector Auto-Regression Model:

Our main hypothesis is that unexpected deviations in the strength and/or intensity of market-relevant news sentiment leads to immediate responses in future expectations of market conditions. In other words, markets are efficient in capturing new information (Fama 1970). We also hypothesize that a positive shock in news sentiment does not necessarily entail a consistent response in expectations. The direction of the response may change with the time horizon of expectations (nearby or 1-year convenience yield).

We estimate a structural vector auto-regression (SVAR) model given the following equation:

$$y_t = \begin{bmatrix} N_t \\ S_t \end{bmatrix} = \beta y_{t-n} + \delta X_t + \varepsilon_t$$

Where $N_t$is the news sentiment, $S_t$is convenience yield, and X_t represent the seasonal monthly dummy variables at time t.

We select the number of lags n following the Akaike information criterion (AIC). The ordering assumption in the above equation is that news sentiment shocks precede convenience yield shocks. This assumption is reasonable for this study, in part, because it provides an upper bound on the size of the impact of a news sentiment shock on convenience yield. Moreover, this assumption allows us to test whether news-sentiment is forward-looking.

Results

Impulse response functions are derived from SVAR estimates in order to present the time path of responses in the dependent variable from shocks in the independent variable. To assess whether news sentiment shocks to immediate, consistent, and significant responses, impulse response functions are presented in the following sections for convenience yield and implied volatility.

Impulse Response Functions: Convenience Yield

If news sentiment is forward-looking, we expect to find an immediate response in convenience yield consistently from market relevant news shocks. The direction of the response may change with the time horizon of expectations (nearby or 1-year convenience yield). The figures 16-21 show the impulse response functions of corn and soybean convenience yields to news sentiment shocks from all four categories (financial, commodity-specific financial, soybean-specific, and corn-specific) for both Bing Liu and General Inquirer sentiment scores. These responses are estimated for both the nearby and 1-year convenience yield time horizons coming from shocks in both absolute polarity and intensity scores. The responses are consistently non-significant in both the short-term and long-term except for a few exceptions. In the few cases where responses are significant, the time delay for the responses are around 7 months and above. The exceptional responses are positively significant, and they appear the strongest in the soybean convenience yield responding to soybean-specific sentiment (figure 20). This suggests that soybean-specific news sentiment shocks create positive responses in soybean convenience yield (both in the nearby and 1-year time horizons) indicating a move towards a contango/normal market. However, this significant response is only seen coming from the General Inquirer's absolute polarity and intensity scores, but none form the Bing Liu scores. Overall, the results do not show any significant, immediate, and consistent convenience yield responses to news sentiment shocks.

Impulse Response Functions: Implied Volatility

The 4-month implied volatility measures the price uncertainty anticipated in the next 4 months on a given date. Implied volatility indices for corn and soybean are used as a robustness check to address the question of whether sentiment is forward-looking in nature by using a different market outcome.

If news sentiment is forward-looking, we expect to find an immediate response in implied volatility consistently from market relevant news shocks. The figures 22-27 below show the impulse response functions of corn and soybean implied volatilities to news sentiment shocks from all four categories (financial, commodity-specific financial, soybean-specific, and corn-specific) for both Bing Liu and General Inquirer sentiment scores. These responses are estimated for shocks in both absolute polarity and intensity scores. The results indicate more significant, immediate, and consistent responses in the implied volatilities compared to the convenience yields. Implied volatility tends to respond negatively to shocks in news sentiment, specifically in financial and soybean-specific news shocks. Figures 22, 23, and 26 shows that absolute polarity shocks in financial news coming from the General Inquirer introduces a significant and immediate (1-3 months ahead) decrease in the implied volatilities of both corn and soybean markets by 0.5 to 1.5 percentage points. This suggests that financial news shocks in the strength of the sentiment leads to stabler expectations in the nearby months in the soybean and corn markets. Similarly, soybean implied volatility responds negatively by 0.5 to 1 percentage points to both commodity-specific and soybean-specific intensity score shocks. This suggests that an increase in the intensity of soybean and commodity news leads to lower volatility expectation in the nearby months. Meanwhile, corn does not exhibit any consistent response to any type of news shock other than financial news.

Conclusion

Seasonal commodity markets are driven by expectations of forthcoming market conditions. Convenience yield acts as a negative cost (i.e. benefit) of holding a physical commodity and causes inverted markets or backwardation. The theory of storage relies on convenience yield to explain variation in relative prices that cannot be explained by changes in storage costs. This is the basis of the 'Working Curve,' which has been empirically documented for storable commodities using physical inventory data and futures spreads (Working 1949, Telser 1958, Brennan 1958, Joseph, Irwin, and Garcia 2016). However, the intra-year dynamics of convenience yield is difficult to explore using inventories data which are typically observed annually. Understanding intra-year convenience yield dynamics in commodity markets is what motivates this work, which is not seen in previous literature. This approach builds on the growing body of literature where sentiment analysis has been applied to provide forecasting power in economic indicators and to help measure specific economic parameters derived from economic theory. This study introduces a novel approach to analyzing convenience yield dynamics in commodity markets by leveraging sentiment analysis, enabling a more granular understanding of convenience yield dynamics in corn and soybean markets. The central premise is that market sentiment, as captured from news articles, reflects forward-looking expectations regarding supply, demand, and price levels. The research questions we are trying to answer are twofold: how forward-looking and efficient is the information captured from news sentiment for storable commodity prices, and can it help explain convenience yield dynamics in these markets?

Using impulse response functions from structural VAR for corn and soybean markets, the results do not show any significant, immediate, and consistent convenience yield responses to news sentiment shocks while more significant, immediate, and consistent responses in the implied volatilities. Implied volatility tends to respond negatively to shocks in news sentiment, specifically in financial and soybean-specific news shocks. Absolute polarity score shocks in financial news coming from the General Inquirer introduces a decrease in corn and soybean implied volatilities by 0.5 to 1.5 percentage points 1-3 months ahead. This suggests that financial news shocks in the strength of the sentiment leads to stabler expectations in the nearby months in the soybean and corn markets. Similarly, soybean implied volatility responds negatively by 0.5 to 1 percentage points to both commodity-specific and soybean-specific intensity score shocks. This suggests that an increase in the intensity of soybean and commodity news leads to lower volatility expectation in the nearby months. Meanwhile, corn does not exhibit any consistent response to any type of news shock other than financial news.

This study leverages sentiment analysis to enable a granular understanding of intra-year convenience yield expectations. In conclusion, news does not seem to provide new information that is not already captured in future prices, both at the short-term (nearby) and long-term (1-year) time horizons. This suggests commodity markets are very efficient in capturing new information (Fama 1970) and news sentiment is not forward-looking. Moreover, news sentiment does not provide forward-looking information regarding convenience yield dynamics but provides some information regarding price uncertainty 1-3 months ahead.

Convenience yield dynamics are difficult to measure directly using news sentiment analysis. Future studies in this area may include using large language models to process the content of each article to categorize relevant information to convenience yield e.g. determining type of shock (demand- and/or supply-side), time horizon of information, source of news etc.

References

Adams, Travis, Andrea Ajello, Diego Silva, and Francisco Vazquez-Grande (2023). More than Words: Twitter Chatter and Financial Market Sentiment, Finance and Economics Discussion Series 2023-034. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2023.034.

Baker, S. R., Bloom, N., & Davis, S. J. (2016). Measuring economic policy uncertainty. The Quarterly Journal of Economics, 131(4), 1593–1636. https://doi.org/10.1093/qje/qjw024

Benhabib, J., & Spiegel, M. M. (2019). Sentiments and economic activity: Evidence from US states. The Economic Journal, 129(618), 715–733. https://doi.org/10.1111/ecoj.12605

Brennan, M. J. (1958). The supply of storage. The American Economic Review, 48(1), 50–72.

Carroll, C. D., Fuhrer, J. C., & Wilcox, D. W. (1994). Does consumer sentiment forecast household spending? If so, why? American Economic Review, 84(5), 1397–1408. https://doi.org/10.2307/2117779

Carter, C.A., Rausser, G.C. and Smith, A., 2017. Commodity storage and the market effects of biofuel policies. American Journal of Agricultural Economics, 99(4), pp.1027-1055.

Chang, S., & Suk, D. Y. (1998). Failed takeovers, methods of payment, and bidder returns. The Financial Review, 33(2), 19–34. https://doi.org/10.1111/j.1540-6288.1998.tb01366.x

Costola, M., Hinz, O., Nofer, M., & Pelizzon, L. (2023). Machine learning sentiment analysis, COVID-19 news and stock market reactions. Review of International Business and Finance, 64, Article 101881. https://doi.org/10.1016/j.ribaf.2023.101881

Du, X., & Dong, F. (2023). Agricultural policy uncertainty and its impact on commodity markets. Journal of the Agricultural and Applied Economics Association, 2(2), 263–277. https://doi.org/10.1002/jaa2.56

Engelberg, J. (2008). Costly Information Processing: Evidence from Earnings Announcements. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1107998

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. The Journal of Finance, 25(2), 383–417. https://doi.org/10.2307/2325486

Fan, J. H., Binnewies, S., & De Silva, S. (2023). Wisdom of crowds and commodity pricing. Journal of Futures Markets, 43(8), 1040–1068. https://doi.org/10.1002/fut.22393

Feng, L., Fu, T., & Shi, Y. (2022). How does news sentiment affect the states of Japanese stock return volatility? International Review of Financial Analysis, 84, 102267. https://doi.org/10.1016/j.irfa.2022.102267

Frydman, R., Mangee, N., & Stillwagon, J. R. (2020, April 21). How market sentiment drives forecasts of stock returns (Institute for New Economic Thinking Working Paper Series No. 115). Institute for New Economic Thinking. https://doi.org/10.36687/inetwp115

Garcia, D. (2012). Sentiment during recessions. Journal of Finance, Forthcoming. https://doi.org/10.2139/ssrn.1571101

Ho, Kin-Yip & Shi, Yanlin & Zhang, Zhaoyong, 2013. "How does news sentiment impact asset volatility? Evidence from long memory and regime-switching approaches," The North American Journal of Economics and Finance, Elsevier, vol. 26(C), pages 436-456.

