Three Essays on High-Frequency Return and Volatility Dynamics in Commodities and Financial Futures Markets

Thèse

Simon-Pierre Boucher

Sous la direction de:

Marie-Hélène Gagnon, directrice de recherche
Gabriel Power, codirecteur de recherche

Résumé

Cette thèse se compose de trois essais qui exploitent des données à haute fréquence pour étudier la dynamique des rendements et de la volatilité sur les marchés de contrats à terme de matières premières et d'actifs financiers. Le fil conducteur est l'identification fine, à l'échelle de la minute, des mécanismes par lesquels l'information — annonces macroéconomiques, activité des intervenants, communications de banque centrale — se transmet aux prix et à la volatilité, mécanismes que les données quotidiennes ne permettent pas de distinguer.

Le premier essai examine si l'intensité de la spéculation sur les marchés à terme de l'énergie amplifie ou atténue la réaction des prix aux nouvelles macroéconomiques. À partir de données à cinq minutes sur six contrats à terme américains (pétrole brut, gaz naturel, or, argent, cuivre, palladium) d'avril 2007 à février 2024, de surprises standardisées tirées de 26 annonces macroéconomiques et d'une mesure d'intensité spéculative construite sur les positions désagrégées de la CFTC, l'essai montre qu'une activité spéculative plus élevée *atténue* la réaction des rendements et de la volatilité aux surprises et resserre les écarts acheteur-vendeur. Ces effets stabilisateurs — plus marqués pour les matières premières procycliques que pour l'or, valeur refuge — proviennent des *money managers* plutôt que des *swap dealers*, ce qui suggère un effet bénéfique de la spéculation sur la liquidité et la découverte des prix.

Le deuxième essai construit une base de données inédite de valeurs liquidatives indicatives (iNAV) à la minute pour quatre fonds négociés en bourse (FNB) de matières premières — or, argent, pétrole et gaz naturel — et mesure la transmission de volatilité entre chaque FNB et son panier sous-jacent. En décomposant la variance réalisée en composantes continue et de sauts et en combinant modèles HAR et autorégressions vectorielles bayésiennes, l'essai établit que la transmission passe principalement par les sauts plutôt que par la diffusion, qu'elle est nettement plus forte dans les données à une minute qu'à trente minutes, et que sa direction dépend du mécanisme d'arbitrage : unidirectionnelle — de l'iNAV vers le FNB — pour les métaux précieux détenus physiquement, bidirectionnelle et asymétrique pour les FNB énergétiques adossés à des contrats à terme.

Le troisième essai décompose les communiqués du FOMC en deux dimensions sémantiques — le ton de politique monétaire (*hawkish*/*dovish*) et la nouveauté informationnelle par rapport au communiqué précédent — à l'aide d'un ensemble de modèles de langage (MiniLM et BERT) entraînés sur les communications de la Réserve fédérale, avec une sélection des références fondée sur l'analyse en composantes principales. Sur 148 annonces du FOMC (2008–2025) et des données à une minute pour sept contrats à terme, l'essai montre que le ton prédit les rendements directionnels, tandis que la nouveauté prédit la volatilité ; l'interaction des deux dimensions sur le VIX constitue l'effet le plus robuste. Les deux canaux suivent des dynamiques opposées — résolution rapide de l'incertitude pour la volatilité, réévaluation progressive sur deux heures pour les rendements — ce qui est incompatible avec un canal d'information unique.

Ensemble, ces trois essais montrent que la granularité intrajournalière n'est pas un raffinement technique mais une condition d'identification : l'ampleur, la direction et même le signe des effets documentés seraient invisibles ou biaisés à fréquence quotidienne. Les résultats éclairent des débats de politique publique — réglementation des positions spéculatives, conception des produits indiciels, stratégie de communication des banques centrales — en identifiant précisément qui transmet l'information aux prix, par quel canal et à quel horizon.

Abstract

This dissertation consists of three essays that use high-frequency data to study return and volatility dynamics in commodities and financial futures markets. The common thread is the minute-level identification of the mechanisms through which information — macroeconomic announcements, trader positioning, and central bank communications — is transmitted to prices and volatility, mechanisms that daily data cannot disentangle.

The first essay asks whether the intensity of speculative trading in energy futures markets amplifies or dampens the reaction of prices to macroeconomic news. Using 5-minute data on six U.S. futures contracts (crude oil, natural gas, gold, silver, copper, and palladium) from April 2007 to February 2024, standardized surprises from 26 macroeconomic announcements, and a time-varying measure of speculative intensity built from disaggregated CFTC positions, the essay shows that higher speculative activity *dampens* the response of returns and volatility to surprises and narrows bid-ask spreads. These stabilizing effects — stronger for procyclical commodities than for gold, a safe haven — are driven by money managers rather than swap dealers, suggesting that speculative trading improves liquidity and price discovery.

The second essay builds a novel dataset of minute-level indicative Net Asset Value (iNAV) observations for four commodity exchange-traded funds (ETFs) — gold, silver, crude oil, and natural gas — and measures volatility transmission between each ETF and its underlying basket. Decomposing realized variance into continuous and jump components and combining HAR models with Bayesian vector autoregressions, the essay establishes that transmission runs primarily through jumps rather than diffusion, that it is substantially larger in 1-minute than in 30-minute data, and that its direction depends on the arbitrage mechanism: unidirectional from iNAV to ETF for physically backed precious metals, bidirectional and asymmetric for futures-based energy ETFs.

The third essay decomposes FOMC statements into two semantic dimensions — policy tone (hawkish/dovish) and informational novelty relative to the previous statement — using an ensemble of language models (MiniLM and BERT) trained on Federal Reserve communications with data-driven, PCA-based reference selection. Across 148 FOMC announcements (2008–2025) and 1-minute data for seven futures contracts, the essay shows that tone predicts directional returns while novelty predicts volatility, with the tone–novelty interaction on VIX

being the most robust effect. The two channels display opposite dynamics — rapid uncertainty resolution for volatility, gradual two-hour repricing for returns — a pattern inconsistent with a single information channel.

Together, the essays show that intraday granularity is not a technical refinement but a condition for identification: the magnitude, direction, and even sign of the documented effects would be invisible or biased at the daily frequency. The results inform policy debates on speculative position limits, index-product design, and central bank communication strategy by identifying who transmits information to prices, through which channel, and at what horizon.

Contents

List of Tables

List of Figures

Remerciements

Je tiens d'abord à remercier ma directrice de recherche, Marie-Hélène Gagnon, et mon codirecteur, Gabriel Power, pour leur encadrement, leur disponibilité et leur confiance tout au long de mon parcours doctoral. Leurs conseils rigoureux et leur générosité intellectuelle ont façonné chacun des trois essais de cette thèse, et bien au-delà, le chercheur que je suis devenu.

Je remercie enfin mes parents, Richard et feu Sylvie, à qui je dois tout. Leur soutien indéfectible et les valeurs qu'ils m'ont transmises m'ont porté jusqu'ici. Cette thèse est aussi la leur.

[À COMPLÉTER — ajouter, s'il y a lieu : membres du comité, examinateur externe, organismes de financement, collègues et proches.]

Avant-propos

Déclaration sur l'utilisation de l'intelligence artificielle générative

Dans le cadre de la préparation de cette thèse, j'ai utilisé l'outil d'intelligence artificielle générative Claude (Anthropic) pour écrire le code informatique servant à l'entraînement et à l'application des modèles de plongements lexicaux (*embeddings*) du chapitre 3 (ensemble bi-modèle MiniLM et BERT). J'ai conçu l'architecture de l'analyse, supervisé la production de ce code, puis vérifié, testé et validé l'ensemble des programmes et des résultats qui en découlent. L'IA générative n'a pas été utilisée pour produire le contenu scientifique de la thèse : les questions de recherche, l'analyse, l'interprétation des résultats et la rédaction du texte demeurent entièrement les miennes et celles de mes coauteurs.

Renseignements sur les articles insérés

Cette thèse comprend trois articles, chacun constituant un chapitre. Les travaux de recherche ont été réalisés au cours de mes études doctorales, sous la direction de Marie-Hélène Gagnon et la codirection de Gabriel Power. Pour chacun des trois articles, je suis premier auteur : j'ai contribué à la conception de la recherche, construit les bases de données, réalisé l'ensemble des analyses empiriques et rédigé les manuscrits. Mes coauteurs, Marie-Hélène Gagnon et Gabriel J. Power (Université Laval), ont contribué à la conception de la recherche, à l'interprétation des résultats et à la révision des manuscrits.

Chapitre 1 — Speculative Trading in Energy Markets : Evidence from Macroeconomic Surprises

Cet article est coécrit avec Marie-Hélène Gagnon et Gabriel J. Power. Statut : [À COMPLÉTER — la source indique une version révisée soumise à *The Energy Journal* (« Revised version for The Energy Journal », avec remerciements à l'éditeur George Filis et à deux évaluateurs anonymes) ; préciser l'état exact : soumis le . . ., en révision, accepté le . . ., publié le . . ., et les modifications entre la version insérée et la version publiée, s'il y a lieu.] Les auteurs remercient les participants de séminaires à Humboldt University Berlin, South Dakota State University (Ness School), University of Illinois Urbana-Champaign (ACE), ainsi que les participants

aux congrès de la Commodity & Energy Markets Association (2021), de la World Finance & Banking Association (2021), de la Société canadienne de sciences économiques (2022), de la 4e conférence Ethical Finance and Sustainability (2022) et de la Multinational Finance Society (2022) ; Jocelyn Grira, Joseph Marks et Alessandro Melone (rapporteurs), ainsi que Scott Irwin, Michel Robe et Zhiguang Wang ; pour le soutien financier, le Conseil de recherches en sciences humaines (CRSH) et la Chaire Industrielle-Alliance Groupe financier.

Chapitre 2 — Seeing Through the ETF : Indicative NAV and Commodity Volatility Transmission

Cet article est coécrit avec Marie-Hélène Gagnon et Gabriel J. Power. Statut : [À COMPLÉTER — la source indique une version préparée pour soumission au *Journal of Futures Markets* ; préciser si l'article a été soumis et à quelle date.]

Chapitre 3 — Returns and Volatility Around FOMC Announcements : A High-Frequency Analysis of Policy Tone and Novelty

Cet article est coécrit avec Marie-Hélène Gagnon et Gabriel J. Power. Statut : [À COMPLÉTER — manuscrit daté du 6 mars 2026 dans la source ; préciser s'il est non soumis, soumis (revue et date) ou accepté.]

