How Economic News Drives Implied Volatility in Agricultural Commodity Markets

Hongqiang Yan, Mark Manfredo, and Ashok Mishra

Suggested citation format:

Yan, H., Manfredo, M., and Mishra, A. 2026. “How Economic News Drives Implied Volatility in Agricultural Commodity Markets.” Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, Chicago, Illinois, April 20-21, 2026. [http://www.farmdoc.illinois.edu/nccc134/paperarchive]

How Economic News Drives Implied Volatility in Agricultural Commodity Markets

Hongqiang Yan[1] **Mark Manfredo**[2] **Ashok Mishra**[3]

Paper prepared for the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management, 2026.

[1]Postdoctoral Researcher, Morrison School of Agribusiness, Arizona State University. Email: hongqiang.yan@asu.edu.

[2]Professor, Morrison School of Agribusiness, Arizona State University.

[3]Professor and Kemper and Ethel Marley Chair in Food Management, Morrison School of Agribusiness, Arizona State University.

How Economic News Drives Implied Volatility in Agricultural Commodity Markets

Practitioner's Abstract

We examine whether broad economic news Granger-causes implied volatility derived from options on agricultural futures markets. Using approximately 1.5 million Wall Street Journal *articles (1984–2026), we construct 180 daily news topic attention series via Latent Dirichlet Allocation and apply a high-dimensional Granger causality framework that combines sparse-group LASSO variable selection with debiased inference and HAC standard errors. Using implied volatility measures provided by Bloomberg, we find that financial and macroeconomic news topics systematically Granger-cause implied volatility across corn, soybeans, soybean meal, and soybean oil markets and across Bloomberg constant-maturity horizons of 30 days, 60 days, and 3 months. Topics related to program trading and index arbitrage, investment banking, and the 2008 financial crisis are the most robust across commodities, consistent with the financialization of agricultural futures markets. The number of significant news topics declines from 23 at the 30-day maturity to 10 at the 3-month maturity, suggesting that financial news primarily influences short-term uncertainty. A crop-specific news topic is not selected for any commodity-maturity combination. This may be due to crop fundamentals being transmitted to implied volatility through scheduled USDA report releases rather than through ongoing national media coverage.*

Key Words: Topic Models, Textual Data, Agricultural Commodities, Implied Volatility, Granger Causality

1 Introduction

The use of textual data for economic research has gained considerable attention in recent years (Gentzkow et al., 2019; García, 2013; Bybee et al., 2024; Adämmer and Schüssler, 2020). This interest has extended to agricultural economists studying commodity markets (Qayum and Janzen, 2025; Li et al., 2024). At the same time, advances in machine learning and natural language processing have shifted news measurement toward approaches that analyze the content and narrative structure of news coverage (Hong et al., 2025; Barbaglia et al., 2023). Despite these developments, empirical research on how broad daily economic news affects agricultural commodity price behavior remains limited. This paper addresses that gap by examining the extent to which major newspaper-based economic news contributes to commodity price uncertainty, measured by the option-implied volatility of major agricultural futures markets.

Implied volatility is the expected volatility of a futures price over the life of an option, backed out from the option's market price. Because agricultural commodity options are written on the

corresponding futures contracts, implied volatility summarizes the market's forward-looking assessment of price risk for the underlying commodity and is widely used by market participants to manage that risk. Aka et al. (2025) find that implied volatility is the single most important predictor of future commodity option returns across a broad set of predictors. However, empirical research on the determinants of agricultural implied volatility has focused primarily on USDA information releases (Massa et al., 2024; He and Bian, 2025; Yang and McKenzie, 2025), seasonal production cycles (Simon, 2002), and energy market spillovers (Trujillo-Barrera et al., 2012). Broad daily economic news has received little attention as a driver of agricultural implied volatility.

Our analysis draws directly on two recent studies. First, Bybee et al. (2024) show that a topic model estimated from the full text of the *Wall Street Journal* (*WSJ*) generates 180 news attention series that track a wide range of macroeconomic conditions and explain approximately 25% of the variation in aggregate stock market fluctuations. Following their approach, we estimate a Latent Dirichlet Allocation (LDA) model with $K = 180$ topics (Blei et al., 2003) on *WSJ* articles as our measure of daily economic news attention. Second, Babii et al. (2024) develop a high-dimensional Granger causality framework that combines sparse-group least absolute shrinkage and selection operator (sg-LASSO) with debiased LASSO inference and HAC standard errors, and apply it to test whether these news topics Granger-cause the CBOE Volatility Index (VIX). We extend their framework to agricultural commodity markets, providing, to our knowledge, the first application of high-dimensional Granger causality tests to implied volatility in corn, soybeans, soybean oil, and soybean meal.

Our news data are drawn from ProQuest TDM Studio. We restrict the sample to items classified under the source type "Newspaper" and the document type "Article" in the *Wall Street Journal*, yielding approximately 1.5 million articles spanning January 1984 to March 2026. Following Bybee et al. (2024), we estimate an LDA model with $K = 180$ topics to obtain daily time series of news topic attention. For the implied volatility data, we use at-the-money (100% moneyness) Bloomberg implied volatility for corn, soybeans, soybean meal, and soybean oil at the 30-day, 60-day, and 3-month maturities, spanning October 2005 to March 2026 (5,023 daily observations per series).

The empirical framework models each commodity-maturity pair as an autoregressive distributed lag (ADL) regression in which all 180 LDA topic lags appear as potential predictors alongside lags of implied volatility. With a lag order of $L = 5$, this gives $180 \times 5 = 900$ topic-lag terms, a dimension that classical ordinary least squares (OLS) Granger tests cannot accommodate. Following Babii et al. (2024), we apply sparse-group LASSO variable selection, keeping the AR(1) component unpenalized, and conduct post-selection inference via the debiased LASSO estimator with HAC standard errors.