Hsu, Y.-J., Lu, Y.-C., & Yang, J. J. (2021). News sentiment and stock market volatility. Review of Quantitative Finance and Accounting, 57(3), 1093–1122. https://doi.org/10.1007/s11156-021-00971-8

Irwin, S. H., Sanders, D. R., & Yan, L. (2022). The order flow cost of index rolling in commodity futures markets. Applied Economic Perspectives and Policy, 44(3), 1212-1234. https://doi.org/10.1002/aepp.13297

Joseph, K., Irwin, S.H. and Garcia, P., 2016. Commodity storage under backwardation: Does the working curve still work?. Applied Economic Perspectives and Policy, 38(1), pp.152-173.

Kaldor, N. (1939). Speculation and economic stability. The Review of Economic Studies, 7(1), 1–27. https://doi.org/10.2307/2967593

Klibanoff, P., Lamont, O., & Wizman, T. A. (1998). Investor Reaction to Salient News in Closed-End Country Funds. The Journal of Finance, 53(2), 673–699. Retrieved June 17, 2023, from https://www.jstor.org/stable/117365

Konnikova, M. (2014, December 17). How headlines change the way we think. The New Yorker. https://www.newyorker.com/science/maria-konnikova/headlines-change-way-think

Larsen, V. H., & Thorsrud, L. A. (2019). The value of news for economic developments. Journal of Econometrics, 209(1), 162-182. https://doi.org/10.1016/j.jeconom.2018.11.013

Lemarié, J., Lorch, R. F., Jr., & Péry-Woodley, M.-P. (2018). Understanding how headings influence text processing. Discours, 22. https://doi.org/10.4000/discours.8600

Maslyuk-Escobedo, S., Rotaru, K., & Dokumentov, A. (2017). News sentiment and jumps in energy spot and futures markets. Pacific-Basin Finance Journal, 45, 186–210. https://doi.org/10.1016/j.pacfin.2016.07.001

Shapiro, A.H. and Wilson, D.J., 2022. Taking the Fed at its word: A new approach to estimating central bank objectives using text analysis. The Review of Economic Studies, 89(5), pp.2768-2805.

Smales, L. A. (2014). News sentiment in the gold futures market. Journal of Banking & Finance, 48, 318-330. https://doi.org/10.1016/j.jbankfin.2014.09.006

Soo, C. (2015, October). Quantifying animal spirits: News media and sentiment in the housing market (Ross School of Business Paper No. 1200). Available at SSRN: https://ssrn.com/abstract=2330392 or http://dx.doi.org/10.2139/ssrn.2330392

Stone, P. J., Bales, R. F., Namenwirth, J. Z. and Ogilvie, D. M. (1962), The general inquirer: A computer system for content analysis and retrieval based on the sentence as a unit of information. Syst. Res., 7: 484-498. doi:10.1002/bs.3830070412

Telser, L. G. (1958). Futures trading and the storage of cotton and wheat. Journal of Political Economy, 66(3), 233–255. https://www.jstor.org/stable/1833216

Tetlock, P. C., Saar-Tsechansky, M., & Macskassy, S. (2008). More than words: Quantifying language to measure firms' fundamentals. Journal of Finance, 63(3), 1437–1467. https://doi.org/10.1111/j.1540-6261.2008.01362.x

Warren, N. L., & Sorescu, A. (2017). Interpreting the stock returns to new product announcements: How the past shapes investors' expectations of the future. Journal of Marketing Research, 54(5), 799–815. https://doi.org/10.1509/jmr.14.0119

Williams, G. K., & Anand, S. S. (2009). Predicting the polarity strength of adjectives using WordNet. Department of Computer Science, University of Warwick, United Kingdom.

Working, H. (1949). The theory of price of storage. The American Economic Review, 39(6), 1254–1262. American Economic Association. https://www.jstor.org/stable/1816601

Zhang, L., & Liu, B. (2011). Identifying noun article features that imply opinions. Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Short Papers (pp. 575–580). Association for Computational Linguistics.

Zhou, K. (2025). Professional investors and media narratives. Working Paper. Retrieved from https://kangyingzhou.com/files/JMP_KangyingZhou.pdf