Introduction

Futures markets are where much of the world's price discovery takes place. Crude oil, natural gas, gold, equity indexes, Treasury notes: for each of these assets, the futures contract is typically the most liquid instrument, the first to react to news, and the reference point from which spot prices, inventories, and investment decisions are set. Understanding how information is impounded into futures prices — and how that process shapes volatility — is therefore central to asset pricing, to market design, and to a series of policy debates that have accompanied the transformation of these markets over the past two decades: the growth of speculative and index-based trading in commodities (Tang and Xiong, 2012; Cheng and Xiong, 2014), the rise of exchange-traded funds (ETFs) as the retail gateway to commodity exposure, and the increasing weight of central bank communication as a market-moving event in its own right.

A common empirical obstacle runs through these debates. Information arrives in minutes, but much of the literature has measured its effects in days. Daily data conflate the announcement of interest with everything else that happens the same day, attenuate estimates through microstructure noise and intraday reversals, and are silent about the dynamics — amplification, absorption, resolution of uncertainty — that unfold in the first minutes and hours after an information event (Andersen et al., 2003; Kothari and Warner, 2007). The high-frequency literature pioneered by Andersen and Bollerslev (1998) and extended by realized-measure econometrics (Andersen et al., 2001; Barndorff-Nielsen and Shephard, 2004; Corsi, 2009) showed that intraday data are not merely more precise: they change what can be identified. Event windows measured in minutes isolate the causal effect of a single release; realized variance turns latent volatility into an observable; and the separation of continuous and jump components distinguishes gradual information diffusion from discrete repricing.

This dissertation applies that identification strategy to three questions about commodities and financial futures markets. Each essay constructs a new high-frequency dataset, each exploits a well-defined information event or transmission channel, and each documents effects whose magnitude, direction, or even sign would be invisible at the daily frequency. The common thread is the same throughout: *who* transmits information to prices, through *which channel*, and at *what horizon*.

Essay 1: Speculative trading and macroeconomic surprises in energy markets

The first essay addresses one of the most persistent controversies in commodity market policy: whether speculative trading destabilizes energy markets. The commodity price run-up of 2004–2008 spawned an influential narrative — the "Masters hypothesis" — according to which the growing presence of financial traders amplifies price movements and volatility, with harmful consequences for the real economy. The theoretical literature shows that such distortions are possible (Basak and Pavlova, 2016; Goldstein and Yang, 2022), but the empirical evidence, largely based on daily data and Granger-causality designs, is inconclusive (Fattouh et al., 2013; Irwin and Sanders, 2012).

The essay takes a different route. Macroeconomic announcement releases provide sharply identified, exogenous information shocks whose surprise component can be standardized following Balduzzi et al. (2001). Using 5-minute futures data for two energy commodities (crude oil and natural gas) and four metals (gold, silver, copper, and palladium) from April 2007 to February 2024, together with a time-varying, commodity-specific measure of speculative intensity built from the CFTC's disaggregated Commitments of Traders data, the essay asks whether the reaction of returns, volatility, and bid-ask spreads to macroeconomic surprises is amplified or dampened when speculative trading is more intense. The answer is unambiguous: the interaction between surprises and speculative intensity is consistently stabilizing. Prices and conditional volatility react *less* to surprises when speculative positions are larger, and bid-ask spreads narrow rather than widen. Disaggregating trader categories shows that these beneficial effects are driven by money managers — informed traders in the sense of Cheng et al. (2015) — rather than swap dealers. For energy markets, where volatility feeds directly into household energy costs and investment under uncertainty, the finding that speculation dampens rather than amplifies news-driven volatility speaks directly to the design of position limits and to the broader regulatory debate.

Essay 2: The iNAV and volatility transmission between commodity ETFs and their underlying assets

The second essay turns from the futures market itself to the fastest-growing channel through which investors access commodities: exchange-traded funds. An ETF is an equity-like wrapper around an underlying basket — physical bullion for gold and silver funds, futures positions for oil and natural gas funds — and the arbitrage mechanism that keeps the wrapper aligned with its contents is also a conduit for volatility (Ben-David et al., 2018). Measuring that conduit precisely requires observing the value of the basket at the same frequency as the ETF price. The essay's central contribution is the construction of a novel minute-level dataset of indicative net asset values (iNAV) for four commodity ETFs (GLD, SLV, USO, UNG), validated against official NAVs and exchange-disseminated values, over more than a decade.

With ETF and basket volatilities observable minute by minute, the essay measures transmission using HAR-type cascade models (Corsi, 2009) extended with cross-asset terms, decomposes realized variance into continuous and jump components, and complements the analysis with Bayesian vector autoregressions. Three results stand out. First, the direction of transmission mirrors the arbitrage technology: for physically backed precious metal funds, volatility flows one way, from the basket to the ETF; for futures-based energy funds, transmission is bidirectional and asymmetric. Second, transmission operates primarily through jumps rather than through the continuous component — a channel that standard connectedness measures, built on daily total volatility, largely obscure. Third, sampling frequency matters quantitatively: 1-minute data reveal transmission up to twice as large as 30-minute estimates. Beyond the substantive findings, the essay demonstrates that the informational content of the iNAV — a series disseminated in real time but rarely archived or studied — provides a sharper image of the ETF–underlying relationship than price-based proxies.

Essay 3: Policy tone, informational novelty, and the high-frequency reaction to FOMC announcements

The third essay studies the information events that arguably matter most for financial futures: announcements of the Federal Open Market Committee. A large literature measures monetary policy surprises through interest-rate futures (Kuttner, 2001; Gürkaynak et al., 2005) or through the joint reaction of rates and equities (Nakamura and Steinsson, 2018). Yet FOMC statements are texts, and their market impact depends not only on what they imply for the policy rate but on how they say it. The essay decomposes each statement into two semantic dimensions: *policy tone* — the hawkish-versus-dovish orientation of the language — and *informational novelty* — the distance between a statement and its predecessor. The measurement apparatus is a dual-model ensemble of transformer language models (MiniLM and BERT) fine-tuned on Federal Reserve communications, with reference statements selected by a data-driven, PCA-based procedure that removes researcher discretion.

Combining these measures with 1-minute data for seven futures contracts — equities, VIX, Treasuries, the dollar, crude oil, and gold — across 148 FOMC announcements between 2008 and 2025, the essay documents a clean division of labour between the two dimensions. Tone predicts directional returns: a one-standard-deviation dovish shift is associated with equity gains that accumulate to roughly twelve basis points over two hours. Novelty predicts volatility, with a sign opposite to the naive prediction: statements that depart more from their predecessor *reduce* uncertainty, consistent with communication acting as clarification rather than noise. The tone–novelty interaction on VIX futures is the most robust effect in the paper, and pre-announcement placebo tests together with five independent inference methods support a causal reading. The two channels also operate on different clocks — volatility effects resolve within roughly twenty minutes, return effects build for two hours — a pattern inconsistent with any single-channel account of central bank communication.

Contributions and structure of the dissertation

Taken together, the three essays make three kinds of contributions. First, they build and document new high-frequency datasets — standardized macroeconomic surprises matched to 5-minute commodity futures, a minute-level iNAV panel for commodity ETFs, and a semantically decomposed corpus of FOMC communications aligned with 1-minute futures data — each of which has value beyond the questions asked here. Second, they deliver identified answers to long-standing questions: speculation stabilizes rather than destabilizes energy futures around news; ETF–underlying volatility transmission is jump-driven and direction-dependent; and central bank communication moves returns through tone but volatility through novelty. Third, they carry policy implications — for the calibration of speculative position limits, for the design and monitoring of commodity index products, and for the communication strategy of central banks.

The remainder of the dissertation is organized as follows. Chapter 1 presents the first essay, on speculative trading and macroeconomic surprises in energy futures markets. Chapter 2 presents the second essay, on iNAV-based volatility transmission between commodity ETFs and their underlying baskets. Chapter 3 presents the third essay, on the high-frequency effects of the tone and novelty of FOMC statements. A general conclusion summarizes the findings, discusses their limitations, and outlines avenues for future research. Appendices A and B collect the mathematical proofs and additional results of Chapter 3.

Chapter 1

Speculative Trading in Energy Markets: Evidence from Macroeconomic Surprises

1.1 Résumé

La spéculation sur les marchés de l'énergie et des matières premières est souvent accusée d'accroître la volatilité, de distordre les prix et de nuire à l'efficience des marchés. Nous adoptons une approche nouvelle, fondée sur les annonces macroéconomiques et des données à haute fréquence, pour mesurer l'effet de l'intensité de la spéculation. Nous étudions l'effet de 26 annonces macroéconomiques sur les contrats à terme énergétiques (pétrole brut, gaz naturel), comparés aux métaux (or, argent, cuivre, palladium). Une intensité spéculative plus élevée atténue l'effet des surprises macroéconomiques sur la dérive des prix, la volatilité et les écarts acheteur-vendeur : la spéculation améliore la liquidité et la découverte des prix tout en réduisant la volatilité. Cet effet d'atténuation est plus marqué pour les matières premières procycliques que pour l'or, valeur refuge. En désagrégeant par type de trader, ces effets bénéfiques proviennent des *money managers* plutôt que des *swap dealers*.

1.2 Abstract

Speculative trading in energy and commodity markets has been blamed for increased volatility, price distortions and market inefficiency, with negative effects on the real economy. We take a new approach to investigate the impact of speculative trading using macroeconomic announcements and high-frequency data. We study the impact of 26 macroeconomic announcement releases on energy commodities (crude oil, natural gas) as our baseline case, which we contrast with metals (gold, silver, copper, and palladium). We find that increased speculative trading lessens the impact of macroeconomic surprises on futures markets, as measured by price

drift, volatility, and bid-ask spreads. Our full-sample results show that increased trading by speculators improves liquidity and price discovery, while reducing volatility. We document a damping effect on volatility that is stronger for procyclical commodities such as crude oil and natural gas than for precious metals such as gold, which is a safe haven. In sub-sample analysis where we separate the effects of money managers and swap dealers, we find that the positive effects that we document are driven by money managers. Since traditional market participants prefer stability, our results suggest a beneficial impact of increased trading and speculation.

Keywords: energy markets, crude oil, futures, high-frequency, speculation, trading, commercial, volatility, macroeconomic, announcements.