Our main findings are as follows. At the 30-day maturity, we identify seven significant topics for corn and soybean oil, four for soybeans, and five for soybean meal. The most consistent topics that Granger cause implied volatility across commodities for the 30-day maturity are three topics related to financial markets: "Program Trading and Index Arbitrage," "Investment Banking and Global Capital Markets," and "The Financial Crisis and Collapse of Investment Banks." The number of significant topic-commodity pairs declines from 23 at the 30-day maturity to 15 at the

60-day maturity and 10 at the 3-month maturity, indicating that financial news has a stronger effect on short-term rather than long-term implied volatility.

We also find that the crop-specific news topic (with top terms including "price, corn, soybean, grain, farmer, commodity") is not selected for any commodity-maturity combination. A possible explanation is that crop-related uncertainty is driven primarily by USDA information events rather than by ongoing national media media coverage. The established literature shows that USDA reports resolve crop-specific uncertainty at scheduled, predictable intervals (Massa et al., 2024; He and Bian, 2025; Yang and McKenzie, 2025; Lee and Park, 2024), delivering officially compiled supply and demand estimates that markets cannot anticipate from publicly available news (Goyal and Adjemian, 2023). *WSJ* crop coverage, by contrast, reflects ongoing media commentary that is freely accessible and likely already reflected in prices. Once this information channel is captured by the USDA announcement mechanism, ongoing newspaper coverage of crop conditions carries no additional predictive power for implied volatility.

The remainder of the paper proceeds as follows. Section 2 reviews the related literature. Section 3 describes the data. Section 4 presents the empirical framework. Section 5 discusses the results. Section 6 concludes.

2 Related Literature

2.1 Text as Data in Finance and Economics

The use of unstructured text for quantitative economic research has grown rapidly across three broad methodological approaches (Gentzkow et al., 2019). The first is dictionary-based sentiment analysis, which counts positive and negative words from predefined lexicons and aggregates them into sentiment scores. García (2013) establishes a foundational result with respect to this approach in that sentiment derived from *New York Times* financial columns predicts Dow Jones Industrial Average returns, with effects three to four times larger during recessions, demonstrating that media tone contains information about financial market outcomes beyond standard economic variables. Barbaglia et al. (2023) extend the sentiment approach with a fine-grained, aspect-specific variant (FiGAS), extracting sentiment along six economic dimensions from 6.6 million articles across six U.S. newspapers. They find that financial sector sentiment helps forecast GDP, industrial production, payrolls, and CPI using a post-LASSO selection framework. The FiGAS methodology is detailed in Consoli et al. (2022), who propose a domain-specific economic and financial sentiment lexicon and show it outperforms general-purpose dictionaries for economic text. Qayum and Janzen (2025) apply sentiment analysis to agricultural commodity markets specifically, finding that positive sentiment shocks reduce corn and soybean implied volatility by 0.5 to 1.5 percentage points in a structural VAR, though they find no significant sentiment effect on convenience yields.

The second approach constructs keyword-based indices by counting the frequency of predefined

term sets relevant to specific economic or policy themes. Baker et al. (2026) develop Equity Market Volatility (EMV) trackers from article counts across eleven U.S. newspapers using term sets covering forty economic categories, finding that their headline tracker correlates at $R^2 \approx 0.60$ with the VIX in-sample and maintains comparable predictive power out-of-sample through 2023. News about commodity markets features in 44% of all EMV articles—the most frequently mentioned category—indicating that commodity market developments are closely connected to the financial news narratives that drive equity market volatility. Kelly et al. (2021) propose Hurdle Distributed Multinomial Regression (HDMR), a high-dimensional method that separately models whether a word appears and how often it recurs, applied to *WSJ* front-page text. They find that the inclusion margin carries economic information beyond word counts and improves out-of-sample fit over earlier count-based text models.

The third approach uses unsupervised topic models, which infer a corpus's thematic structure from word co-occurrence patterns without predefined term lists or labeled data. Thorsrud (2020) extracts LDA topics from a Norwegian business newspaper and feeds them into a time-varying dynamic factor model, producing a daily index that tracks the business cycle and dates recessions accurately. Bybee et al. (2024) apply LDA at scale to the full *WSJ* corpus (800,000 articles), validating a 180-topic specification via Bayes factor and demonstrating that topic attention series track NBER recessions, GDP growth, and unemployment. Most directly relevant to our work, Hong et al. (2025) use the same 180-topic *WSJ* decomposition to forecast CPI inflation, showing that narrative attention improves out-of-sample root mean squared error (RMSE) by 3–10% at the 3-month horizon and 10–25% at the one-year horizon, with outperformance concentrated during recessions. Our paper extends this strand of the LDA literature to agricultural commodity implied volatility, asking whether the same news narrative structure that predicts equity market outcomes and inflation also carries predictive content for commodity price risk expectations.

Adämmer et al. (2025) further demonstrate the breadth of applications for topic-based text predictors. They apply the Correlated Topic Model (CTM; Blei and Lafferty, 2007), which is a variant of LDA that allows topic proportions to co-occur, to approximately 800,000 articles from the *New York Times* and *Washington Post*, and show that textual predictors substantially improve quantile (tail risk) forecasts of employment, inflation, industrial production, and consumer sentiment relative to a rich set of FRED-MD economic indicators. Ardia and Bluteau (2026) take a different approach, using a genetic algorithm to select which words in *WSJ* text are most predictive of a target variable such as inflation. Applied to the *WSJ* corpus, their methodology both recovers the keyword structure underlying the Economic Policy Uncertainty (EPU) index and improves out-of-sample inflation forecasting accuracy, illustrating the gains from data-driven token selection relative to predetermined term lists. Our paper contributes to this growing literature by extending text-based prediction to agricultural commodity implied volatility—a financial market variable not previously studied in this context.