Appendix

Correlations

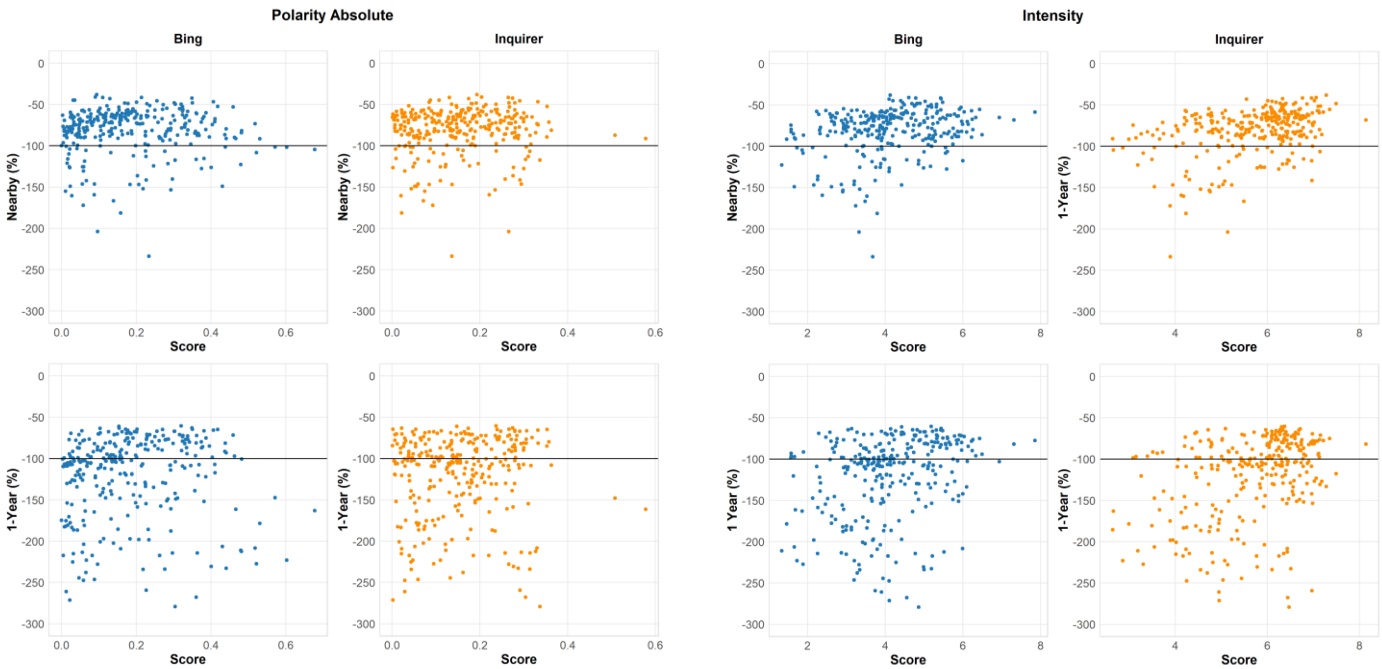

Figure 5: Soybean convenience yield and soybean-specific sentiment scores

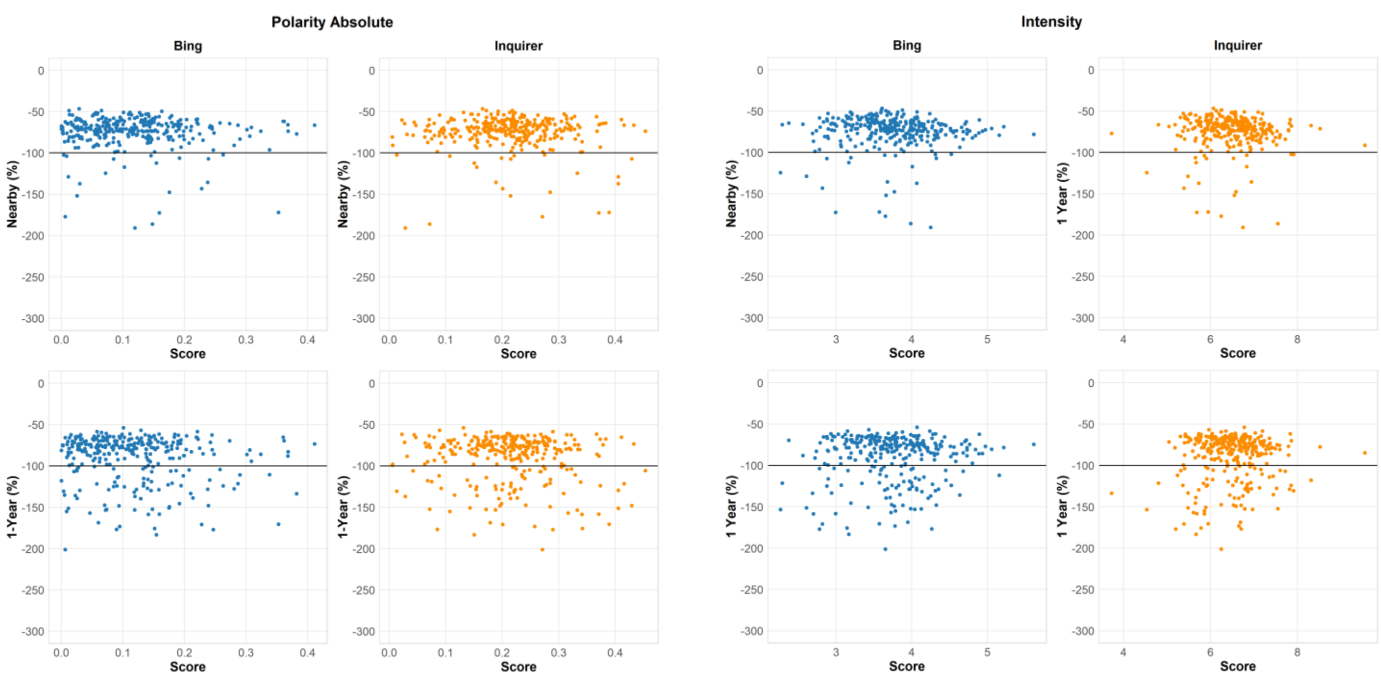

Figure 6: Corn convenience yield and corn-specific sentiment scores

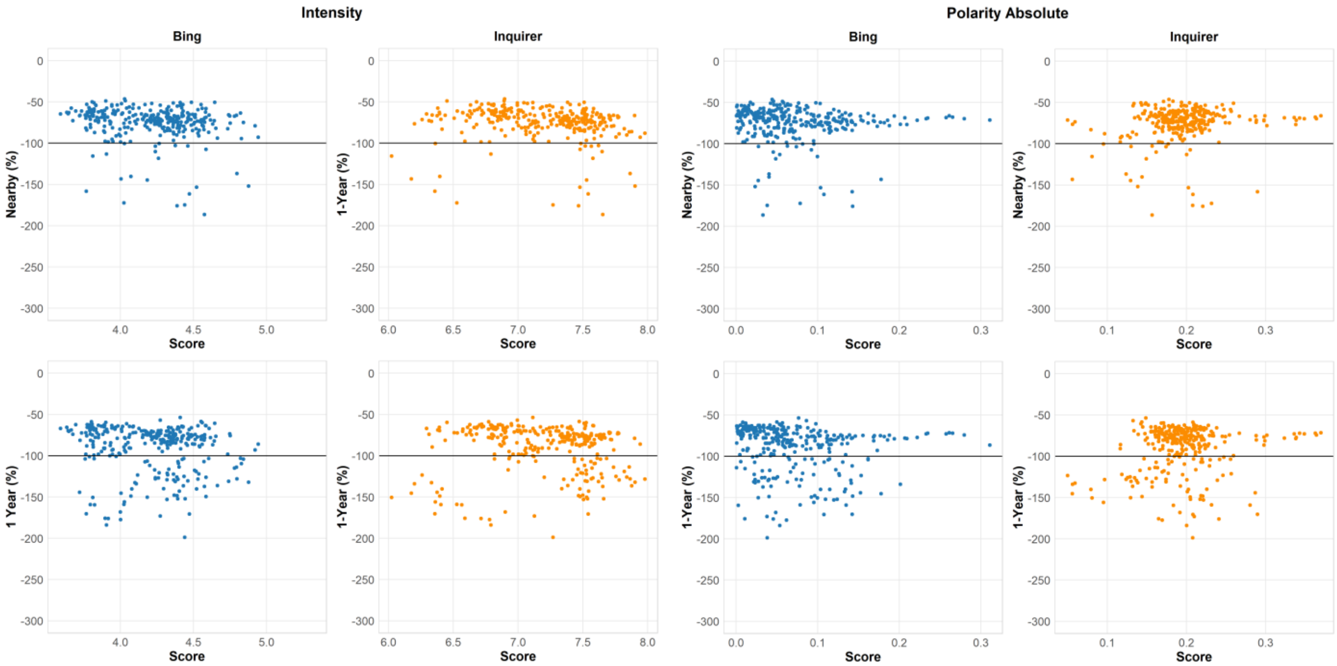

Figure 7: Corn convenience yield and financial sentiment scores

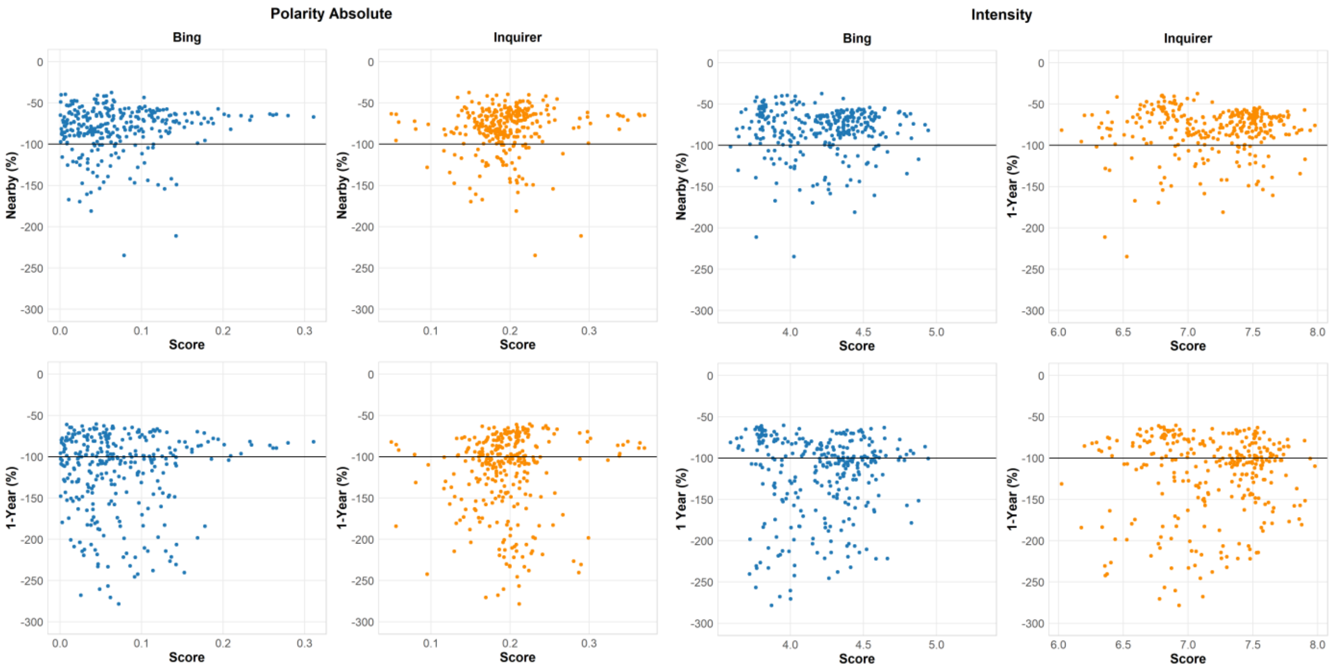

Figure 8: Soybean convenience yield and financial sentiment scores

Figure 9: Soybean convenience yield and commodity-specific financial sentiment scores

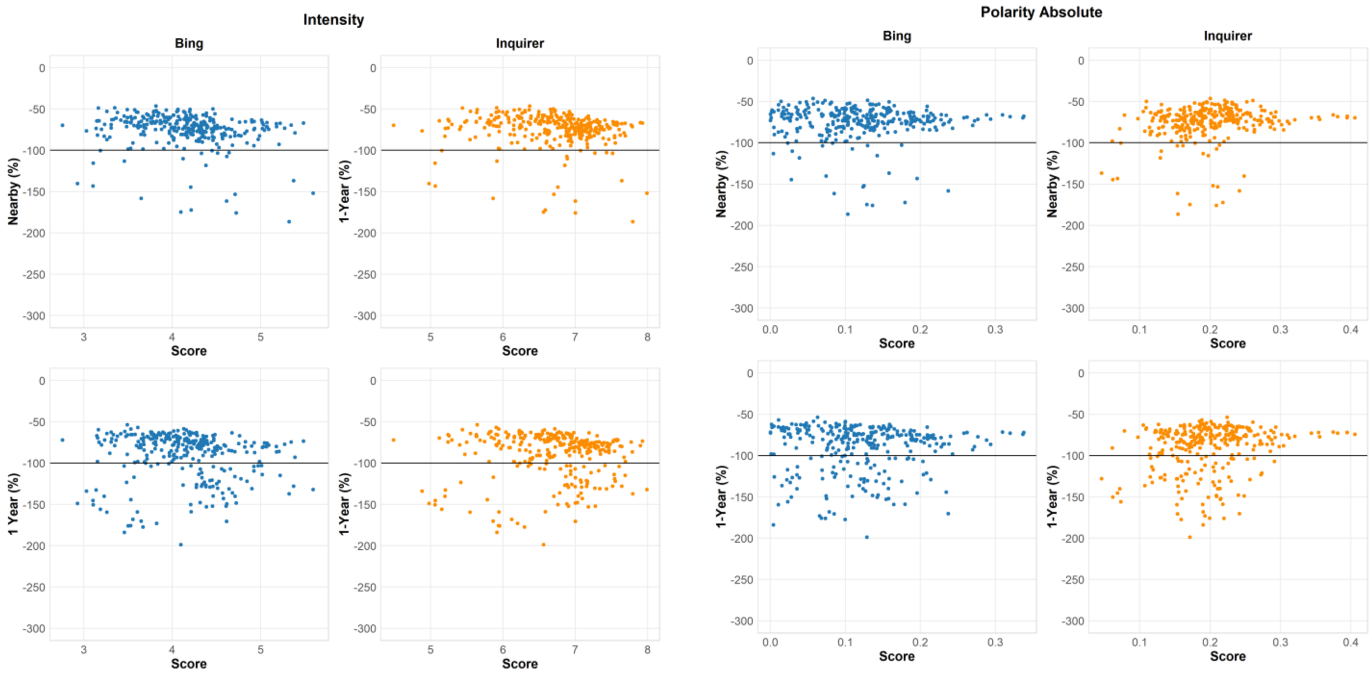

Figure 10: Corn convenience yield and commodity-specific financial sentiment scores