1.3 Introduction

The appeal of energy and other commodities as an asset class has grown since the Commodity Futures Modernization Act of 2000 (CFMA). By partially deregulating derivatives, the CFMA has made it easier to trade commodity futures contracts for investment purposes. Rather than invest in physicals or in shares of commodity-linked firms, investors can use futures to gain exposure to "commodity beta" (Boons et al., 2014). This evolution is particularly relevant for energy markets, where futures trading volume has grown tremendously. For instance, the total trading volume of commodity derivatives was 137.3 bn contracts in 2023, which is 64% more than in 2022 (Futures Industry Association, 2024). An important reason for this trend is that commodities have periodically benefited from bull cycles (most notably in 2004-2008), attracting a growing number of speculators and institutional investors. As a result, the commodities asset class has become an important but volatile revenue source for trading firms and investment banks. Three firms, Goldman Sachs, Citi, and Macquarie earned together $20 bn from commodities trading in 2022, much of it from energy-related contracts.[1]

Is the presence of more financial investors harmful to traditional market participants, such as hedgers? While many think so, the evidence is unclear. The idea that poorly informed investors can disrupt markets has a long history (Shleifer and Summers, 1990) and has been revived in a recent theoretical literature on financialization (Basak and Pavlova, 2016; Goldstein and Yang, 2022). These papers are motivated by the commodity price run-up of 2004-2008, which occurred shortly after the CFMA was passed (Domanski and Heath, 2007). Critics argue that the activities of financial investors, who are not directly involved in producing or processing commodities, can distort prices and increase volatility.[2] These critics claim that energy and commodity markets have become more sensitive to financial market fluctuations, and less to

[1]Sources: The Financial Times, Bloomberg, S&P Global and Euronews.

[2]A high-profile example is Masters (2009), who testified before the U.S. Congress in 2008 and before the CFTC in 2009 about "Ending excessive speculation in commodity markets." While his argument is not supported by empirical evidence, as shown by Irwin and Sanders (2012), the Masters hypothesis reflects beliefs held at the time by many market participants.

supply and demand fundamentals. Energy futures markets have attracted attention due to the popular perception that large price swings affect the real economy (e.g., through higher gasoline and heating costs) (Cheng and Xiong, 2014). Research, however, generally does not support this claim (Baumeister and Kilian, 2014). Whether or not these fears are justified, policymakers have taken notice and the CFTC has progressively implemented rule changes such as new position limits.

A large empirical literature debates the causes of periodic price and volatility run-ups in commodity and energy markets. Researchers emphasize the importance of differentiating between speculative traders and passive investors (e.g., index traders). The latter category of traders is more recent and trades energy and commodity contracts for diversification purposes rather than for speculative profit. While Singleton (2014) suggests that financial investors may be to blame for higher energy prices, Kilian and Murphy (2014) use a structural model to show that speculation can be ruled out as a cause of the oil price surge during 2003-2008 – even though speculative demand played a role in previous oil price spikes. Further evidence against the hypothesis that index traders are responsible for the sharp increase in commodity prices is provided by Irwin and Sanders (2011) and Irwin and Sanders (2012). In a different strand of the literature, Büyükşahin and Harris (2011) use Granger causality tests and daily data to investigate whether speculators increase crude oil futures prices. They find little evidence to support that claim. Also using daily position-level data, Brunetti et al. (2016) show that speculators reduce price volatility in commodity and energy markets. Reviewing this early literature, Fattouh et al. (2013) conclude that speculation is unlikely to explain the commodity and energy bull cycle of 2004-2008.

Recent research provides new theoretical grounds to establish how trading activity could affect energy and commodity prices (Basak and Pavlova, 2016; Goldstein and Yang, 2022). The subsequent empirical literature, however, does not reach a consensus. Henderson et al. (2015) use data on commodity-linked notes to show that uninformed trading flows affect commodity prices, but Ready and Ready (2022) argue that the economic magnitude of this effect is too small to matter. Other recent papers find instances of futures price overshooting, reversals, and greater noise in markets (Da et al., 2024). They also find that commodities seem to display higher correlations with equities and with each other (Kang et al., 2023).

Thus, our main contribution is to provide sharply identified evidence on the impact of speculative trading on energy (crude oil and natural gas) and metal markets (gold, silver, copper, and palladium), with additional evidence on sub-categories of traders. We focus on speculation rather than financialization, which has been linked to the trading activities of passive (index) traders (Tang and Xiong, 2012). Specifically, our paper investigates the impact of speculative trading of sub-categories of non-commercial traders. Our rationale is that speculative trading is likely to reflect informed trades, which is our focus, in contrast to index traders who are considered uninformed. Energy commodities serve as our baseline case due to their economic importance and high trading volumes, while metals offer a useful comparison, particularly as gold is perceived as a safe-haven asset. We use high-frequency (5-minute) data to measure the instantaneous reaction of commodity futures returns, volatility, and bid-ask spreads to the surprise component in macroeconomic announcement releases (Andersen et al., 2007; Kurov et al., 2019). The data runs from April 4th, 2007, to February 11th, 2024. Our framework also accounts for the time-varying intensity of speculative trading activity. This study builds on Kilian and Vega (2011), who find no evidence, at a daily frequency, that energy prices react to macroeconomic announcements. By using intraday data, we can better identify the impact of specific macro surprises. We also avoid a common criticism of event study methods, namely that using daily frequency data may reduce the power of statistical tests and could lead the researcher to misattribute the effect of a specific announcement, as other market events occur the same day (Kothari and Warner, 2007).

We find evidence of beneficial effects (price stability and market efficiency) from increased trading activity in energy and commodity markets. Our first finding is a damping effect on price reactions: while macro surprises generate a positive abnormal return for good news (and negative for bad news), the magnitude of this reaction is significantly weaker when speculative trading is higher. Second, we find a similar damping effect on volatility reactions. While all surprises (good or bad) generate a volatility increase, this reaction is lessened when the futures market shows more speculative trading. Third, we document lower bid-ask spreads when speculative trading is higher, controlling for the surprise environment. Fourth and last, these beneficial effects are linked to the trading activities of money managers. In contrast, increased trading by swap dealers appears to have an amplifying effect on reactions to macro surprises. These new insights are made possible by investigating this issue using a new angle, namely their sensitivity to macroeconomic surprises, and with high-frequency data. Our findings have important implications for energy market investment and regulation, and to the broader debate about speculation in energy and commodity markets.

By investigating a broad range of traders in energy and commodity markets, our findings also extend the work of Brunetti et al. (2016). They find that financial investors, especially money managers and hedge funds, help commodity markets by supplying liquidity, reducing volatility, and generally improving market efficiency. Moreover, our results relate to Cheng et al. (2015) who show that financial investors, being better informed about markets, contribute to price discovery and liquidity. This is particularly relevant for energy markets, where accurate price discovery is crucial for physical market participants and investors. Thus, speculative traders help markets by distributing and assimilating new information into prices. These insights are valuable given the ongoing energy transition and the importance of efficient price discovery in energy markets. Two papers are probably closest to ours. First, Brunetti et al. (2016) who find that hedge funds add liquidity to commodity markets, resulting in more efficient prices and lower volatility. They argue that it is merchant positions (i.e., hedgers) that are linked to greater volatility, and that the presence of hedge funds allows for faster and more efficient price discovery. Second, using daily data, Kilian and Vega (2011) study how energy prices react to macroeconomic announcements. Our paper extends this work to high frequency data.

1.4 Background

Price discovery in commodity markets occurs mainly in futures markets and is affected by informational frictions around supply and demand. Thus, risk sharing and information discovery represent a potential channel for financial investors to generate distortions in energy and commodity markets (Cheng and Xiong, 2014). In their model, Basak and Pavlova (2016) predict that the increased presence of financial actors can increase commodity futures volatility, as well as correlations between commodity and equity returns. Goldstein and Yang (2022) also argue that under some conditions, a greater presence of financial investors can be harmful to commodity markets. Theory shows how a change in the participant mix in commodity markets could generate undesirable distortions, but the empirical literature is far from settled. Singleton (2014) argues that trading activity by financial investors creates informational frictions, leading commodity prices to become more volatile and to diverge from their fundamental values. Using a no-arbitrage argument, however, Hamilton and Wu (2014) show that the positions of commodity traders included in index funds cannot be used to achieve excess returns in futures markets. Ready and Ready (2022) show that while index traders do have a positive price impact, it is much too small to explain the apparent price distortions or bull cycles observed since 2004. In addition, financial investors do not have a uniform impact on market liquidity. Investors affect liquidity risk by either providing liquidity to meet the hedging needs of other traders or consuming liquidity when they trade for their own needs (Kang et al., 2020). Indeed, Brunetti and Reiffen (2014) show using data on commodity trader positions that index traders provide insurance against price risk.

1.4.1 Macroeconomic announcements

Surprises in macroeconomic announcements affect financial markets, whether in stocks (Scholtus, Van Dijk, and Frijns, 2014) or in bonds (Fleming and Remolona, 1997). In a key study, Balduzzi, Elton, and Green (2001) find that 17 public news releases affect bond prices, trading volume, and bid-ask spreads. Karali and Ramirez (2014) show that energy futures markets exhibit asymmetric responses to macroeconomic news, with significant volatility spillovers between natural gas and crude oil markets. Cao et al. (2024) document a time-varying relationship between U.S. monetary policy and crude oil prices, finding that unexpected oil price increases can push monetary policy from expansionary to restrictive stance. Kang et al. (2020) further show that after 2004, short-term oil price volatility is driven by industrial production, term spreads, and credit spreads, along with traditional market factors.

The literature on commodity-specific announcements is smaller and less conclusive. Hollstein, Prokopczuk, and Würsig (2020) look at how different economic variables affect the term structure of commodity futures volatility. They show that speculation and jobs-related macro variables have the largest impact on volatility. Zhu et al. (2022) further show that stock market anomalies can be explained to some extent by oil price shocks, separately from the effect of other macroeconomic variables and investor sentiment. While the literature finds a clear impact of macroeconomic announcements on stock and bond prices, there is no clear answer as to whether they affect commodity futures prices, or whether increased trading by financial participants accentuates these reactions. This issue is especially relevant for energy markets, given their macroeconomic importance. Our research provides new insights by using high-frequency data, expanding the set of announcements, and considering a time-varying measure of speculative trading intensity to capture trading activities for each of the commodities in the sample.

1.5 Data

We now present a detailed description of our data. Since this paper relies on several types of data, we describe: i) how to obtain the macroeconomic announcement surprises, ii) the commodity futures data, and iii) how to capture speculative trading activity.