2.2 Agricultural Commodity Market Volatility and Information

Implied volatility is an economically important variable in agricultural markets, where it is often used as a key risk measure. For example, the USDA Risk Management Agency derives the volatility factor underlying revenue-protection crop insurance directly from the implied volatility of close-to-the-money options over the final five trading days of the price-discovery period (Goodwin et al., 2014). As well, implied volatility in agricultural options markets has been shown to forecast realized price volatility and to anticipate its changing patterns (Simon, 2002; Egelkraut et al., 2007). Beyond its role in measuring and forecasting price risk, implied volatility is important in predicting options returns. For example, across a broad set of commodity options, including agricultural commodities, Aka et al. (2025) find it to be the single most important return predictor.

The established literature on what drives implied volatility in agricultural markets is centered on the USDA announcement mechanism. Massa et al. (2024) review more than five decades of research on USDA report market impacts. The central finding is that implied volatility rises in anticipation of scheduled World Agricultural Supply and Demand Estimates (WASDE) releases and declines sharply upon publication as uncertainty resolves. He and Bian (2025) show that the implied volatility embedded in corn, soybeans, and wheat options on WASDE release days exceeds the realized volatility that follows, which they interpret as an announcement volatility risk premium (e.g., the compensation market participants pay to hedge announcement-induced price risk). Yang and McKenzie (2025) find that implied volatility falls by approximately one percentage point around August WASDE releases but incorporates crop news only sluggishly, suggesting that non-crop factors influence implied volatility beyond announcement windows.

Two features of the USDA announcement mechanism are relevant to our paper. First, the information it delivers is genuinely new. For example, Goyal and Adjemian (2023) show that USDA crop forecasts incorporate field survey data that is not otherwise publicly available, arriving at discrete intervals rather than continuously through media coverage. Second, the uncertainty it resolves dissipates quickly. Lee and Park (2024) find that realized volatility spikes at the 11:00 AM CST release time and dissipates in under sixty minutes, while pre- and post-release days are statistically indistinguishable. The USDA channel therefore explains discrete, high-frequency movements in implied volatility, leaving room for lower-frequency news channels to explain persistent variation.

The USDA announcement channel also reaches beyond commodity futures. Cao et al. (2024) show that food-sector equity prices react significantly to WASDE, Grain Stocks, and Acreage reports, with input- and output-side firms responding in opposite directions to the same USDA surprise. Cao and Robe (2022) show that commodity implied volatility declines significantly in the days following USDA releases. Agricultural information events thus shape uncertainty expectations across the financial system.

Beyond scheduled USDA announcements, news sentiment provides a lower-frequency information channel. Chi et al. (2024) find that both novel and recycled news sentiment predict commodity futures returns, with the effect strongest during periods of high financial stress. Rogmann et al. (2024) find analogous results for energy commodities, where newspaper sentiment significantly

affects consumer energy prices during periods of high media coverage.

Adjemian and Jo (2025) show using a proxy SVAR that an adverse agricultural supply news shock raises the VIX and reduces real GDP and industrial production—confirming that shocks originating in agricultural markets propagate more broadly into the economy. Our paper asks the reverse question: does news about the broader financial system predict agricultural commodity implied volatility?

2.3 Commodity Market Financialization

Commodity futures markets changed markedly over the first decade of the 2000s. Index investment in commodity-related instruments rose from about $15 billion in 2003 to more than $200 billion by mid-2008 (Tang and Xiong, 2012). Furthermore, as institutional investors, pension funds, and exchange-traded funds entered these markets, they increasingly treated them as another financial asset class rather than tools for hedging price risks (Irwin and Sanders, 2012). Grain futures volumes more than tripled, and trading shifted from telephone and open-outcry to electronic order matching, which lowered costs and broadened access (Irwin and Sanders, 2012).

The key consequence for our purposes is that index investors allocate to commodities as an asset class and adjust positions across all indexed commodities simultaneously in response to financial market conditions, so any shift in their broader financial allocation flows into agricultural futures regardless of crop conditions.

Tang and Xiong (2012) document that after 2004 the futures prices of non-energy commodities such as corn and soybeans comoved far more strongly with oil prices, with the rise in comovement concentrated among index-member commodities, and this increase was significantly larger for commodities included in major commodity indices than for off-index commodities. Financial market shocks therefore propagate into agricultural futures even when crop fundamentals are unchanged. This integration motivates our focus on broad economic news, alongside crop-specific news, as a predictor of agricultural implied volatility.

3 Data

The implied volatility measure we use is the constant-maturity, at-the-money implied volatility estimated at the daily market close. These implied volatility estimates are calculated and retrieved from Bloomberg (Cao and Robe, 2022). Because agricultural futures options are listed only for selected contract months rather than every calendar month, we take constant-maturity implied volatilities from Bloomberg's OVDV volatility surface, which is fitted from out-of-the-money options and interpolated across maturities using Hermite interpolation of total variance to yield series at fixed 30-day, 60-day, and 3-month horizons (Cui, 2011). We obtain implied volatility at these

three maturities for four commodities: corn, soybeans, soybean meal, and soybean oil.

Corn and the soybean complex play a central role in both U.S. and global agricultural markets, and have been subject to heightened uncertainty in recent years owing to, among other factors, trade policy developments affecting U.S. soybean export demand. The sample runs from October 28, 2005, through March 2026, providing 5,023 daily observations for each commodity–maturity pair and 12 dependent-variable series in all.

Agricultural implied volatility series exhibit several features relevant to our empirical framework. First, all series are highly persistent, with first-order autocorrelations typically exceeding 0.97, motivating the autoregressive component of our ADL specification. Second, implied volatility is higher on average and more variable at the 30-day maturity than at the 3-month maturity, reflecting the greater sensitivity of short-dated options to near-term uncertainty. Third, the series exhibit pronounced peaks during the 2008–2009 financial crisis, the 2012 U.S. drought, and the 2020 COVID-19 pandemic—episodes associated with both commodity-specific shocks and broader financial market dislocations consistent with the financialization hypothesis (Tang and Xiong, 2012; Irwin and Sanders, 2012). Table 1 reports summary statistics for each commodity-maturity pair.