Auto-Correlation Functions

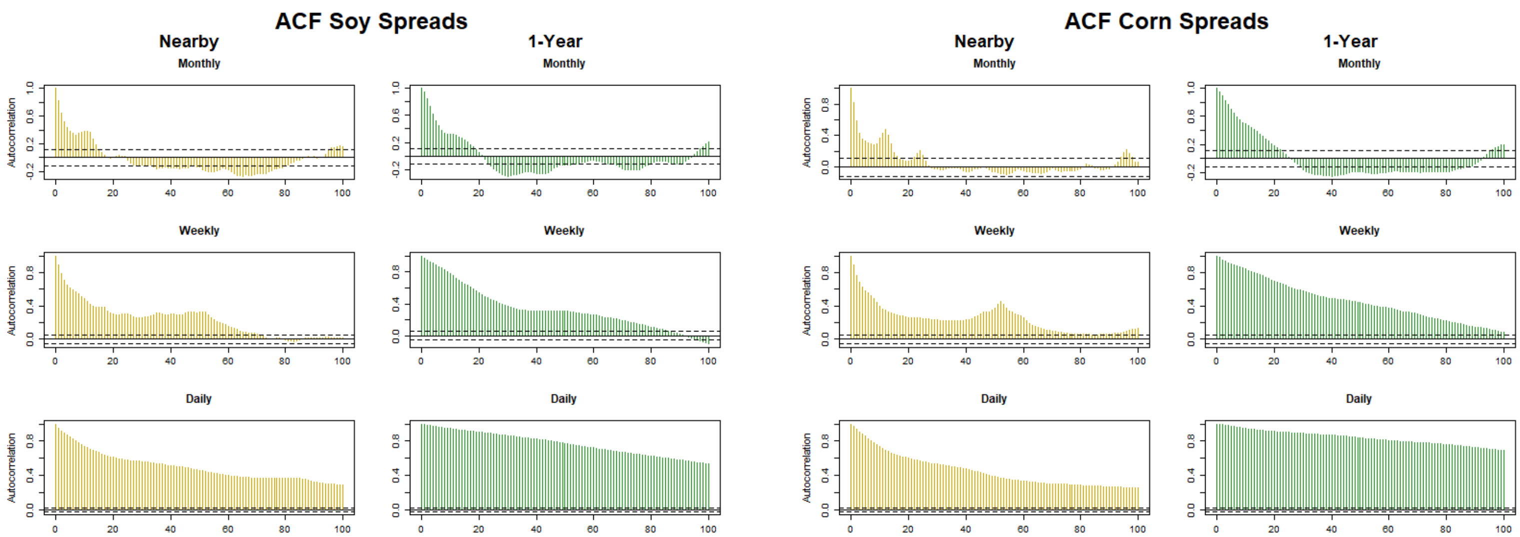

Figure 11: Auto-correlation plots for soybean and corn convenience yields at the monthly, weekly, and daily levels.

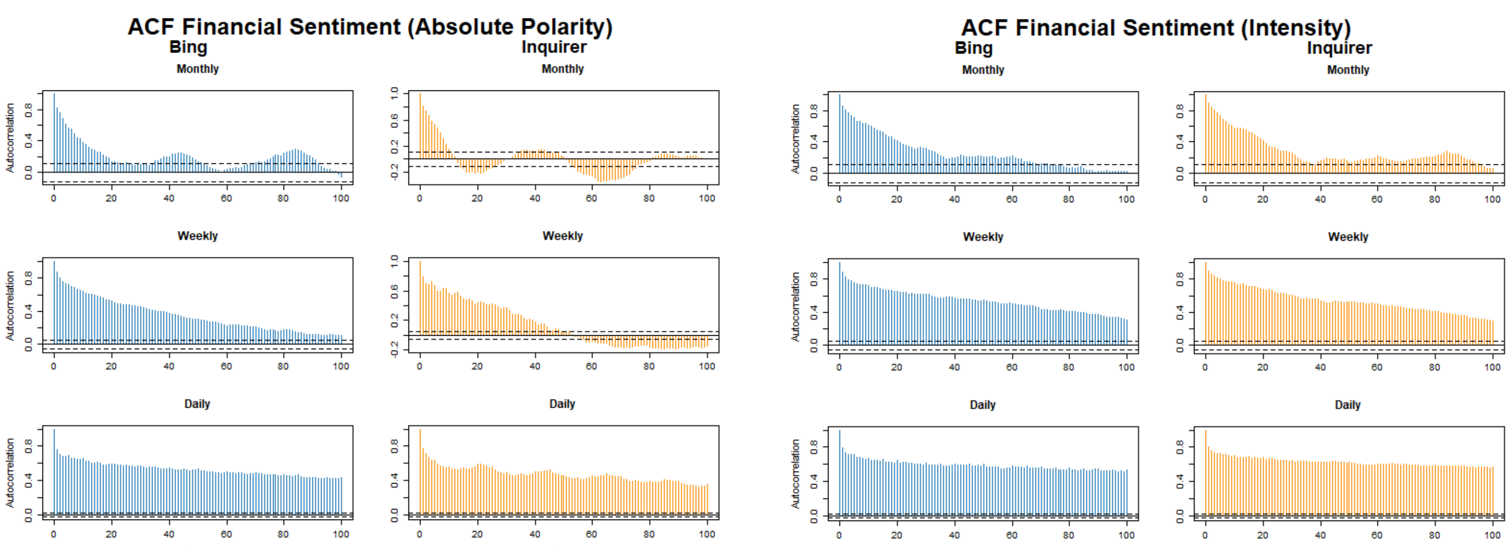

Figure 12: Auto-correlation plots for financial sentiment scores at the monthly, weekly, and daily levels.

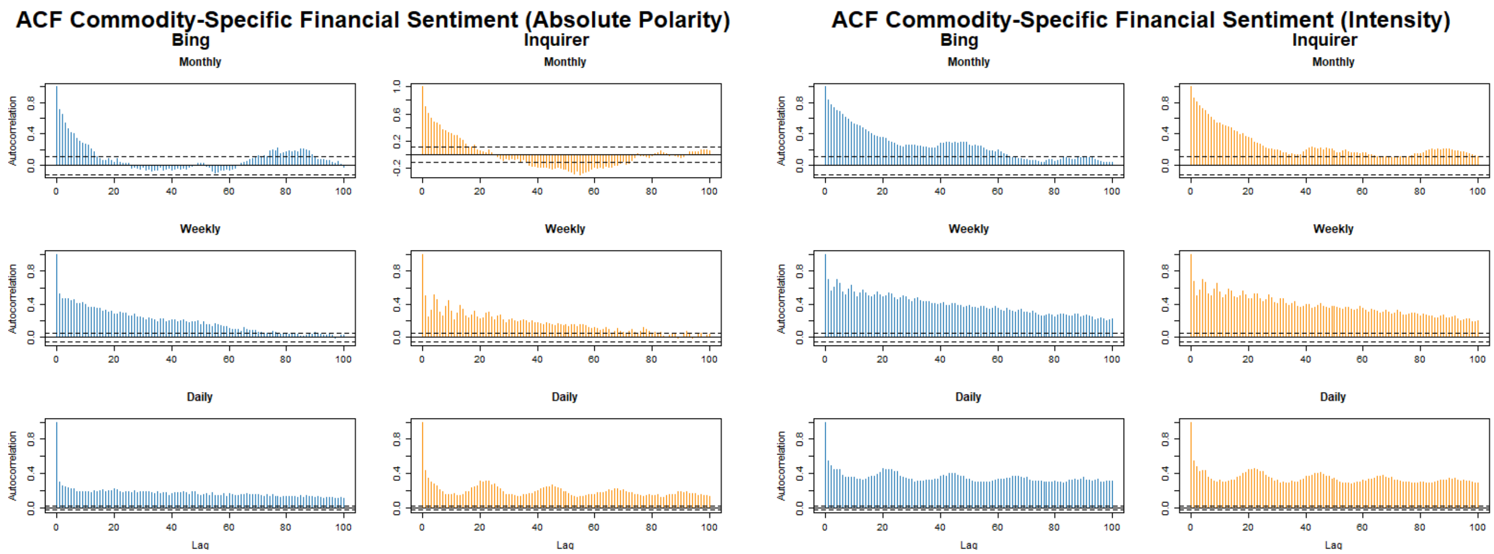

Figure 13: Auto-correlation plots for commodity-specific financial sentiment scores at the monthly, weekly, and daily levels.