1.5.1 Data on macroeconomic announcements

The macroeconomic announcement release data are obtained from Bloomberg and Refinitiv Eikon. We collect information on 22 announcements that are standard to the literature (see e.g., Andersen et al., 2003). Our sample for macroeconomic announcements is matched to our high-frequency data and therefore runs from April 2nd, 2007 to February 11th, 2024. The announcements belong to ten categories: Income, Employment, Industrial Activity, Investment, Consumption, Housing Sector, Government, Net Exports, Inflation, and Forward-looking. Most of the announcements are released on a monthly basis. Table 1.2 summarizes the announcements and provides more detail such as the number of observations, release frequency, source, unit of measure, and time of release. Bloomberg provides analyst forecasts for all announcements, as well as the actual value of the announcement release. For all announcements except the Consumer Price Index and Initial Jobless Claims releases, a positive surprise will be interpreted by investors as signaling a strong economy (Fleming and Remolona, 1997). In addition, we include energy sector-specific announcements published by the U.S. Energy Information Administration. The first is the weekly crude oil storage report, which provides an update on the quantity of crude oil held in storage in the U.S. The second is the weekly natural gas storage report. We do not include OPEC announcements, as they cannot be reliably used in a high-frequency econometric design (Känzig, 2021).[3]

[3]There are a few issues with the OPEC announcements: First, they are not released at a specific time. Second, it is impossible to know precisely when a given OPEC announcement was made available to investors.

It is common practice in this literature to use the standardized surprise of an announcement rather than its realized value to quantify the unexpected component of the release. To calculate surprises, we follow Balduzzi, Elton, and Green (2001). Let A_{kt} be the realized value (i.e., release) of macroeconomic announcement k at time t, and let E_{kt} be the median value of all Bloomberg analyst forecasts for announcement k at time t. To standardize the surprise, we divide the raw surprise $(A_{kt} - E_{kt})$ by σ_k, the sample standard deviation of the surprise for announcement k. Thus, equation (1.1) describes the standardized surprise for announcement k at time t:

$$S_{kt} = \frac{A_{kt} - E_{kt}}{\sigma_k} \tag{1.1}$$

The sample period is used to compute σ_k, as in Balduzzi et al. (2001) and Kurov et al. (2019).[4] Table 1.3 presents the minimum, 1st quartile, median, mean, 3rd quartile, and maximum of the surprise for each announcement.

1.5.2 Commodity futures price data

For intraday data on commodity futures prices, we use Barchart's API.[5] Our dataset for prices contains some of the most economically significant commodity futures contracts traded in the U.S. We use a high-frequency price series that runs from April 2nd, 2007 to February 11, 2024. Among these contracts, crude oil and natural gas are pro-cyclical, while gold and silver behave as safe havens. High-grade copper and palladium are industrial metals used in the manufacturing of consumer products.

For each of the commodities in our sample, price returns R_t are calculated as the log return over a 5-minute period ($\tau = 5$) beginning at time t. The database provides the futures contract close price (p_t^{close}) of each 5-minute period. Thus, R_t is obtained as in equation (1.2):

$$R_t^{t+\tau} = \ln\left(\frac{p_{t+\tau}^{close}}{p_t^{close}}\right) = \ln(p_{t+\tau}^{close}) - \ln(p_t^{close}) \tag{1.2}$$

Descriptive statistics for the 5-minute log returns are presented in Table 1.4. The most extreme outlier observations belong to crude oil, while gold has the fewest outliers.[6]

Third, OPEC's influence has weakened since the 1980s.

[4]The literature argues that measuring σ_k in this way is reasonable because the standardized surprise is not used for forecasting purposes. Using raw surprises is not recommended due to scaling issues, nor is using analyst dispersion for σ_k because announcement coverage sometimes involves only a few analysts. For robustness, we also estimate our models using surprises where σ_k is computed using only past observations. The main findings are unchanged. In this case, we exclude the first M observations (e.g., $M = 10$) to get a reasonable sample size for σ_k.

[5]See the `https://www.barchart.com/futures` website.

[6]The main findings are robust to using different window lengths by estimating equation 1.2 using 30-minute returns.

1.5.3 Measures of speculative trading activity and trader categories

The index of speculative trading is constructed using data in the *Commitment of Traders (CoT) Report* published weekly by the Commodity Futures Trading Commission (CFTC). The data provided by the CFTC includes the number of positions held by different types of participants in commodity markets. The CFTC separates trader types as follows: *Commercials* refer to trader-reported futures positions which the trader claims are used for hedging purposes, while *Non-Commercials* is obtained by subtracting the total long and short commercial positions from the total open interest.[7] We use the following information presented in the *CoT* report: for a futures contract i, the number of long and short positions held by Non-Commercial traders are SL_i and SS_i, respectively, while for Commercial traders they are HL_i and HS_i.[8]

The specific measure we use follows Hedegaard (2011), who suggests an index of speculative activity computed as the ratio of net long speculative positions over total open interest (NLS_i):

$$NLS_i = \frac{SL_i - SS_i}{OI_i} \tag{1.3}$$

In addition to computing NLS using the full sample data, we use disaggregated data from the CFTC to compute the NLS index separately for money manager (MM) and swap dealer (SD) positions, which allows for additional empirical analysis. The data on money manager and swap dealer positions come from Quandl's API.[9] *Money managers* typically refer to Non-Commercial market participants who are involved in managing funds and investing in commodity futures and options markets (Fishe and Smith, 2012).[10] Their activities are influenced by financial and economic factors related to commodities. Money managers are often considered to be more informed investors because they actively manage portfolios and adjust their positions based on market information and analysis. *Swap dealers* are considered as Non-Commercial traders by the CFTC. They typically use futures contracts to hedge risk generated by their swap positions. Swap dealers have been studied in relation to index investors, as their positions are distinct from those of other market participants. Their activities are influenced by the need to manage large exposures and to facilitate trading for clients. While money managers regularly take long

[7] The CFTC defines commercial traders as participants in commodity markets who primarily use futures contracts to hedge their business activities (e.g., buying or selling commodities). All traders who are not classified as Commercial are automatically classified as Non-Commercial traders. To obtain the number of long positions held by Non-Commercial traders, we subtract the total long Commercial positions from the total open interest. For the number of short positions held by Non-Commercial traders, we subtract the total short Commercial Positions from the total open interest.

[8] In an earlier draft, we also reported results based on two alternative proxies as well as a proxy constructed using principal component analysis. The alternative proxies are Working's T (Working, 1960) and the market share of non-commercial traders (Büyükşahin and Robe, 2014). These results, which are available upon request, are consistent with our main findings and do not change the paper's implications.

[9] See the `https://data.nasdaq.com/data/CFTC-commodity-futures-trading-commission-reports` website.

[10] This category is also called "Managed money." The CFTC writes that they are "registered commodity trading advisor (CTA); a registered commodity pool operator (CPO); or an unregistered fund identified by CFTC." There is some overlap between Money managers and hedge funds, but they are distinct.

or short futures positions, swap dealers mainly take long positions (Fishe and Smith, 2012). Hedgers, who we exclude from the analysis, tend to take short positions.

1.6 Econometric framework and methods

1.6.1 Modeling the impact of surprises on returns

Our high-frequency regression model is based on Kurov et al. (2019).[11] We run the following regression using the specification in equation (1.4):

$$R_t^{t+\tau} = \alpha + \sum_{m=1}^{22} \gamma_m S_{m,t} + \delta X_j + \sum_{m=1}^{22} \theta_m (S_{m,t} \cdot X_j) + \beta R_{t-\tau}^{t} + \epsilon_t \tag{1.4}$$

where $R_t^{t+\tau}$ is the continuously compounded futures return from time t to $t+\tau$, S_{mt} is the surprise for macroeconomic announcement m published at time t, and X_j is the NLS speculative trading intensity variable, which is updated at a weekly frequency, with j the index for the week. The impact of macro announcements on commodity futures returns can be assessed by looking at the γ_m coefficient in the mean equation, while the δ coefficient controls for the level of speculative trading intensity as it relates to futures returns. The key coefficient to help answer our main research question is θ_m, which relates the effect of time-varying speculative trading intensity on the impact of the news release.The regression is estimated using a two-step weighted least squares (WLS) procedure. To account for heteroskedasticity, we construct a volatility estimate by means of an exponential moving average, using the regression residuals obtained in the first step. This auxiliary regression is presented in equation (1.5), with a smoothing parameter $\alpha = 0.9$ and a starting parameter value set to $\sigma_1 = \epsilon_t$:

$$\sigma_t = \alpha \sigma_{t-1} + (1-\alpha) \mid \epsilon_t \mid \tag{1.5}$$

After obtaining σ_t for each observation, we apply the transformation $w_t = \hat{\sigma_t}^{-2}$ to obtain the WLS regression weight. Then, we multiply each variable by w_t and run an OLS regression to estimate the model.

1.6.2 Modeling the impact of surprises on volatility

To estimate the volatility equation, we use a GARCH specification, as it is well known that the variance of commodity futures returns displays time variation and clustering (see e.g., Brunetti and Reiffen, 2014). We specify a GARCH (1,1) model and extend the equation by including our NLS speculative intensity proxy as well as the macroeconomic news surprise variables. First, we estimate the mean equation (1.6):

[11] In unreported results, we run the regressions using the approach shown in Andersen et al. (2003). The results are similar.

$$R_t^{t+\tau} = \alpha + \sum_{m=1}^{22} \gamma_m S_{m,t} + \beta R_{t-\tau}^t + \epsilon_t \tag{1.6}$$

Then, we estimate the following equation for conditional variance:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \alpha_2 \epsilon_t^2 + \sum_{m=1}^{22} \Phi_m D_{m,t} + \beta X_j + \sum_{k=1}^{n} \phi_k I_{kt} + \sum_{h=1}^{23} \rho_h D_h \tag{1.7}$$

where $I_{k,t} = D_{m,t} \cdot X_j$ and $D_{m,t}$ is a dummy variable for macro announcement m. The latter equals 1 if an announcement takes place at time t (5-minute frequency) and equals 0 otherwise. X_j is the NLS speculative intensity variable as defined earlier. The ρ_h coefficient captures intraday periodicity, while the D_h dummy equals 1 at hour h and 0 otherwise. The impact of macro announcement m on conditional variance is captured by the Φ_m coefficient in equation (1.7), while the β coefficient shows the impact of the speculative intensity variable X_j. Finally, the ϕ_k coefficient shows the interaction effect from speculative trading and the macro surprise m. The standard errors are computed using the Newey-West heteroskedasticity and autocorrelation consistent (HAC) estimator with automatic lag selection, following the procedure outlined in Newey and West (1994).This methodology mirrors the approach used in Andersen et al. (2003, 2007) and Kurov et al. (2019) to study announcement effects in other asset classes. In unpublished results, we consider the mixed-data sampling (MIDAS) approach proposed by Ghysels et al. (2004). We find that the results are similar.