Our news dataset is extracted from ProQuest TDM Studio. We collect full-text articles published in the *Wall Street Journal*, restricting the sample to items classified under the source type "Newspaper" and the document type "Article." The resulting corpus comprises approximately 1.5 million articles and approximately two billion words spanning January 1984 to March 2026. Using the *Wall Street Journal* as the news source follows Bybee et al. (2024) directly and ensures that our LDA topic structure is fully comparable to the 180-topic model they estimate and validate.

To organize and decompose the unstructured text into a lower-dimensional set of common "news topics," we apply LDA (Blei et al., 2003). LDA is an unsupervised generative model in which each topic is represented as a probability distribution over the vocabulary, and each document is represented as a mixture of topics. Topics can be thought of as clusters of words that tend to co-occur across articles. The formation of topics is optimized so that a relatively small number captures as much of the variation in word usage across documents as possible (Bybee et al., 2024). Importantly, this approach captures what the news is about—the narrative content—rather than merely its emotional tone, which distinguishes it from dictionary-based sentiment measures (García, 2013; Barbaglia et al., 2023) and from keyword-frequency approaches (Baker et al., 2026). The unsupervised nature of LDA means that the topic structure is determined entirely by word co-occurrence patterns in the corpus, without requiring the researcher to specify a list of relevant terms in advance.

Following Bybee et al. (2024), we set the number of topics to $K = 180$, which they find to be a statistically optimal specification according to a Bayes factor criterion. For each day t, we compute the daily topic attention as

$$z_{j,t} = \frac{1}{N_t} \sum_{i=1}^{N_t} \hat{\theta}_{ij}, \tag{1}$$

where $\hat{\theta}_{ij}$ is the estimated share of article i devoted to topic j on day t, and N_t is the number of articles published on day t. This yields a $T \times 180$ matrix of daily topic attention series that serve

as the independent variables in our Granger causality analysis.

Table 3 illustrates selected topics from the estimated model, along with their top keywords. The 180 topics broadly reflect the major themes covered in the *Wall Street Journal* throughout the sample period, including financial markets (e.g., "Program Trading and Index Arbitrage"; "The Financial Crisis and Collapse of Investment Banks"), macroeconomic policy ("Interest Rates and Monetary Policy"; "Tax Policy and Revenue Legislation"), geopolitical events ("Russia–Ukraine Conflict and European Security"), and agricultural commodities ("Crop", with top terms including "price, corn, soybean, grain, farmer, commodity"). The topic model thus spans a broad spectrum of economic narratives, enabling us to test which, if any, exhibit incremental predictive power for agricultural implied volatility.

4 Empirical Framework

4.1 Autoregressive Distributed Lag Model and Granger Causality

Following Babii et al. (2024), we model each commodity-maturity pair using an autoregressive distributed lag (ADL) regression:

$$y_t = \mu + \sum_{l=1}^{L} \phi_l \, y_{t-l} + \sum_{j=1}^{180} \sum_{l=1}^{L} \beta_{jl} \, z_{j,t-l} + \varepsilon_t, \tag{2}$$

where y_t denotes the implied volatility series for a given commodity-maturity pair, $z_{j,t}$ denotes the daily attention to news topic j as defined in Equation (1), L is the lag order set to 5 (one trading week), capturing the short-run dynamics over which news is likely to be incorporated into option prices, and ε_t is an error term. Both the implied volatility series and the topic attention series are used in levels. We verify the stationarity of the implied volatility series prior to estimation via Phillips–Perron unit root tests (Table 2). All series reject the unit root null at the 5% level or better. Topic attention shares are bounded in $[0,1]$ by construction and therefore stationary.

Under this specification, news topic j does not Granger-cause y_t if and only if its distributed lag coefficients are jointly zero:

$$H_0^{(j)}: \quad \beta_{j1} = \beta_{j2} = \cdots = \beta_{jL} = 0. \tag{3}$$

Rejecting $H_0^{(j)}$ implies that lagged attention to topic j contains incremental predictive information for implied volatility, beyond the history of y_t itself (Granger, 1969).

4.2 Estimation and Inference

With $L = 5$ lags, the design matrix in Equation (2) contains $180 \times 5 = 900$ topic-lag regressors in addition to the autoregressive terms, far exceeding what classical OLS Granger tests can accommodate. We therefore apply the two-step high-dimensional inference procedure of Babii et al. (2024). In the first step, the ADL model is estimated via sg-LASSO regularization, which simultaneously selects individual topic lags and entire topic groups, with the AR(1) term always retained unpenalized. In the second step, a debiased estimator corrects the shrinkage bias of the sg-LASSO and yields asymptotically normal coefficient estimates with HAC standard errors, enabling valid post-selection inference. We refer the reader to Babii et al. (2024) for full technical details, including the debiased central limit theorem and the HAC long-run variance estimation.

We test the Granger non-causality null $H_0^{(j)}$ via a Wald χ^2 statistic constructed from the debiased z-statistics for all L lags of topic j. Under $H_0^{(j)}$, the statistic is asymptotically $\chi^2(L)$. We reject at the 5% significance level.

As a complementary test, we also estimate the reverse direction, asking whether lagged agricultural implied volatility Granger-causes news topic attention:

$$z_{j,t} = \alpha_j + \sum_{l=1}^{L} \gamma_{jl} z_{j,t-l} + \sum_{l=1}^{L} \delta_{jl} y_{t-l} + u_{jt}. \tag{4}$$

Evidence of reverse Granger causality would indicate that financial media coverage responds to agricultural market volatility, informing our interpretation of the main results.

5 Results

5.1 Granger Causality at the 30-Day Maturity

Table 4 reports the number of news topics that Granger-cause implied volatility at the 5% significance level. At the 30-day maturity, we find 7 significant topics for corn and soybean oil, 5 for soybean meal, and 4 for soybeans, for 23 significant commodity-topic pairs in total.