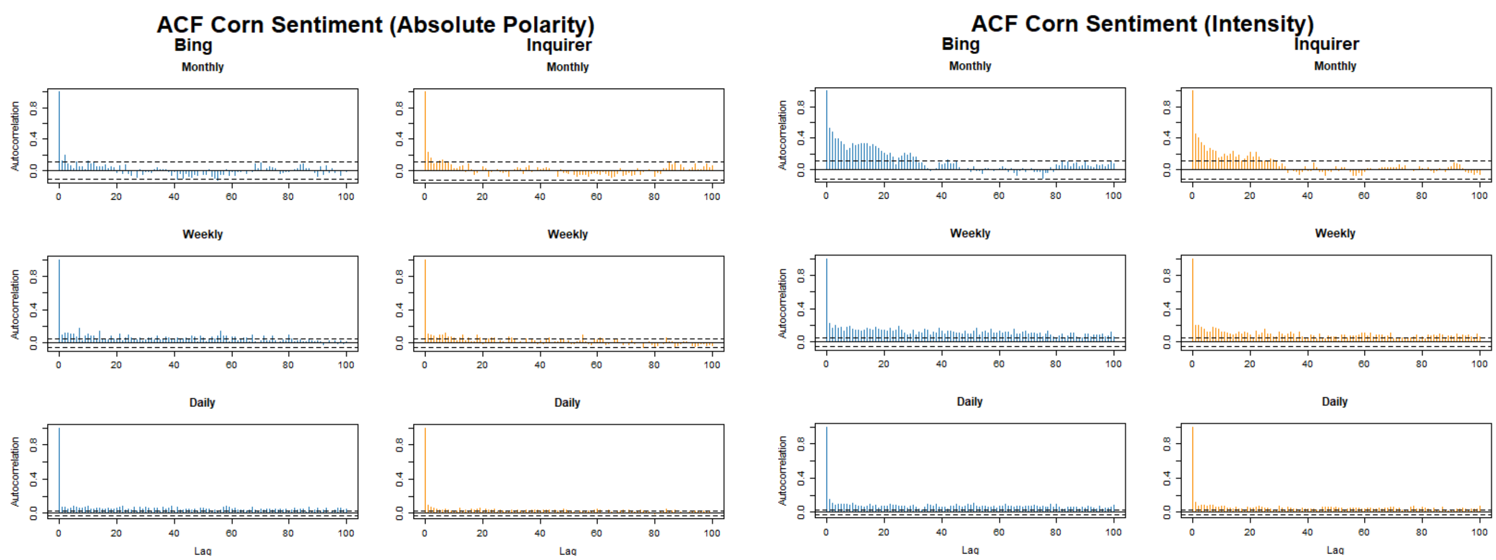

Figure 14: Auto-correlation plots for corn-specific sentiment scores at the monthly, weekly, and daily levels.

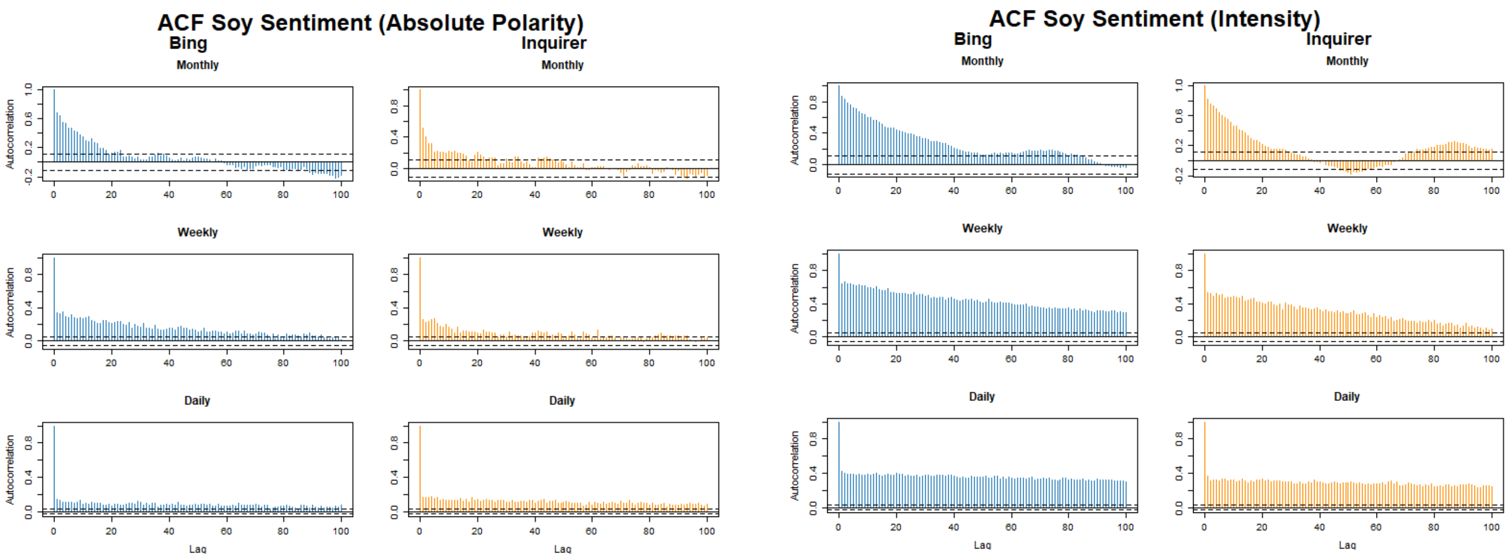

Figure 15: Auto-correlation plots for soybean-specific sentiment scores at the monthly, weekly, and daily levels

Augmented-Dickey Fuller Tests

Variable	Test Statistic
Financial News	
Absolute Polarity Score (Bing)	-2.50*
Absolute Polarity Score (Inquirer)	-3.26**
Intensity Score (Bing)	-2.58*
Intensity Score (Inquirer)	-3.00**
Commodity-Specific Financial News	
Absolute Polarity Score (Bing)	-3.22**
Absolute Polarity Score (Inquirer)	-2.78*
Intensity Score (Bing)	-1.99*
Intensity Score (Inquirer)	-2.70**
Corn News	
Absolute Polarity Score (Bing)	-5.38***
Absolute Polarity Score (Inquirer)	-7.51***
Intensity Score (Bing)	-4.57***
Intensity Score (Inquirer)	-4.54***
Soybean News	
Absolute Polarity Score (Bing)	-2.82**
Absolute Polarity Score (Inquirer)	-3.88***
Intensity Score (Bing)	-2.97**
Intensity Score (Inquirer)	-2.33*

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$
Test specification includes drift and lagged differences.

Table 7: Augmented Dickey-Fuller test: sentiment scores

Variable	Test Statistic (τ)
Corn	
Nearby Convenience Yield	-6.02***
1-Year Convenience Yield	-2.63**
Implied Volatility	-3.8388***
Soybean	
Nearby Convenience Yield	-4.42***
1-Year Convenience Yield	-3.18**
Implied Volatility	-4.44***

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$
Test specification includes drift and lagged differences.

Table 8: Augmented Dickey-Fuller test: convenience yield, implied volatility

Impulse Response Functions

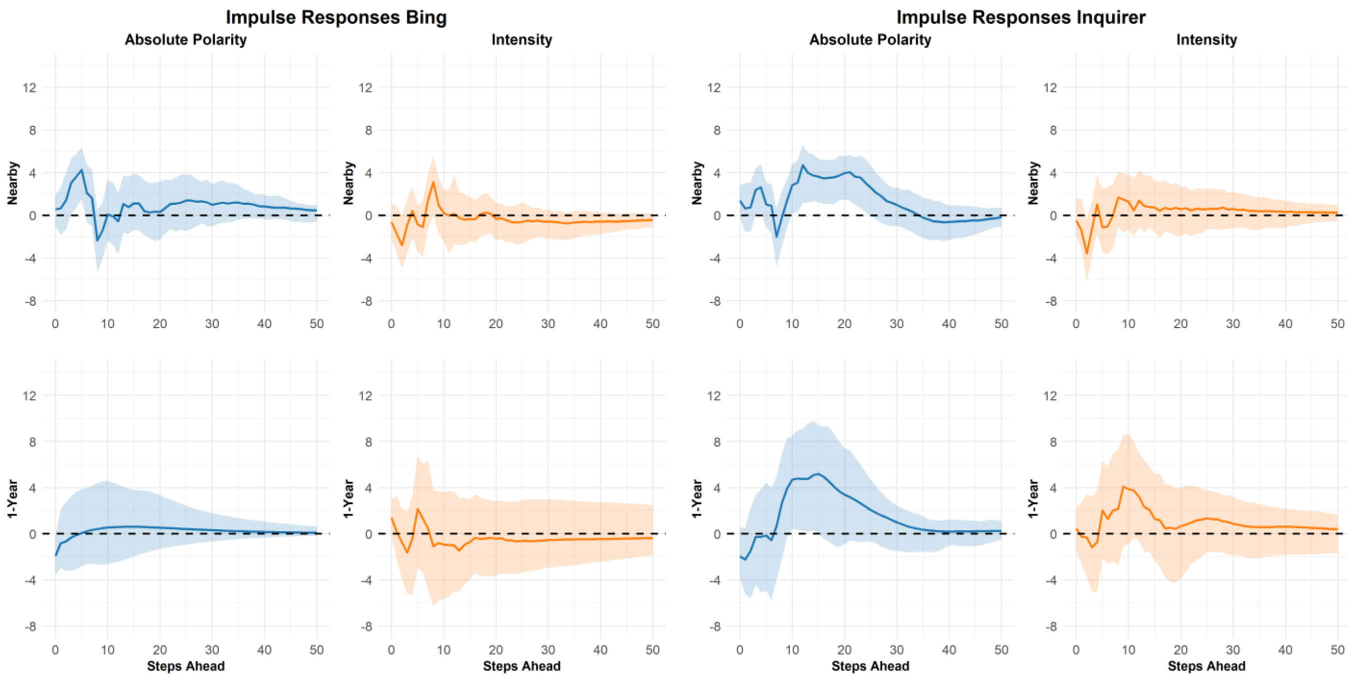

Figure 16: Soybean convenience yield (in percentage points) responses to a financial sentiment shock