1.6.3 Modeling the impact on bid-ask spreads

Speculative trading could make markets more efficient by improving information. We test this hypothesis by measuring the effect of macro surprises on the futures price bid-ask spread in high-frequency regressions. The bid-ask spread is widely recognized as a measure of market efficiency. A narrower spread suggests less uncertainty about the asset's true value and reflects lower transaction costs, improved liquidity, and lower information asymmetry (Roll, 1984). Furthermore, Chordia et al. (2008) show that the bid-ask spread is an indicator of market quality and market efficiency. Since a smaller spread is associated with a more efficient price discovery process, an increase in the quality of market information should decrease the spread. Therefore, we estimate the following equation, where the relative bid-ask spread is defined as $(Ask_t - Bid_t)/Mid_t$:

$$\text{Spread}_t = \alpha + \sum_{m=1}^{22} \gamma_m D_{m,t} + \delta X_j + \sum_{m=1}^{22} \theta_m (D_{m,t} \cdot X_j) + \beta \, \text{Spread}_{t-\tau} + \epsilon_t. \tag{1.8}$$

In equation (1.8) , Spread_t is the relative bid-ask spread measured at 5-minute frequency t using the high-frequency data and $\text{Spread}_{t-\tau}$ is the lagged spread. In addition, $\sum_{m=1}^{22} \gamma_m D_{m,t}$

accounts for the macroeconomic announcements, where each dummy variable $D_{m,t}$ is multiplied by its respective coefficient γ_m. We also include the speculative trading intensity variable X_j with its coefficient δ. The interaction terms $\sum_{m=1}^{22} \theta_m(D_{m,t} \cdot X_j)$ capture the effect of speculative trading intensity on the bid-ask spread at the time of a release, with each term multiplied by its respective coefficient θ_m. To test whether increased speculative trading activity improves informational efficiency at the time of a macroeconomic announcement, we check whether the sign on θ_m is negative and significant, thus reducing the spread.

1.7 Results

1.7.1 The impact of surprises on cumulative abnormal returns

We begin by documenting the impact of macroeconomic announcement surprises on commodity futures returns. The impact of surprises is shown in the following graphs of high-frequency cumulative abnormal returns (CARs). Figures 1.1 to 1.6 show CARs for each commodity as measured over a window of 60 minutes before to 60 minutes after a macroeconomic announcement release. The figures are constructed similarly to those shown in Kurov et al. (2019). The magnitude of the CARs after announcement releases is comparable to those shown in their paper for stock index and Treasury futures. Unlike them, however, we do not see evidence of a pre-announcement drift.

A red line denotes the average CAR for announcement releases that are seen as negative surprises (i.e., worse than anticipated news), while a green line denotes the average CAR for positive surprises. For brevity, we discuss only the CARs for crude oil and gold, as they are representative of pro-cyclical energy markets and safe haven assets. Figure 1.1 shows the average CAR for crude oil futures. The CAR increases following a positive surprise and decreases following a negative one, confirming that crude oil is a pro-cyclical commodity. In contrast, figure 1.2 shows that gold futures react in the opposite manner. The red line indicates that CAR is positive after bad news, while the green line shows that CAR is negative after good news.

1.7.2 Macroeconomic surprises, speculative trading activity and futures returns

Table 1.6 presents the results of high-frequency regressions that explain commodity futures returns immediately after a macroeconomic announcement release. Our discussion focuses on coefficients that are statistically significant at the 5% level.[12] We begin with energy commodities, which serve as our baseline case. The γ_m coefficient shows the immediate impact

[12] In an earlier draft, we also reported results for different sub-periods, such as the Zero Lower Bound period, the 2008-2010 financial crisis and Great Recession, and the COVID-19 period. These results do not materially affect our findings or conclusions, and they are available upon request.

of a macro surprise on commodity futures returns. For crude oil futures, we find that several macroeconomic announcements exhibit significant effects. Consider for instance Initial Jobless Claims, for which a greater than expected value indicates bad economic news. The table shows that for this announcement, γ_m is negative, indicating that crude oil prices tend to drop in response to unexpected increases in jobless claims. This result is consistent with the expectation that higher jobless claims signal weaker economic conditions, which reduce the demand for crude oil. The corresponding θ_m coefficient for Initial Jobless Claims is positive, however, suggesting that increased speculative trading mitigates the negative impact of bad macroeconomic news on crude oil prices. This damping effect suggests that markets are better informed as a result of the increased participation of speculators. Indeed, the announcement release creates a smaller surprise and a smaller shock.

In the case of natural gas futures, we find that the results across announcements are less frequently significant. Increased speculative trading tends to increase the magnitude of the surprise's effects, when the results are significant. An exception is the natural gas market-specific announcement release (inventories), for which the coefficient is negative but not significant. Overall for natural gas futures, we do not find as much support that speculative trading dampens the effect of macroeconomic announcements on returns. The reason why the results for natural gas futures are less conclusive is most likely that this futures contract displays greater volatility, that it has a lower trading volume (Irwin and Sanders, 2012), and it has less speculative trading activity (thus, fewer informed traders) (Büyükşahin and Robe, 2014).

Copper is a pro-cyclical, industrial commodity, so it is expected that the results for copper futures should resemble those for crude oil. The results are highly significant for many announcements. If we look at announcements such as ADP Employment and CB Consumer Confidence, which are "good news", the γ_m coefficients are positive, confirming copper's pro-cyclical nature and its ties to industrial production. The corresponding θ_m coefficients are negative, consistent with the damping effect that we document for energy commodities. Thus, the main finding for copper futures is that, as with crude oil, increased speculative trading intensity has the effect of weakening the impact of a macro surprise.

Gold is considered to be a safe haven asset (Baur and Lucey, 2010). Therefore, it is expected that the reaction of gold futures returns to macro surprises will be the opposite to what we have found for crude oil, natural gas, and copper. This is indeed what we find: gold futures returns are lower after "good news" and higher after "bad news". These results support the idea that energy commodities are pro-cyclical, while gold is a safe haven asset. In particular, the positive γ_m coefficient for Initial Jobless Claims indicates that gold prices rise in response to unexpected increases in jobless claims, as investors seek safety in gold positions during economic uncertainty. The θ_m coefficient is negative, indicating that speculative trading activity tempers this flight to safety, leading to less pronounced price increases.In the case of the ADP Employment release, the γ_m coefficient is negative for gold futures, suggesting that strong employment figures reduce gold prices as investors move away from safe-haven assets towards pro-cyclical assets, such as energy commodities. The positive θ_m coefficient suggests that speculative trading reduces the extent of this price drop. The CB Consumer Confidence and Advance Retail Sales announcements also generate negative γ_m coefficients for gold futures and positive θ_m coefficients at a 10 percent level, providing additional evidence of a moderating influence of speculative trading on the reaction of gold to economic news. The main finding is therefore that whether a commodity is pro-cyclical or a safe haven asset, increased speculative trading has a damping effect on the reactions to macro surprises.

The last two commodities in our sample, silver and palladium futures, behave more like gold futures. We find that for Initial Jobless Claims for palladium and ADP Employment in silver, significant γ_m coefficients are found in directions consistent with their status as safe haven assets. The θ_m coefficients for silver and palladium also indicate a damping effect on price reactions. Taken together, our results show that the damping effect of speculative trading is stronger in energy markets, suggesting that the beneficial effects of financial participants are particularly important for energy commodities.

1.7.3 Macroeconomic surprises, speculative trading and volatility

Table 1.7 presents regression results to explain the conditional variance of high-frequency commodity futures returns after macroeconomic announcements. We first examine our baseline assets, energy commodities, as volatility in energy markets is of particular concern given their economic importance and direct impact on consumer prices. For crude oil, macroeconomic surprises generally lead to an increase in conditional variance, as the Φ_m coefficients are consistently positive. For instance, a surprise in Initial Jobless Claims significantly increases crude oil volatility. Natural gas exhibits similar patterns, with some variations in magnitude and a greater impact on inventory-related announcements. This result suggests that unexpected economic news generates greater uncertainty and price fluctuations in energy markets. In contrast, the interaction coefficients ϕ_m, which inform us about the impact of speculative trading, tend to be negative for the two energy commodities. The implication is that increased trading activity by speculative traders lowers volatility following a macro surprise. Taking Initial Jobless Claims as an example, we find that higher levels of speculation reduce the impact of news on volatility in crude oil futures markets. Thus, speculative trading can act as a stabilizing force by damping the heightened volatility that occurs after a surprise in macroeconomic news.

Comparing these results with those for other commodities, we find that gold, copper, silver, and palladium also show positive Φ_m coefficients across various announcements, indicating increased volatility following macro surprises. However, the magnitude of these effects is generally smaller than what we see in energy markets, especially in the case of precious metals. The corresponding ϕ_m coefficients, measuring the damping effect of financial investor activity, are negative across commodities and announcements, and the strongest effects are found in crude oil futures markets. This pattern holds for other macroeconomic announcements as well. The coefficients for ADP Employment, CB Consumer Confidence, and Advance Retail Sales, for instance, generally indicate increased volatility after surprises, as shown by the positive Φ_m coefficients, while the corresponding ϕ_m coefficients are negative, supporting the finding of a stabilizing effect of increased speculative trading.

Therefore, our findings highlight the valuable role of speculative traders in reducing volatility, particularly in crude oil futures markets, where price stability has important implications for the broader economy. While macroeconomic news tends to increase volatility across commodity markets, the damping effect of speculative trading appears strongest in energy markets. We show that this relationship is consistent across different types of announcements and remains robust when controlling for various market conditions.

1.7.4 Macroeconomic surprises, speculation and bid-ask spreads

Table 1.8 shows our results for the impact of macroeconomic surprises and speculative trading intensity on futures prices bid-ask spreads. This empirical analysis provides a test of informational efficiency. We first focus on energy markets. In the bid-ask spread regressions, the γ_m coefficient denotes the effect of surprises on the spread, while the θ_m coefficient shows the interaction effect between surprises and speculative trading intensity. The main hypothesis is whether $\theta_m < 0$, which would indicate that greater speculative activity improves informational efficiency through narrower spreads. Such a finding would be consistent with what we have documented above for futures returns and volatility. In the case of crude oil futures, we find that γ_m is negative for the initial jobless claims announcement, which means that the market becomes more efficient immediately after a news release. This is consistent with the resolution of uncertainty. The θ_m coefficient is always negative when it is significant, indicating that the bid-ask spread narrows even more (implying greater informational efficiency) after a macro announcement if crude oil futures markets benefit from greater speculative activity relative to hedging activity, as measured by the NLS proxy. The results for natural gas futures are similar to those for crude oil. The relationship between speculative trading and market efficiency is clearest during periods of higher trading volume, such as the release of storage reports and weather-related announcements.