Three topics related to financial markets account for the most consistent results across commodities (Table 5). "Program Trading and Index Arbitrage" is the only topic significant for all four commodities. Its top terms (*williams, program, trading, stewart, stock, securities, arbitrage*) reflect equity market microstructure and index arbitrage activity, and its significance across all four markets is consistent with the commodity financialization hypothesis (Tang and Xiong, 2012; Irwin and Sanders, 2012): as institutional investors increasingly allocate to agricultural futures as a financial asset class, equity-market trading activity becomes a common driver of agricultural

price risk. "Investment Banking and Global Capital Markets" (key terms: *capital, bank, boston, credit, suisse, deutsche, markets*) is significant for corn, soybeans, and soybean meal, suggesting that news about investment bank health and capital market conditions propagates into grain market uncertainty. "The Financial Crisis and Collapse of Investment Banks," with top terms (*smith, lehman, brothers, salomon, bond, barney, bear stearns*) capturing coverage of the 2008 financial collapse, is likewise significant for corn, soybeans, and soybean meal. Babii et al. (2024) find an analogous *Financial crisis* topic to be the dominant Granger cause of the VIX, significant at the 1% level regardless of sample period and model specification; our results extend this finding to agricultural implied volatility, consistent with commodity–equity co-movement during periods of systemic financial stress.

Beyond these three, each commodity has additional significant topics that do not generalize across markets. For corn, "Presidential Administration and White House Policy" (key terms: *trump, president, biden, white, house, administration, policy*) is associated with a negative coefficient: higher attention to White House policy is associated with lower corn implied volatility, consistent with policy announcements resolving rather than amplifying uncertainty. "Tax Policy and Revenue Legislation," "Equity Markets and Exchange Trading," and "Mutual and Hedge Funds" are also significant for corn. For soybean meal, "Interest Rates and Monetary Policy" (key terms: *rate, interest, term, treasury, federal reserve, yield, bond*) and an M&A topic (key terms: *deal, bid, billion, stake, buy, acquire, takeover*) are significant. For soybeans, a corporate earnings topic (key terms: *cent, stock, share, company, earning, quarter, revenue*) is statistically significant.

Soybean oil exhibits a different pattern from the grain markets. Beyond "Program Trading and Index Arbitrage," significant topics include a real estate topic (key terms: *hotel, million, property, house, room, resort, real estate*), "Russia–Ukraine Conflict and European Security," "Public Health Crisis" (key terms: *test, vaccine, virus, health, disease, pandemic, coronavirus*), and "Interest Rates and Monetary Policy." Two topics enter with negative coefficients for soybean oil: "New York City Urban Development and Policy" (key terms: *mayor, building, street, manhattan, authority, city, new york*) and an entertainment topic (key terms: *movie, film, million, studio, cola, entertainment, hollywood*).

5.2 Granger Causality at the 60-Day Maturity

At the 60-day maturity, 15 topic-commodity pairs are significant, down from 23 at 30 days. The financial stress topics that dominated the 30-day results largely recede: "Investment Banking and Global Capital Markets" and "The Financial Crisis and Collapse of Investment Banks" are no longer significant for any commodity. "Program Trading and Index Arbitrage" is the exception, remaining significant for corn and soybean meal.

The most notable new entry is "Japan's Economy and Corporate Earnings" (key terms: *yen, billion, billion yen, profit, export, japan, corporate*). While it was not significant at the 30-day maturity, it is for soybeans and soybean meal at the 60-day maturity. Japan is the largest single-country importer of U.S. corn and a major buyer of U.S. soybeans. The delayed appearance suggests that

macroeconomic conditions in Asian export markets shape longer-run demand uncertainty rather than short-term price risk. "Executive Compensation" (key terms: *option, pay, compensation, bonus, executive, stock option*) is also significant for corn. For soybean meal, the M&A topic and "Mutual and Hedge Funds" persist from 30 days.

For soybean oil, three topics that were significant at 30 days also remain significant at this maturity: "Russia–Ukraine Conflict and European Security," "New York City Urban Development and Policy," and the real-estate topic (key terms: *hotel, million, property, house, room, resort, real estate*).

5.3 Granger Causality at the 3-Month Maturity

At the 3-month maturity, 10 topic-commodity pairs are significant. "Japan's Economy and Corporate Earnings" broadens from the soybean complex to now include corn, becoming significant for corn, soybeans, and soybean meal. This is the most broadly significant topic at this maturity. "Executive Compensation" and "Program Trading and Index Arbitrage" persist for corn. "Mutual and Hedge Funds" persists for soybean meal.

For soybean oil, "Russia–Ukraine Conflict and European Security" no longer Granger-causes implied volatility, whereas "New York City Urban Development and Policy" continues to Granger-cause implied volatility with a negative coefficient. Two technology topics are newly significant at this maturity: a computing topic (key terms: *system, software, apple, computer, device, hardware, chip*) for soybean oil, and an internet and technology topic (key terms: *internet, technology, company, online, tech, web, digital*) for corn.

5.4 Absence of the Crop-Specific News Topic

The "Crop" topic (key terms: *price, market, corn, soybean, grain, farmer, crop, commodity*) is not selected for any commodity-maturity combination. This is not a result of low variation in the topic series: the topic is active throughout the sample. The result holds at all three maturities and all four commodities.

The most direct explanation is the structure of how crop information reaches options markets. He and Bian (2025) report straddle returns of -7.12% for corn and -5.04% for soybeans on WASDE announcement days—discrete, predictable episodes in which uncertainty is resolved at a single point in time. Lee and Park (2024) show that the realized volatility spike at the 11:00 AM WASDE release dissipates within sixty minutes. Crop-specific uncertainty therefore arrives through scheduled USDA reports that carry genuinely new field-survey data (Goyal and Adjemian, 2023), not through the ongoing *WSJ* coverage that our topic series measures. Once the autoregressive dynamics of implied volatility absorb that information, daily crop media attention has no remaining predictive power.