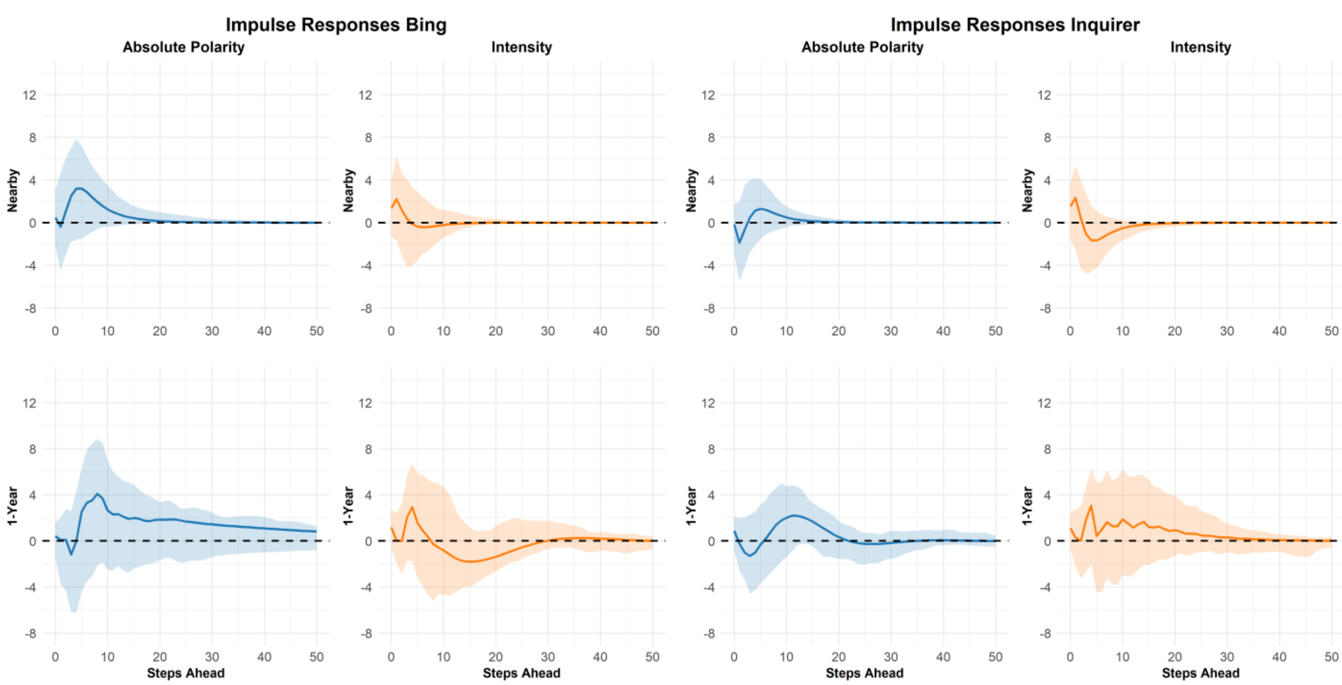

Figure 17: Corn convenience yield (in percentage points) responses to a financial sentiment shock

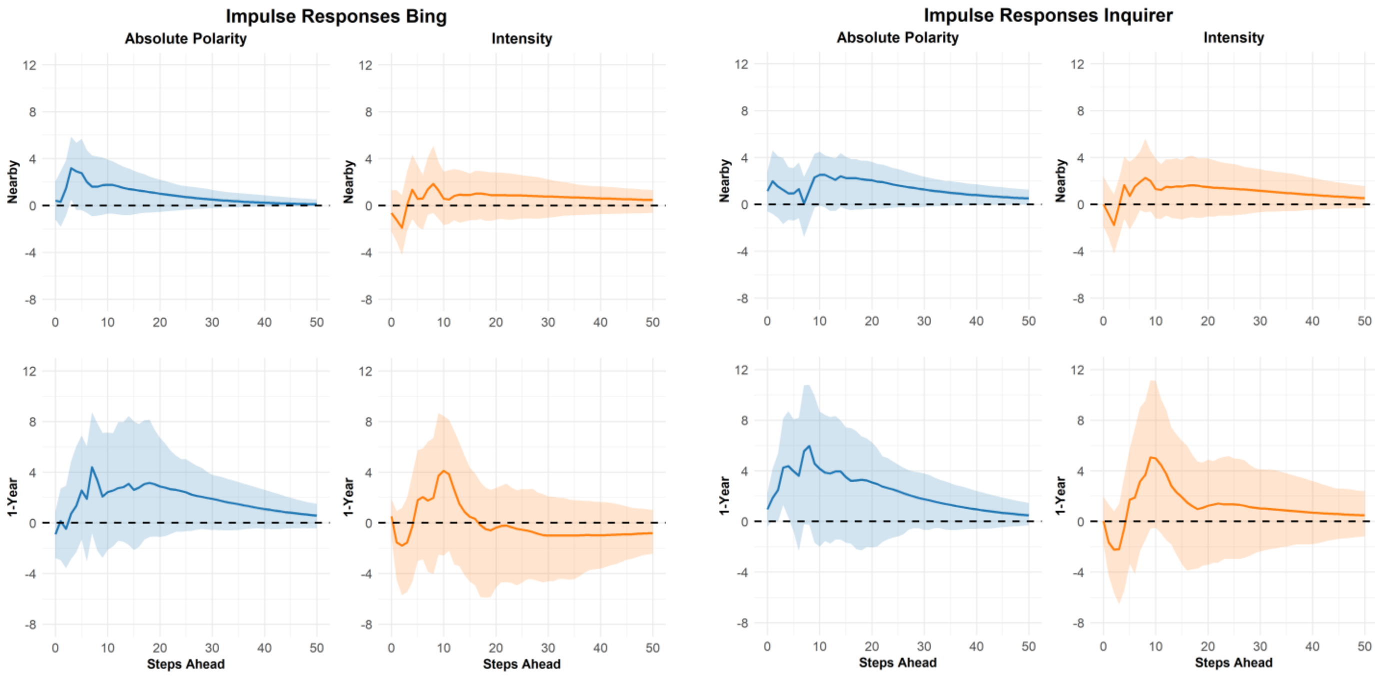

Figure 18: Soybean convenience yield (in percentage points) responses to a commodity-specific financial sentiment shock

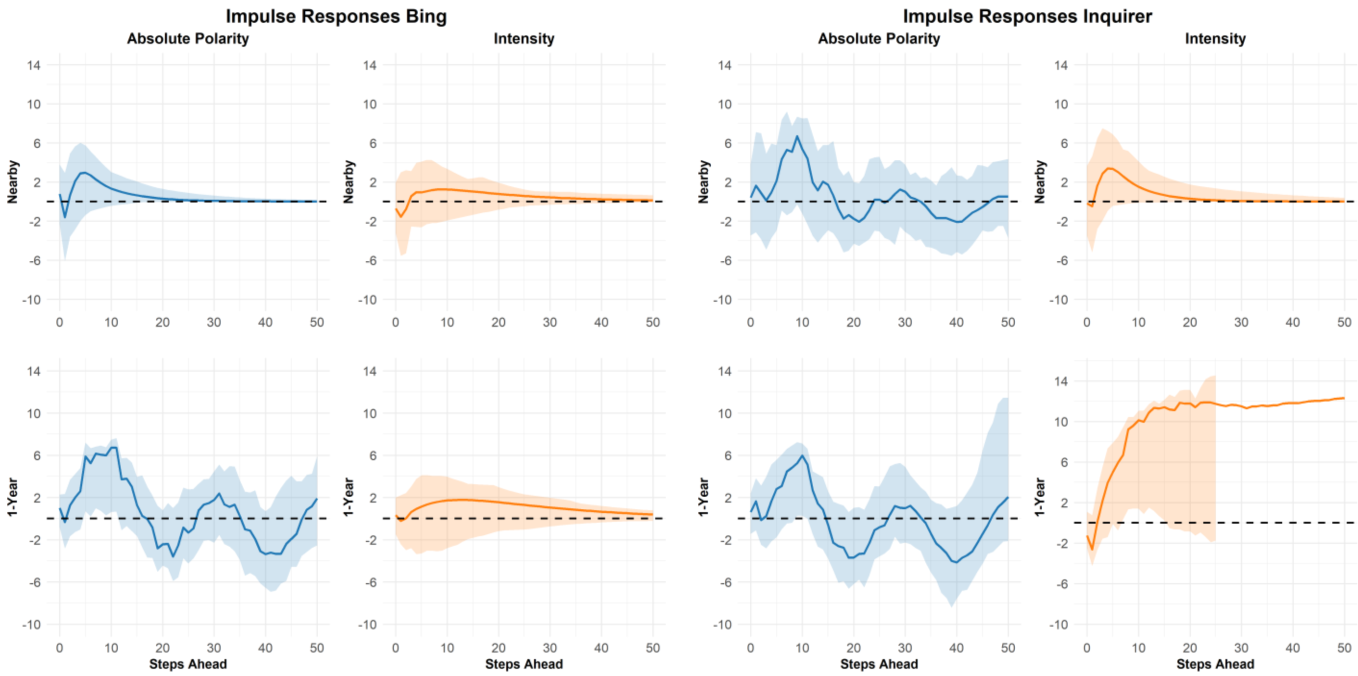

Figure 19: Corn convenience yield (in percentage points) responses to a commodity-specific financial sentiment shock