Comparing these results to those obtained for the other commodities in our sample, we find that the results for θ_m (speculative intensity) are less often significant, but that they are negative when they are significant. The θ_m coefficient being negative suggests that greater speculative trading intensity improves market efficiency, as bid-ask spreads tend to be lower when the coefficient is significant. This improvement in market efficiency appears to be most pronounced in energy markets, where accurate price discovery is particularly important given their relevance for the real economy. Overall, the results for the bid-ask spread provide additional support for our claim that greater trading activity has beneficial effects on commodity derivatives markets, with these benefits being especially notable in energy markets where efficient price discovery has important implications for both market participants and the broader economy.

1.7.5 Differences in results according to trader type

To investigate whether differences in trader type are relevant in explaining our findings, we provide disaggregated results in this section. To this end, we estimate equations (1.6), (1.7) and (1.8) for two categories of Non-Commercial traders, namely, swap dealers (SD) and money managers (MM). For each of the two, the CFTC reports the number of long and short positions in their disaggregated Commitment of Traders (COT) reports. We compute the NLS index for each trader category over time, allowing us to separately quantify the intensity of trading activity by money managers and swap dealers.

First, we examine the returns equation for money manager positions, as shown in table 1.9. Increased trading activity by money managers has the same effect as in our baseline results. If we consider crude oil futures, for example, the γ_m macro surprise coefficient is positive while the θ_m coefficient for speculative trading is negative. Since θ has the opposite sign to γ, the implication is that increased futures trading activity by money managers lowers the impact of surprises on futures returns. This is similar to our baseline, aggregate findings. Next, table 1.10 presents results using only swap dealer positions. Here we find a notable difference relative to money managers. This table shows that the speculative intensity coefficient θ for swap dealers has the same sign as the macro surprise coefficient γ. This result means that increased swap dealer trading activity seems to amplify the reaction of futures returns to macro surprises. The exception to these results is in the case of natural gas futures, where increased trading by swap dealers for some announcements appears to have the opposite effect on returns. To contrast this finding with prior research, Brunetti et al. (2016) find, using daily data, that the positions of swap dealers are not correlated with contemporaneous returns and volatility in commodity futures markets. They further show that hedge funds decrease, and hedgers increase, volatility. Our empirical analysis extends their findings using high-frequency data and the setting of macro surprises as a source of new information affecting energy and commodity futures markets.

Tables 1.11 and 1.12 show our disaggregated results for the variance equation using sample data for money managers and swap dealers, respectively. The money manager results are similar to what we find in the aggregate sample, namely that both good and bad surprises increase volatility (as shown by $\Phi_{mm} > 0$), while greater money manager trading activity lowers the impact of news on volatility (as shown by $\phi_{mm} < 0$). This result strengthens prior evidence in the literature about beneficial effects of speculators, which were based on daily data (Brunetti et al., 2016). In contrast, our results suggest that increased activity by swap dealers appears to increase volatility after macro news (since $\phi_{sd} > 0$). Together, the results line up with our findings for the returns equation and with our main message, which is that informed traders stabilize markets by contributing new information. Swap dealers, as intermediaries, typically do not trade based on information but rather following the needs of their clients. This economic motivation can explain why our results suggest that their trading activities amplify price and volatility reactions to news. Although our discussion centers on crude oil as a benchmark commodity, the results for the other pro-cyclical commodities support our interpretation of the findings. In the case of gold futures, the disaggregated results continue to support a safe haven interpretation (Erb and Harvey, 2013). Indeed, gold has features of a commodity and a currency, but earlier studies have also found that its value increases with investor risk aversion, since it is perceived as a safe haven during times of economic uncertainty and market volatility.

Lastly, we examine whether the effects on bid-ask spreads differ between trader types. Table 1.13 presents the results for money managers. As the θ_m coefficients tend to be negative, it seems that money managers enhance market efficiency by lowering bid-ask spreads. The results for swap dealers are shown in Table 1.14. Unlike for money managers, we find that θ_m coefficients for swap dealers tend to be positive, indicating that greater swap dealer activity increases the bid-ask spread after a macroeconomic release. The results for swap dealers therefore suggest a decline in market efficiency due to their increased trading activities. These findings are in line with our main results and further support the claim that traders in energy and commodity markets do not all have the same effect on market efficiency. The results point to the importance of information acquisition and investor attention as an economic channel.

1.8 Discussion and implications

Our findings contribute new insights to an unsettled literature on the impact of different types of financial participants in energy and commodity markets (Ready and Ready, 2022). In addition to being an important alternative asset class, energy commodities are central to economic activity, while energy futures trading is important for price stability. Being closely related to macroeconomic risk (Cheng et al., 2015), energy commodities therefore play a key role in macro-finance. By taking a novel angle of high-frequency market reactions to macroeconomic surprises, our results contribute a more nuanced picture of the impact of speculative trading and help to reconcile previous findings in the literature.

Goldstein and Yang (2022) develop a theoretical model to examine how financial traders improve information transmission between futures and spot markets. They show how market efficiency should benefit from the presence of informed financial traders, whose actions lead to increases in pricing transparency and to lower information asymmetry. Our empirical results provide support for this model. We show that increased speculative trading activity lowers bid-ask spreads after a macro news release, in addition to reducing the magnitude of price and volatility reactions to macro surprises. This improvement in informational efficiency is especially valuable in energy markets, where price discovery has direct implications for industrial users and consumers. We find that this effect is driven by money managers. In addition, our results point to smaller, negative effects linked to swap dealers. The difference between the two sets of disaggregated empirical results can be explained by the fact that they trade for different purposes. Money managers aim to make a profit based on information, while swap dealers are intermediaries who, while perhaps informed, primarily provide services to customers such as institutional investors, in addition to hedging their swap positions. In line with this interpretation, our results also build on Fishe and Smith (2012), who show that some financial traders (e.g., money managers) are better informed than others in commodity markets.

Our findings also extend and build on the results of an earlier empirical literature that uses daily-level data. This literature includes Brunetti et al. (2016) who find that speculative traders, especially money managers and hedge funds, are helpful to energy and commodity markets, as well as Büyükşahin and Harris (2011) and Alquist and Gervais (2013) who show that the futures positions of financial firms such as hedge funds do not predict next-day changes in crude oil prices. Moreover, our results relate to Cheng et al. (2015) who show that financial investors, being better informed about markets, contribute to price discovery and liquidity. The key message is therefore that speculative trading is helpful to markets by distributing and assimilating new information into prices. The findings shown in this paper have important implications for energy markets regulation and policy. First, they suggest that attempts to limit speculative trading in energy markets could, in fact, increase price volatility and reduce informational efficiency and price discovery. Second, they indicate that different types of financial participants have distinct effects on market quality, suggesting that regulatory frameworks should pay careful attention to market composition. Third, they highlight the importance of maintaining a robust price discovery mechanism in energy markets, given their crucial role in the economy.

1.9 Conclusion

This paper investigates the impact of speculative trading on the real economy and energy markets through a new angle, namely high-frequency surprises in macroeconomic announcement releases, which allows for better-identified effects. We empirically test whether increased speculative trading activity amplifies or dampens the impact of macro surprises on prices and volatility in commodity futures markets. Our results suggest that increased speculative trading activity has beneficial effects for energy and commodity markets. This is accomplished by reducing volatility and improving price discovery, as indeed price stability and efficient price discovery are crucial for market participants and the broader economy. We find that a greater intensity of speculative trading does not amplify the effects of macro announcement surprises on prices or volatility. On the contrary, an increase in speculative trading in a given commodity has a damping effect: prices and volatility react *less* to macro surprises when speculative trading is higher. This stabilizing effect of speculation is especially valuable to energy markets, where price volatility can have significant economic consequences.

What is more, our findings are consistent with information diffusion economic arguments. Our analysis of bid-ask spreads in futures contracts further confirms that speculative trading tends to improve market efficiency. Our results show that these effects are mostly linked to the trading activities of money managers rather than swap dealers. Thus, we contribute to a literature that emphasizes how non-commercial market participants such as money managers are beneficial to commodity markets by supplying liquidity, reducing volatility, and generally improving market efficiency. This finding is particularly relevant for energy markets, which have seen substantial increases in trading volume and complexity.

Our findings have important implications for energy market regulation and policy. The damping effect on volatility shocks documented in this paper implies that speculative traders contribute to market stability, which is essential for energy security and economic planning. This stability could also facilitate investment in energy infrastructure and support the ongoing energy transition. By lowering the magnitude of volatility shocks, and thus reducing the real option value of delaying investments, our findings suggest that a greater involvement by speculative traders may also help with sustainability efforts to finance a green energy transition, alongside other instruments such as green bonds and portfolio screens for sustainable investments. Looking forward, our results suggest several promising avenues for future research in energy markets. First, the role of financial investors (speculators as well as passive investors) in facilitating the energy transition warrants further investigation. Second, the connection between speculative trading and energy market regulation remains an important area for study, given that we find different impacts for money managers and swap dealers. Finally, the impact on energy price discovery of new trading technologies and market participants is an emerging research frontier. These questions are particularly relevant given the increasing importance of energy markets in addressing climate uncertainty and in ensuring economic stability.

1.10 Tables

Table 1.1: Summary of the literature: Effect of non-commercial trading activity on commodity futures volatility

References	**Proxy used for financialization or speculation**	**Impact on volatility**
Chang et al. (1997)	CFTC's definition of speculators	**Increase**
Daigler and Wiley (1999)	CFTC's definition of speculators	
Irwin and Holt (2004)	Set speculators	
Tang and Xiong (2012)	Commodity index trader (CIT) positions	
Irwin and Brorsen (1987)	Amount of money invested in traded futures funds	**No change**
Irwin and Yoshimaru (1999)	Trading volume of large-commodity pool operators	
Bryant et al. (2006)	CFTC's definition of speculators	
Haigh et al. (2007)	Number and positions of commodity pool operators and hedge funds	
Brunetti et al. (2016)	Net positions of hedge funds and floor brokers	**Decrease**
Aulerich et al. (2012)	Commodity index trader (CIT) positions	

This table summarizes the findings of a range of studies on the effect of financialization and speculation on commodity futures volatility. The impact on volatility is categorized as Positive, Neutral, or Negative based on the results reported by each study. The proxies used for financialization or speculation include definitions and positions from the CFTC, set speculators, commodity index trader (CIT) positions, trading volume, and net positions of hedge funds and floor brokers.