5.5 Reverse Granger Causality

Our results in previous sections identify three financial-market themes as the topics most strongly associated with 30-day implied volatility: "Program Trading and Index Arbitrage," "Investment Banking and Global Capital Markets," and "The Financial Crisis and Collapse of Investment Banks." This raises a natural concern about the direction of causality. That is, rather than anticipating volatility, financial media may simply report on market turbulence after it has occurred, so that coverage lags the market rather than leading it. If that were the case, lagged implied volatility would predict subsequent attention to these topics. We therefore test whether 30-day implied volatility Granger-causes attention to the three topics, estimating the reverse regression with the debiased LASSO and controlling for the lags of all other topics among the 180. Finding that implied volatility Granger-causes attention would indicate that coverage is at least partly reactive, whereas the absence of such an effect supports our forward-looking interpretation, in which news attention precedes rather than follows movements in implied volatility.

"Investment Banking and Global Capital Markets" and "The Financial Crisis and Collapse of Investment Banks" are Granger-caused by implied volatility for all four commodities. Higher agricultural implied volatility predicts greater subsequent media attention to financial stress narratives, consistent with commodity market turbulence feeding back into financial news coverage. For "Program Trading and Index Arbitrage," the pattern is more varied: implied volatility Granger-causes this topic for soybeans and soybean meal, but not for corn, and the effect for soybean oil is negative, indicating that elevated soybean oil implied volatility is followed by reduced attention to program trading.

The bidirectionality for "Investment Banking" and "Financial Crisis" does not invalidate the forward results. The two directions measure different objects: the forward Granger tests ask whether news predicts implied volatility after conditioning on implied volatility's own history, while the reverse tests ask whether implied volatility predicts news after conditioning on each topic's own history. Both findings can hold simultaneously and together indicate that financial narratives and agricultural market uncertainty are jointly determined.

6 Conclusion

We examine the Granger-causal relationship between broad daily economic news and implied volatility in agricultural commodity futures markets. Drawing on approximately 1.5 million *Wall Street Journal* articles spanning 1984 to 2026, we extract 180 news topic attention series via Latent Dirichlet Allocation following Bybee et al. (2024), and apply the high-dimensional Granger causality framework of Babii et al. (2024) to test whether these topics predict implied volatility for corn, soybeans, soybean meal, and soybean oil at the 30-day, 60-day, and 3-month maturities.

Three findings emerge from our analysis. First, economic news Granger-causes agricultural im-

plied volatility across all four markets and all three implied volatility maturities. The news channel for agricultural price risk is thus empirically substantive, complementing the established roles of USDA information releases (Massa et al., 2024; He and Bian, 2025; Yang and McKenzie, 2025), seasonal production cycles (Simon, 2002), and energy market spillovers (Trujillo-Barrera et al., 2012). Second, financial channels dominate the news-to-implied volatility transmission. Topics related to program trading and index arbitrage, investment banking, and the 2008 financial crisis are the most robust Granger causes across commodities, consistent with the financialization hypothesis. That is, as agricultural futures markets became increasingly integrated with financial markets through commodity index investment (Tang and Xiong, 2012), systemic financial shocks, rather than crop-specific news, became the dominant driver of agricultural price risk expectations. Third, the results exhibit a pronounced term structure. The combined count of significant topic-commodity pairs falls monotonically from 23 at the 30-day maturity, to 15 at the 60-day maturity, and 10 at the 3-month maturity. Most financial topics are short-lived, with their significance concentrated at the 30-day maturity and often absent by the 3-month maturity. A smaller set of topics exhibits either delayed effects, such as Japan's economy for the soybean complex, or persistent effects, such as mutual funds and hedge funds for soybean meal.

The absence of the crop-specific news topic from the selected set reinforces both the view that agricultural futures have become integrated with financial markets (Tang and Xiong, 2012; Irwin and Sanders, 2012) and the efficiency of the USDA announcement mechanism. Crop fundamentals appear to be transmitted to agricultural implied volatility through scheduled USDA report releases (He and Bian, 2025; Yang and McKenzie, 2025; Lee and Park, 2024), which carry genuinely new supply and demand estimates compiled under lockup conditions (Goyal and Adjemian, 2023). Daily media commentary on crop conditions, by contrast, carries no marginal predictive power once the autoregressive history of implied volatility is controlled. This finding is consistent with markets efficiently incorporating publicly available crop narratives and other information regarding the expected supply and demand conditions for these commodities.

These findings carry several practical implications. For agricultural market participants, including farmers, elevator operators, and commodity traders, monitoring broad financial market narratives rather than crop-specific news alone may provide early signals of changes in commodity price risk. The dominance of financial topics at the 30-day maturity has direct relevance for short-dated option pricing and hedging strategies. Furthermore, since Aka et al. (2025) establish that implied volatility is the single most important predictor of future commodity option returns, understanding news-driven implied volatility dynamics has direct implications for commodity option trading strategies. For policymakers, the commodity market integration with general financial markets identified here suggests that agricultural market stress can be partly driven by financial market disruptions unrelated to crop fundamentals, implying that financial stability policy has spillover effects on commodity market risk.

Table 1: Descriptive Statistics: Bloomberg Implied Volatility Series

Commodity	Maturity	Obs	Mean	Std. Dev.	Min	Max
Corn	30-day	5,098	26.82	9.20	9.62	62.06
	60-day	5,098	26.98	8.63	10.06	60.06
	3-month	5,096	27.07	8.11	11.44	51.81
Soybeans	30-day	5,100	21.92	7.51	9.56	61.69
	60-day	5,100	22.01	7.33	10.44	66.25
	3-month	5,100	22.09	7.09	10.88	57.81
Soybean Meal	30-day	5,100	24.86	6.81	10.99	57.99
	60-day	5,100	24.38	6.31	11.39	50.04
	3-month	5,100	23.99	5.92	11.93	47.96
Soybean Oil	30-day	5,096	24.55	6.85	12.02	64.37
	60-day	5,096	24.25	6.57	12.58	50.27
	3-month	5,096	24.11	6.47	13.33	48.54

Notes: Sample: October 28, 2005, to March 2026. Values reported in percentage points (annualized). Obs counts differ slightly across maturities due to sporadic missing values in Bloomberg data.