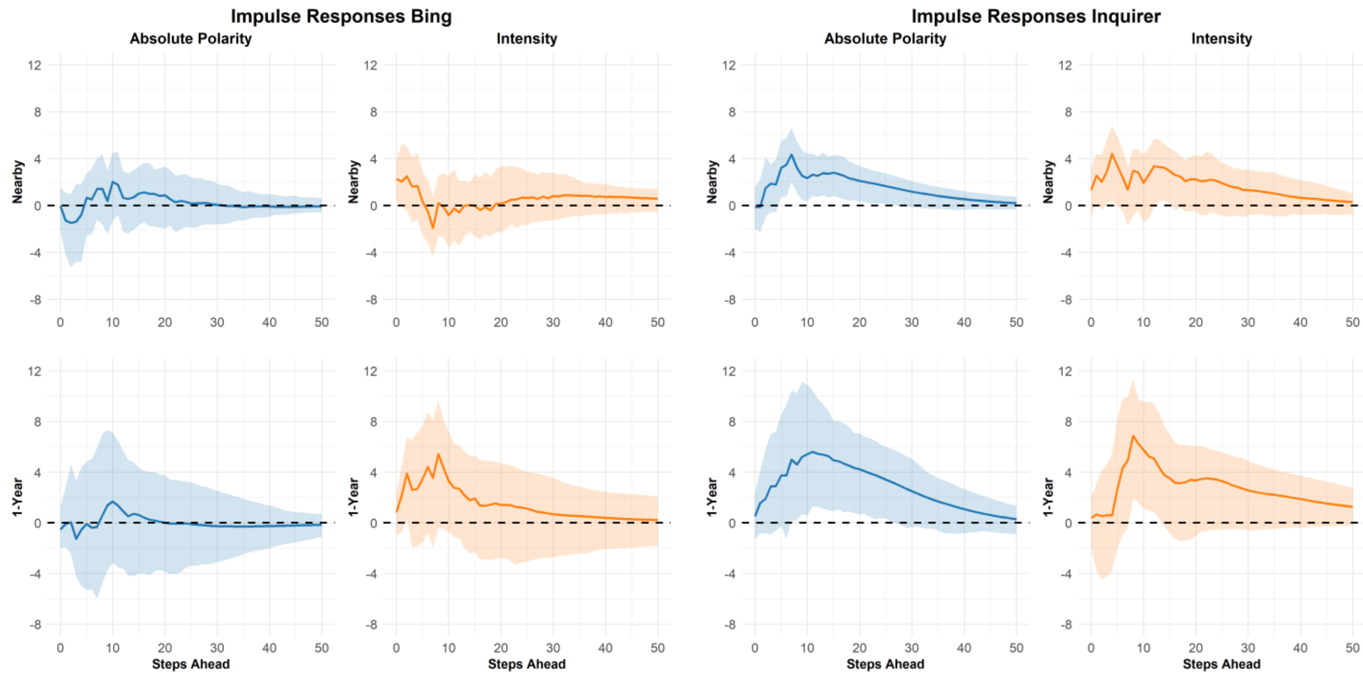

Figure 20: Soybean convenience yield (in percentage points) responses to a soybean-specific sentiment shock

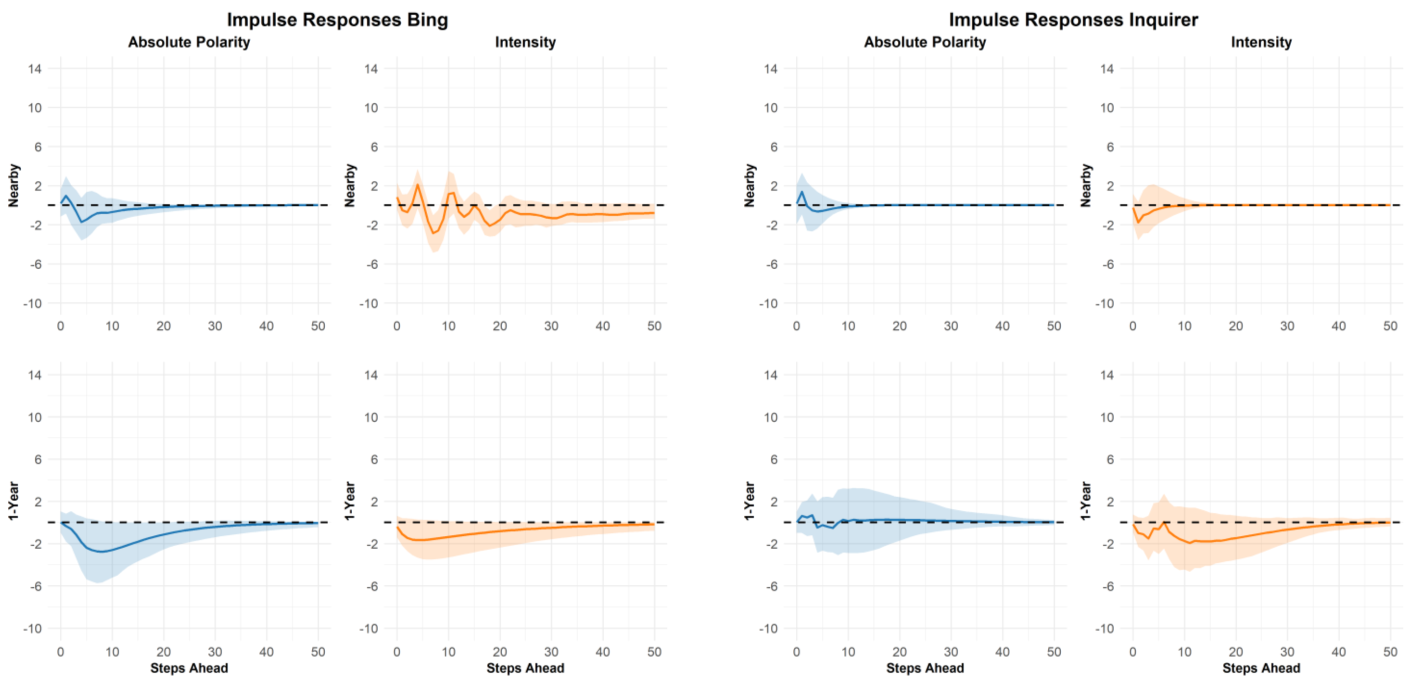

Figure 21: Corn convenience yield (in percentage points) responses to a corn-specific sentiment shock