Table 1.2: List of macroeconomic announcements in our sample

Announcement	Frequency	Source*	Unit	Time
GDP advance	Quarterly	BEA	%	8:30
GDP preliminary	Quarterly	BEA	%	8:30
GDP final	Quarterly	BEA	%	8:30
Personal income	Monthly	BEA	%	8:30
ADP employment	Monthly	ADP	Number of jobs	8:15
Initial jobless claims	Weekly	ETA	Number of claims	8:30
Non-farm employment	Monthly	BLS	Number of jobs	8:30
Factory orders	Monthly	BC	%	10:00
Industrial production	Monthly	FRB	%	9:15
Construction spending	Monthly	BC	%	10:00
Durable goods orders	Monthly	BC	%	8:30
Advance retail sales	Monthly	BC	%	8:30
Consumer credit	Monthly	FRB	USD	15:00
Personal consumption	Monthly	BEA	%	8:30
Building permits	Monthly	BC	Number of permits	8:30
Existing home sales	Monthly	NAR	Number of homes	10:00
Housing starts	Monthly	BC	Number of homes	8:30
New home sales	Monthly	BC	Number of homes	10:00
Pending home sales	Monthly	NAR	%	10:00
Trade balance	Monthly	BEA	USD	8:30
Consumer price index	Monthly	BLS	%	8:30
Producer price index	Monthly	BLS	%	8:30
CB Consumer confidence index	Monthly	CB	Index	10:00
UM Consumer sentiment	Monthly	TR/UM	Index	9:55
Weekly Crude Oil Stock	Weekly	EIA	number of barrels	10:30
Weekly Natural Gas Stock	Weekly	EIA	number of cubic feet	11:00

This table shows the category, frequency, source, unit of measure, and release time for each macroeconomic announcements. *(Automatic Data Processing, Inc. (ADP), Bureau of the Census (BC), Bureau of Economic Analysis (BEA), Bureau of Labor Statistics (BLS), Conference Board (CB), Employment and Training Administration (ETA), Federal Reserve Board (FRB), Institute for Supply Management (ISM), National Association of Realtors (NAR), Thomson Reuters/University of Michigan (TR/UM), and U.S. Department of the Treasury (USDT).)

Table 1.3: Descriptive statistics for the standardized surprise calculated for each of the macroeconomic announcements

Announcements	**Nb. obs.**	**Min.**	**1st Qu.**	**Med.**	**Mean**	**3rd Qu.**	**Max.**
Initial jobless claims	825	-3.407	-0.0720	-0.007	0.068	0.065	22.672
ADP Employment	202	-2.751	-0.0640	0.008	0.046	0.078	12.880
CB Consumer	201	-2.635	-0.4638	0.093	0.102	0.872	2.412
Advance retail sales	202	-4.028	-0.3661	-0.092	0.023	0.183	8.879
Building permit	198	-2.375	-0.5356	0.025	0.082	0.627	3.205
Construction spending	202	-3.054	-0.5912	-0.099	-0.130	0.493	4.335
Consumer_credit	202	-2.055	-0.5217	0.104	0.061	0.619	3.131
Consumer price index	201	-3.483	-0.6966	0.000	-0.035	0.697	4.180
Durable goods orders	193	-2.702	-0.5757	0.022	0.037	0.531	6.688
Existing home sales	202	-4.729	-0.4627	0.000	-0.067	0.488	2.467
Factory orders	202	-3.040	-0.3378	0.000	0.081	0.507	2.534
GDP	185	-2.589	-0.3698	0.000	-0.026	0.370	2.958
Housing starts	199	-2.285	-0.6178	0.000	0.042	0.624	3.401
Industrial production	385	-4.773	-0.5727	0.000	-0.080	0.573	2.291
Michigan Sentiment Index	202	-3.922	-0.4100	0.036	-0.049	0.463	3.244
New home sales	202	-3.062	-0.3466	0.116	0.118	0.631	3.562
Non-farm employment	201	-0.892	-0.0564	0.005	0.078	0.071	13.169
Pending home sales	202	-2.949	-0.4244	0.022	0.075	0.581	5.674
Personal consumption	201	-3.666	-0.3666	0.000	-0.026	0.367	2.566
Personal income	201	-1.079	-0.0771	0.000	0.095	0.077	13.108
Producer price index	188	-3.168	-0.5760	0.000	0.083	0.864	2.880
Trade balance	202	-1.831	-0.1801	-0.018	0.013	0.207	2.359

This table presents descriptive statistics for the standardized surprise $(A_{kt} - E_{kt})/\sigma_{kt}$ for each of the macroeconomic announcements. The column (Nb. Observations) shows the number of individual surprises that can be calculated over the whole analysis period. The columns (Min.), (1st Qu.), (Median), (Mean), (3rd Qu.) and (Max) present respectively the minimum value, the first quartile, the median, the mean, the third quartile and the maximum value for the standardized surprise of each macroeconomic announcement

Table 1.4: Descriptive statistics: 5-minute intraday futures price returns

Commodity Futures	Min (%)	1st Qu. (%)	Med. (%)	Mean (%)	3rd Qu. (%)	Max (%)
Crude Oil (CL=F)	-33.91	-0.0441	0.000	0.000	0.0447	41.64
Gold (GC=F)	-2.782	-0.0241	0.000	0.0001	0.0244	3.064
Copper (HG=F)	-4.534	-0.0363	0.000	-0.0001	0.0365	8.877
Natural Gas (NG=F)	-6.735	-0.0528	0.000	-0.0003	0.0532	15.62
Palladium (PA=F)	-13.350	-0.034	0.000	0.0001	0.0348	9.467
Silver (SI=F)	-7.504	-0.0394	0.000	0.0001	0.0415	4.242

Shows descriptive statistics of the 5-minute intraday returns, for each commodity futures. The columns (Min.), (1st Qu.), (Median), (Mean), (3rd Qu.) and (Max) present respectively the minimum value, the first quartile, the median, the mean, the third quartile and the maximum value for the 5 minute intraday returns.

Table 1.5: Descriptive statistics: Computed value of the NLS proxy (respectively: full sample, money managers-only sample, swap dealers-only sample)

	CL	**GC**	**HG**	**SI**	**PA**	**NG**
			NLS			
Min.	-0.1667	-0.4505	-0.3238	-0.1364	-0.6246	-0.2745
1st Qu.	0.0358	0.2285	-0.1067	0.1547	0.3181	-0.1731
Median	0.0642	0.2397	0.0235	0.2163	0.3056	-0.0563
Mean	0.0774	0.1780	0.0278	0.2123	0.2685	-0.0579
3rd Qu.	0.1834	0.4012	0.1417	0.3334	0.5652	-0.0321
Max.	0.2941	0.5269	0.4413	0.5748	0.7343	0.0794
			NLS_{MM}			
Min.	-0.0370	-0.2353	-0.2727	-0.2303	-0.6005	-0.2471
1st Qu.	0.0358	0.2177	0.0238	0.1550	0.3919	-0.0345
Median	0.1059	0.2177	0.0238	0.1550	0.3919	-0.0345
Mean	0.1011	0.2125	0.0350	0.1425	0.3294	-0.0371
3rd Qu.	0.1834	0.4012	0.1417	0.3334	0.5652	-0.0321
Max.	0.2051	0.4563	0.3923	0.4477	0.7330	0.1867
			NLS_{SD}			
Min.	-0.2645	-0.4092	-0.2263	-0.2263	-0.3640	-0.0689
1st Qu.	-0.1119	-0.1457	0.0238	-0.0255	-0.0251	0.1091
Median	-0.1119	-0.1457	0.2398	-0.0255	-0.0251	0.1091
Mean	-0.0897	-0.1612	0.2395	-0.0276	0.0086	0.1087
3rd Qu.	0.1847	0.1290	0.3923	0.3334	0.5652	0.2743
Max.	0.1847	0.1290	0.3923	0.3334	0.5652	0.2743

This table provides descriptive statistics of the NLS proxy for speculative trading intensity for each commodity futures contract in our sample. The lines (Min.), (1st Qu.), (Median), (Mean), (3rd Qu.) and (Max) present respectively the minimum value, the first quartile, the median, the mean, the third quartile, and the maximum value. CL: crude oil, GC: gold, HG: high-grade copper, SI: silver, PA: palladium, NG: natural gas.

1.11 Figures

Table 1.6: Effects of Macro Announcements and Speculative Trading Intensity (NLS) on Futures Returns - Full Sample

Commodities	Crude Oil		Gold		Copper		Silver		Palladium		Natural Gas	
Announcements	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m
						Macroeconomic News Announcements						
Initial jobless claims	−0.198***	0.879***	0.700***	−1.340***	−0.449***	1.988***	−0.007	−0.015	0.029**	−0.221***	−0.003	−0.109
ADP Employment	0.343***	−1.068**	−1.187***	2.525***	0.581***	−2.370***	−0.026***	0.169	0.008	−0.142	−0.115***	−2.798***
CB Consumer	0.117***	−0.469***	−0.059***	0.070*	0.080***	−0.152**	0.000	0.006	−0.068***	0.116**	0.029	−0.014
Advance retail sales	0.200***	−0.745***	−0.229***	0.447***	0.104***	−0.425**	0.013	−0.086	−0.018	0.007	−0.027	−0.471
Building permit	0.000	0.040	−0.027**	0.037	0.063***	−0.157**	0.007	−0.004	−0.002	−0.037	0.043*	0.362**
Construction spending	0.022	−0.146	−0.041**	0.092**	0.040*	−0.144*	−0.001	−0.189***	0.010	−0.003	−0.029	−0.179
Consumer credit	−0.016	0.083	−0.014*	0.030	−0.016*	0.062	0.002	0.024	0.011	−0.010	−0.012	−0.112
Consumer price index	0.084***	−0.293**	−0.172***	0.310***	0.269***	−0.759***	−0.030***	−0.023	−0.142***	0.231***	0.013	0.150
Durable goods orders	0.153***	−0.738***	−0.073***	0.147***	0.064***	−0.149*	−0.011*	0.000	−0.054***	0.145***	−0.003	−0.100
Existing home sales	0.087***	−0.592***	−0.020	0.050	−0.022	0.108	−0.022***	−0.069	0.012	−0.045	0.041	0.219
Factory orders	−0.001	0.010	−0.029	0.003	−0.060**	0.146	0.004	0.025	−0.019	0.000	0.107***	0.685***
Gross domestic product	0.046*	−0.184	−0.161***	0.242***	0.170***	−0.331***	0.009	−0.005	−0.054***	0.008	−0.036	−0.337*
Housing starts	0.032	−0.127	−0.063***	0.114***	0.071***	−0.149**	0.005	0.006	0.002	−0.036*	0.023	0.383**
Industrial production	0.021	−0.154	−0.001	−0.071	−0.013	−0.116	−0.009	−0.008	−0.022	0.021	−0.023	−0.023
New home sales	0.099***	−0.436**	−0.069***	0.136***	−0.033*	−0.041	−0.024***	0.087*	−0.008	0.022	−0.003	−0.084
Non-farm employment	1.404***	−5.111***	−3.194***	6.812***	1.389***	−6.044***	−0.027***	0.624**	0.000	0.153	−0.218***	−6.294***
Pending home sales	0.073***	−0.349**	−0.018	0.003	−0.013	−0.023	−0.017**	−0.111	−0.009	0.031	−0.024	−0.399
Personal consumption	−0.015	0.068	−0.036**	0.063*	0.009	−0.058	0.003	0.087	0.013	−0.050	0.004	0.369
Personal income	0.009	−0.136	−0.072	0.150	0.254***	−1.052**	−0.017*	0.279**	−0.013	−0.022	−0.028	−0.820
Producer price index	0.035*	−0.236**	−0.077***	0.150***	−0.015	−0.042	−0.005	−0.013	−0.020	0.045	−0.049*	−0.223
Trade balance	0.003	−0.056	−0.052**	0.119**	−0.011	0.041	−0.004	−0.132	−0.039	0.077	−0.047	−0.451
Crude Oil Weekly inventory	−0.077***	−0.396***										
Natural Gas Weekly inventory											−0.365***	−0.077
Observations	1,193,455		1,190,001		1,180,816		1,138,696		749,168		1,101,836	
$R^2(\%)$	0.198		0.204		0.101		0.0290		0.0246		0.118	