Table 2: Phillips–Perron Unit Root Tests

Commodity	**Maturity**	**PP Statistic (Z_τ)**
Corn	30-day	-8.200^{**}
	60-day	-6.198^{**}
	3-month	-5.297^{**}
Soybeans	30-day	-7.168^{**}
	60-day	-6.016^{**}
	3-month	-4.545^{**}
Soybean Meal	30-day	-8.069^{**}
	60-day	-5.886^{**}
	3-month	-4.942^{**}
Soybean Oil	30-day	-7.045^{**}
	60-day	-4.232^{**}
	3-month	-3.573^{*}

Notes: Phillips–Perron Z_τ test with constant and trend; long-run variance estimated by Newey–West. Critical values: -3.97 (1%), -3.41 (5%), -3.13 (10%). H_0: unit root. $^{*}p < 0.05$; $^{**}p < 0.01$.

Table 3: Selected LDA Topics and Top Keywords

T#	Label	Top Keywords (LDA output)
Financial markets		
13	Program Trading & Index Arbitrage	williams, program, trading, stewart, stock, securities, arbitrage
21	Investment Banking & Global Capital Markets	capital, bank, boston, credit, suisse, deutsche, markets
77	Financial Crisis & Collapse of Investment Banks	smith, lehman, brothers, salomon, bond, barney, bear stearns
63	Equity Markets & Exchange Trading	stock, trading, market, exchange, trade, nasdaq, investor
175	Mutual & Hedge Funds	fund, mutual, asset, money, mutual fund, hedge, fidelity
25	Mergers & Acquisitions	deal, bid, billion, stake, buy, acquire, takeover
Macroeconomic and trade		
107	Presidential Administration & White House Policy	trump, president, biden, white, house, administration, policy
125	Tax Policy & Revenue Legislation	tax, irs, income, bill, deduction, revenue, legislation
28	Interest Rates & Monetary Policy	rate, interest, term, treasury, federal reserve, yield, bond
5	Japan's Economy & Corporate Earnings	yen, billion, billion yen, profit, export, japan, corporate
4	Russia–Ukraine Conflict & European Security	russia, russian, ukraine, european, moscow, nato, security
6	Executive Compensation & Corporate Governance	option, pay, compensation, bonus, executive, stock option
Other		
47	Real Estate & Hotels	hotel, million, property, house, room, resort, real estate
155	Public Health Crisis	test, vaccine, virus, health, disease, pandemic, coronavirus
31	New York City Urban Development & Policy	mayor, building, street, manhattan, authority, city, new york
45	Corporate Losses & Restructuring	million, loss, company, bankruptcy, restructure, charge
22	Internet & Technology Companies	internet, technology, company, online, tech, web, digital
89	Computing & Technology (Apple)	system, software, apple, computer, device, hardware, chip
86	Tobacco Industry	tobacco, cigarette, morris, chemical, philip, smoking, rj
68	Entertainment & Media	movie, film, million, studio, cola, entertainment, hollywood
145	Corporate Earnings Reports	cent, stock, share, company, earning, quarter, revenue
174	Quarterly Financial Results	million, share, cent, quarter, profit, results, revenue
34	*Crop (not selected by sg-LASSO in any specification)*	price, market, corn, soybean, grain, farmer, crop, commodity

Notes: Topics are grouped to match Table 5. Top keywords are the highest-probability terms from the LDA model of Bybee et al. (2024). Topic 34 (Crop) appears in none of the significant Granger causality results across all 12 commodity-maturity pairs.

Table 4: Number of Significant News Topics Granger-Causing Implied Volatility ($p < 0.05$)

Implied Volatility Maturity	Corn	Soybeans	Soybean Meal	Soybean Oil	Total
30-day	7	4	5	7	23
60-day	5	1	4	5	15
3-month	5	1	2	2	10

Notes: Each cell reports the number of LDA topics (out of 180) for which the Granger non-causality null is rejected at the 5% significance level based on the Wald χ^2 statistic from the debiased LASSO of Babii et al. (2024). Lag order $L = 5$. Sample: October 2005–March 2026 ($T = 5{,}023$).

Table 5: Granger Causality by Topic, Commodity, and Implied Volatility Maturity

		Corn			Soybeans			Soybean Meal			Soybean Oil		
T#	**Topic**	**30D**	**60D**	**3M**	**30D**	**60D**	**3M**	**30D**	**60D**	**3M**	**30D**	**60D**	**3M**
Financial markets													
13	Program Trading	***	***	*	*			***	***		***		
21	Investment Banking	*			***			**					
77	Financial Crisis	*			*			**					
63	Stock Exchange	*	*										
175	Mutual/Hedge Funds	*							*	*			
25	M&A							*	*				
Macroeconomic and trade													
107	White House Policy†	***											
125	Tax Policy	**											
28	Interest Rates							*			*		
5	Japan Economy			*		**	*		*	**			
4	Russia–Ukraine										***	*	
6	Executive Comp.		**	***									
Other													
47	Real Estate										***	**	
155	Public Health										***		
31	NYC Urban†										***	**	*
45	Corporate Losses		*	*								**	
22	Internet/Technology			**									
89	Technology (Apple)												**
86	Tobacco		**										
68	Entertainment†										*		
145	Corporate Earnings				*								
174	Quarterly Earnings											*	
	Crop (not selected)												
	Total	7	5	5	4	1	1	5	4	2	7	5	2

Notes: Significance of the Wald χ^2 Granger non-causality test from the two-step debiased LASSO of Babii et al. (2024). $*p < 0.05$; $**p < 0.01$; $***p < 0.001$. † indicates a negative estimated coefficient. Blank cells indicate $p > 0.05$. The "Crop" topic (top terms: price, market, corn, soybean, grain, farmer, crop, commodity) is not selected by the sg-LASSO in any commodity-maturity specification. Lag order $L = 5$. Sample: October 2005–March 2026.