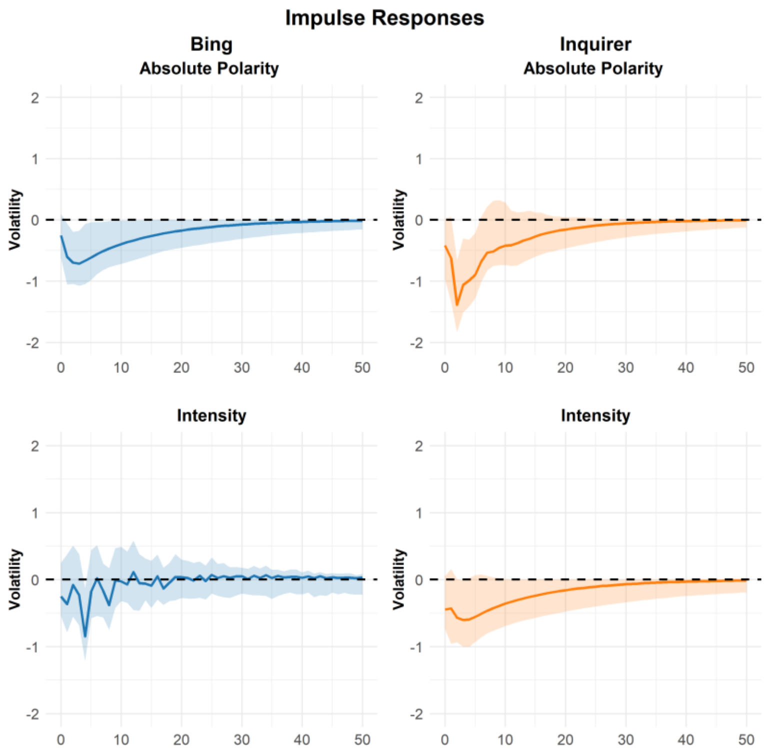

Figure 22: Soybean implied volatility responses to a financial sentiment shock

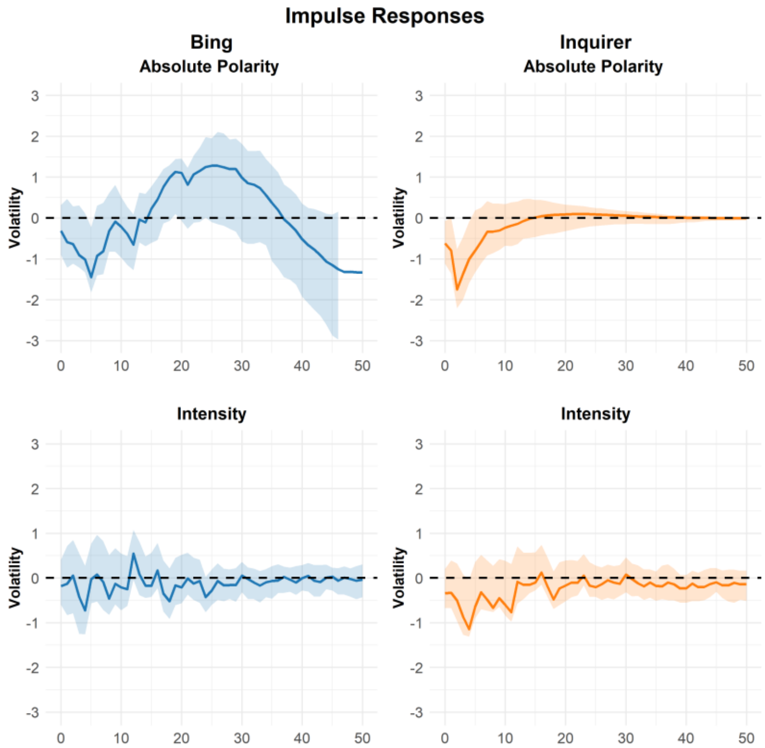

Figure 23: Corn implied volatility responses to a financial sentiment shock

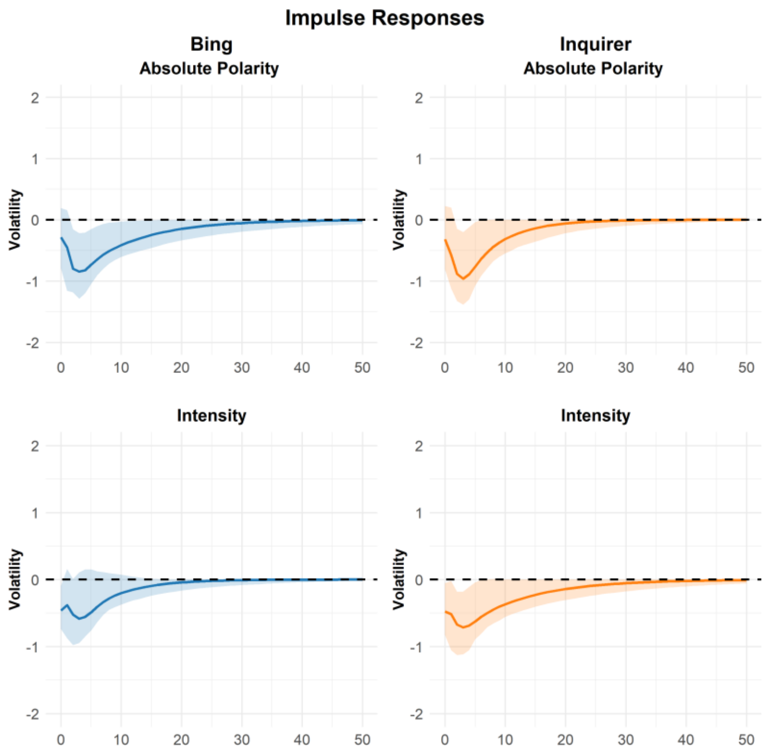

Figure 24: Soybean implied volatility responses to a commodity-specific financial sentiment shock

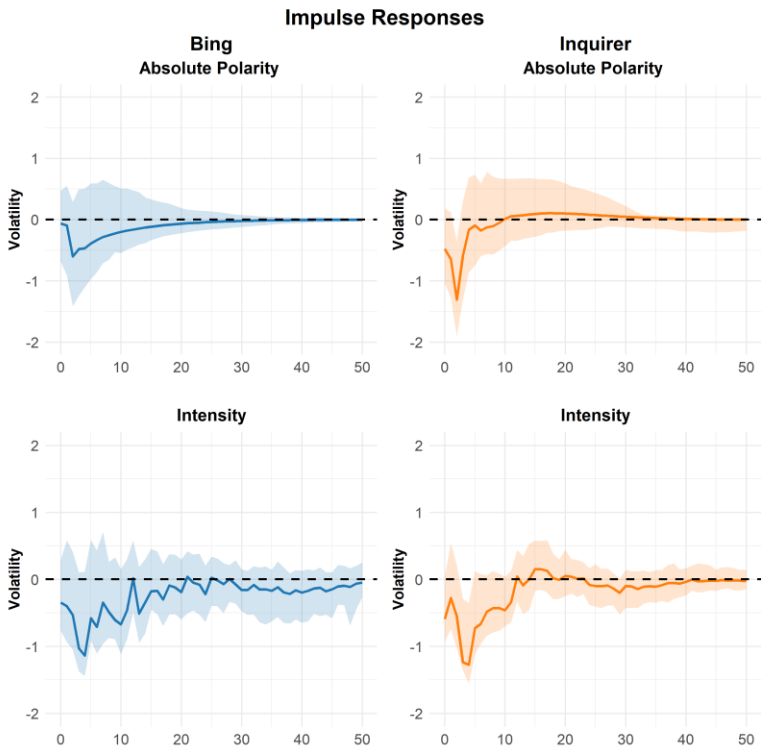

Figure 25: Corn implied volatility responses to a commodity-specific financial sentiment shock

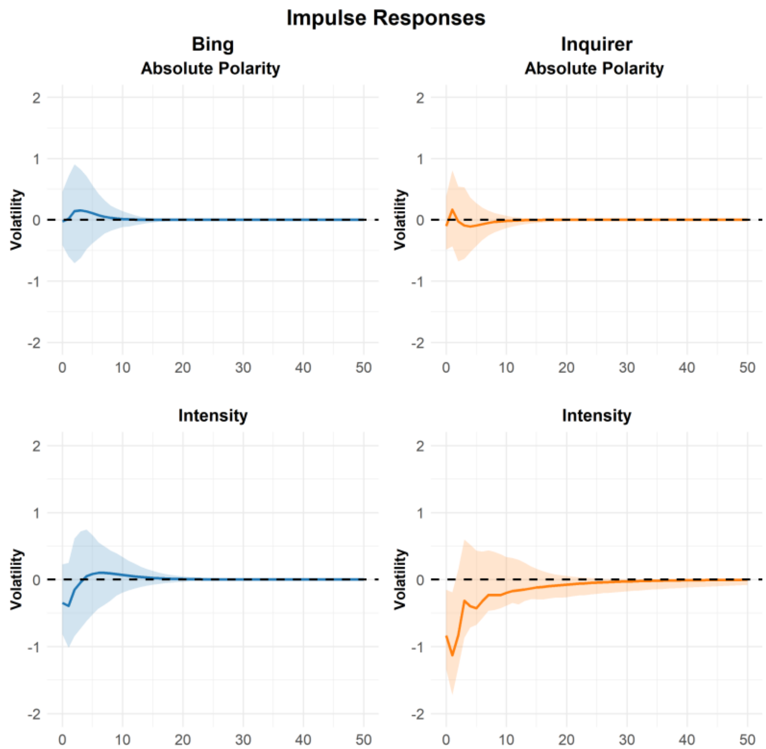

Figure 26: Soybean implied volatility responses to a soybean-specific sentiment shock

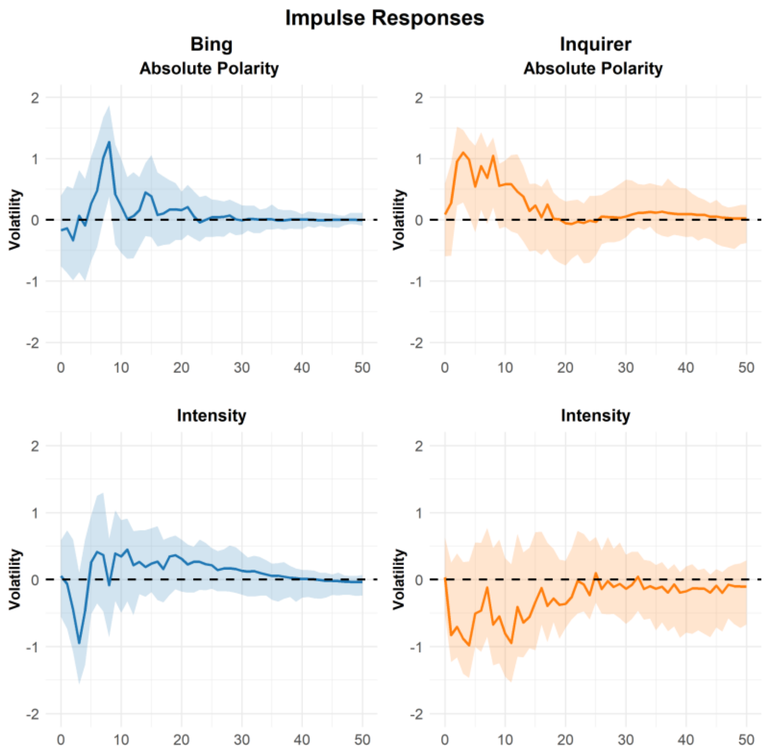

Figure 27: Corn implied volatility responses to a corn-specific sentiment shock