References

Adämmer, P., Prüser, J., and Schüssler, R. A. (2025). Forecasting macroeconomic tail risk in real time: Do textual data add value? *International Journal of Forecasting*, 41:307–320.

Adämmer, P. and Schüssler, R. A. (2020). Forecasting the equity premium: mind the news! *Review of Finance*, 24(6):1313–1355.

Adjemian, M. K. and Jo, J. (2025). The macroeconomic effects of agricultural supply news. SSRN Working Paper 5405106.

Aka, B. F., Gagnon, M.-H., and Power, G. J. (2025). Commodity option return predictability. *Journal of Futures Markets*, 45:1544–1578.

Ardia, D. and Bluteau, K. (2026). Optimal text-based time-series indices. *International Journal of Forecasting*, 42:44–60.

Babii, A., Ghysels, E., and Striaukas, J. (2024). High-dimensional Granger causality tests with an application to VIX and news. *Journal of Financial Econometrics*, 22(3):605–635.

Baker, S. R., Bloom, N., Davis, S. J., and Kost, K. (2026). Policy news and stock market volatility. *Journal of Financial Economics*, 175:104187.

Barbaglia, L., Consoli, S., and Manzan, S. (2023). Forecasting with economic news. *Journal of Business & Economic Statistics*, 41(3):708–719.

Blei, D. M. and Lafferty, J. D. (2007). A correlated topic model of Science. *The Annals of Applied Statistics*, 1(1):17–35.

Blei, D. M., Ng, A. Y., and Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3:993–1022.

Bybee, L., Kelly, B., Manela, A., and Xiu, D. (2024). Business news and business cycles. *The Journal of Finance*, 79(5):3105–3147.

Cao, A. N., Heckelei, T., Ionici, O., and Robe, M. A. (2024). USDA reports affect the stock market, too. *Journal of Commodity Markets*, 34:100384.

Cao, A. N. and Robe, M. A. (2022). Market uncertainty and sentiment around USDA announcements. *Journal of Futures Markets*, 42(2):250–275.

Chi, Y., El-Jahel, L., and Vu, T. (2024). Novel and old news sentiment in commodity futures markets. *Energy Economics*, 140:108006.

Consoli, S., Barbaglia, L., and Manzan, S. (2022). Fine-grained, aspect-based sentiment analysis on economic and financial lexicon. *Knowledge-Based Systems*, 247:108781.

Cui, C. (2011). Commodity implied volatility surface. Technical report, Bloomberg L.P., Quantitative Research and Development, Equities Team. OVDV methodology documentation, Version 2.7.

Egelkraut, T. M., Garcia, P., and Sherrick, B. J. (2007). The term structure of implied forward volatility: Recovery and informational content in the corn options market. *American Journal of Agricultural Economics*, 89(1):1–11.

García, D. (2013). Sentiment during recessions. *The Journal of Finance*, 68(3):1267–1300.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–574.

Goodwin, B. K., Harri, A., Rejesus, R. M., Coble, K. H., and Knight, T. O. (2014). Actuarial review for price volatility factor methodology. Technical report, Sumaria Systems, Inc., prepared for the USDA Risk Management Agency.

Goyal, R. and Adjemian, M. K. (2023). Information rigidities in USDA crop production forecasts. *American Journal of Agricultural Economics*, 105:1405–1425.

Granger, C. W. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3):424–438.

He, X. and Bian, X. (2025). The announcement volatility risk premium in agricultural markets: Evidence from USDA reports. *Journal of Agricultural Economics*, 76:651–665.

Hong, Y., Jiang, F., Meng, L., and Xue, B. (2025). Forecasting inflation using economic narratives. *Journal of Business & Economic Statistics*, 43(1):216–231.

Irwin, S. H. and Sanders, D. R. (2012). Financialization and structural change in commodity futures markets. *Journal of Agricultural and Applied Economics*, 44(3):371–396.

Kelly, B., Manela, A., and Moreira, A. (2021). Text selection. *Journal of Business & Economic Statistics*, 39(4):859–879.

Lee, K. and Park, E. (2024). Exploring calendar effects: the impact of WASDE releases on grain futures market volatility. *Applied Economics Letters*.

Li, Z., Wang, Z., and Diersen, M. (2024). Do agricultural commodity price spikes always stem from news? *Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*.

Massa, O. I., Karali, B., and Irwin, S. H. (2024). What do we know about the value and market impact of the U.S. Department of Agriculture reports? *Applied Economic Perspectives and Policy*, 46:698–736.

Qayum, M. and Janzen, J. (2025). A sentiment-based approach to convenience yield dynamics in commodity futures markets. *Proceedings of the NCCC-134 Conference on Applied Commodity Price Analysis, Forecasting, and Market Risk Management*.

Rogmann, J., Beckmann, J., Gaschler, R., and Landmann, H. (2024). Media sentiment emotions and consumer energy prices. *Energy Economics*, 130:107278.

Simon, D. P. (2002). Implied volatility forecasts in the grains complex. *Journal of Futures Markets*, 22(10):959–981.

Tang, K. and Xiong, W. (2012). Index investment and the financialization of commodities. *Financial Analysts Journal*, 68(6):54–74.

Thorsrud, L. A. (2020). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, 38(2):393–409.

Trujillo-Barrera, A., Mallory, M., and Garcia, P. (2012). Volatility spillovers in US crude oil, ethanol, and corn futures markets. *Journal of Agricultural and Resource Economics*, pages 247–262.

Yang, M. and McKenzie, A. M. (2025). Do corn options update volatility expectations in the wake of USDA reports? *Journal of Futures Markets*, 45:1852–1868.