This table presents estimates of eq. 1.4, $R_t^{t+\tau} = \alpha + \sum_{m=1}^{22} \gamma_m S_{m,t} + \delta X_{t,i} + \sum_{m=1}^{22} \theta_m (S_{m,t} \cdot X_t) + \beta R_{t-\tau}^t + \epsilon_t$ using the method proposed by Kurov et al. (2019) and speculative trading intensity variable $X_t = NLS_t$. The period covered is from 2007-04-01 to 2024-02-11. The γ_m coefficients capture the instantaneous change in the futures return when an announcement has just occurred. The θ_m coefficients capture the instantaneous change in return when an announcement has just occurred in conjunction with the level of speculative trading intensity.

Table 1.7: Effects of Macro Announcements and Speculative Trading Intensity (NLS) on Futures Conditional Variance - Full Sample

Commodities	**Crude Oil**		**Gold**		**Copper**		**Silver**		**Palladium**		**Natural Gas**	
Announcements	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m
					Macroeconomic News Announcements							
Initial jobless claims	0.101***	−0.436***	0.047***	0.020	0.117***	−0.188***	0.025***	−0.056***	0.118***	−0.114***	−0.012	−0.021
ADP Employment	0.028	−0.200	0.064***	−0.084***	0.053***	−0.030	0.028***	0.034	−0.029*	0.015	0.016	0.057
CB Consumer	0.104***	−0.414***	0.031***	0.012	0.045***	0.028	0.028***	−0.103**	0.061***	−0.099***	−0.008	−0.013
Advance retail sales	0.169***	−0.862***	0.096***	−0.056*	0.124***	−0.054	0.028***	−0.156***	0.114***	−0.169***	0.058***	0.251*
Building permit	0.105***	−0.540***	0.046***	−0.060*	0.061***	−0.193***	0.005	−0.074*	0.077***	−0.070***	0.022	−0.026
Construction spending	0.153***	−0.627***	0.095***	−0.119***	0.086***	−0.098	0.041***	−0.190***	0.046***	0.013	−0.019	−0.321**
Consumer credit	0.048*	−0.261*	0.008	−0.003	0.017	0.004	0.003	0.039	−0.006	0.034	0.002	0.014
Consumer price index	0.072**	−0.067	0.139***	−0.075*	0.266***	−0.529***	0.094***	−0.375***	0.182***	−0.274***	0.027	0.039
Durable goods orders	0.098***	−0.411***	0.023**	0.016	0.042***	0.022	0.008	−0.024	0.068***	−0.037	−0.022	−0.256*
Existing home sales	0.048*	−0.059	0.019*	0.041	0.038***	−0.030	0.022***	−0.132***	0.048***	−0.067**	0.022	−0.113
Factory orders	0.088***	−0.311**	−0.007	0.162***	0.018	0.151**	0.026***	−0.132***	0.040***	−0.040	−0.030	−0.695***
Gross domestic product	0.116***	−0.618***	0.047***	0.053	0.100***	−0.063	0.025***	0.017	0.110***	−0.121***	−0.051**	−0.309**
Housing starts	0.113***	−0.607***	0.037***	−0.031	0.062***	−0.190***	0.008	−0.095**	0.060***	−0.049	0.018	0.040
Industrial production	0.096***	−0.638***	0.019*	0.003	0.021	−0.007	0.002	−0.048	0.020	−0.071**	−0.033	−0.187
New home sales	0.101***	−0.527***	0.041***	0.015	0.058***	0.028	0.025***	−0.001	0.068***	−0.115***	−0.025	−0.229*
Non-farm employment	0.365***	−1.206***	0.236***	0.047	0.381***	−0.141**	0.129***	−0.156***	0.173***	−0.050	0.028	−0.354**
Pending home sales	0.109***	−0.494***	0.020*	−0.016	0.025*	0.021	0.028***	−0.008	0.059***	−0.085***	−0.014	−0.308**
Personal consumption	0.006	0.021	0.062***	−0.081**	0.102***	−0.268***	0.007	−0.008	0.111***	−0.148***	0.024	−0.376**
Personal income	0.007	0.047	0.091***	−0.150***	0.125***	−0.341***	0.013**	−0.023	0.094***	−0.109***	0.003	−0.432***
Producer price index	0.097***	−0.350**	0.017	0.085**	0.090***	−0.099	0.011*	−0.051	0.120***	−0.067**	0.008	0.005
Trade balance	0.078***	−0.128	−0.007	0.206***	0.003	0.393***	0.009	0.030	0.103***	−0.087***	0.018	0.112
Crude Oil Weekly inventory	0.005	0.624***										
Natural Gas Weekly inventory											0.531***	−0.687***
Observations	1,193,455		1,190,001		1,180,816		1,138,696		749,168		1,101,836	
R^2(%)	7.065		8.064		7.396		7.057		7.629		13.60	

This table presents estimates of eq. 1.7 using the speculative trading intensity variable NLS_t. The equation is $\sigma_t^2 = \alpha_0 + \alpha_1 \sigma_{t-1}^2 + \alpha_2 \epsilon_t^2 + \sum_{m=1}^{22} \Phi_m D_{m,t} + \beta X_{i,t} + \sum_{k=1}^{n} \phi_k I_{kt} + \sum_{h=1}^{23} \rho_h D_h$ where $I_{kt} = D_{m,t} \cdot X_{i,t}$ and $D_{m,t}$ is a macro announcement dummy variable for release m. The period covered is from 2007-04-01 to 2024-02-11. The Φ_m coefficients capture the instantaneous change in the conditional variance when an announcement has just occurred. The ϕ_m coefficients capture the conditional variance when an announcement has just occurred in conjunction with the level of speculative trading intensity.

Table 1.8: Effects of Macro Announcements and Speculative Trading Intensity (NLS) on Futures Price Bid-Ask Spreads - Full Sample

Commodities	**Crude Oil**		**Gold**		**Silver**		**Natural Gas**		**Copper**		**Palladium**	
Announcements	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m
Macroeconomic News Announcements												
Initial jobless claims	−2.710***	13.698	5.425	−10.380 ***	5.850	−26.397 ***	−0.249	−4.543***	0.015	0.857	−0.191	0.537
ADP Employment	3.629	−10.942 ***	−10.542 ***	22.354	−7.399***	31.336	−0.652	−13.889 ***	0.250	2.043	0.036	−0.467
CB Consumer	1.419	−6.221***	−0.850	1.447	−1.132	3.048	0.512	0.302	0.026	0.242	−0.775	1.928
Advance retail sales	2.089	−7.055***	−2.279	4.431	−0.850	2.848	−0.227	−2.023	0.351	−3.220***	−0.419	1.513
Building permit	−0.197	1.782	−0.437	0.651	−0.774	2.561	0.301	3.765	0.030	0.368	−0.199	−0.203
Construction spending	0.177	−1.035	−0.064	0.076	−0.265	0.847	−0.187	−2.696***	0.061	−2.420***	−0.126	0.298
Consumer credit	0.371	−1.899	−0.027	0.046	−0.029	0.136	−0.041	0.047	0.055	0.359	−0.007	0.021
Consumer price index	−0.829	1.332	−2.239	4.869	−3.053	9.401	−0.042	1.982	−0.562	1.096	−1.828	3.268
Durable goods orders	1.318	−7.368***	−0.956	2.313	−1.310	5.553	−0.121	−1.497	0.203	−0.184	−0.388	1.251
Existing home sales	1.054	−5.366***	−0.398	1.067	−0.161	1.420	1.064	5.295	0.422	−1.812	0.016	−0.041
Factory orders	−0.338	1.588	−0.851	1.343	−0.769	1.833	1.549	11.321	0.014	1.088	−0.033	−0.325
GDP	0.358	−1.600	−1.891	3.738	−2.248	6.383	−0.394	−3.634***	0.121	−0.732	−0.565	0.374
Housing starts	0.190	−1.233	−0.990	1.896	−0.970	2.572	0.541	6.538	0.083	0.511	0.036	−0.563
Industrial production	0.120	−2.099	−0.195	−0.414	−0.061	−1.994	−0.203	−1.192	−0.058	−0.402	−0.181	0.282
Manufacturing capacity	−0.103	2.279	−0.612	0.922	−0.822	1.975	−0.366	−2.898***	−0.037	0.697	−0.141	0.549
New home sales	1.631	−8.343***	−0.906	1.652	−0.385	−0.407	−0.252	−0.885	0.369	−0.456	−0.010	0.057
Non-farm payroll	10.788	−38.888 ***	−29.711 ***	63.342	−10.286 ***	43.957	−2.491	−71.471 ***	0.272	8.522	0.000	1.467
Pending home sales	0.825	−2.545	−0.422	0.524	−0.194	−0.157	−0.575	−9.013***	0.157	−2.360***	−0.354	0.282
Personal consumption	−0.215	−0.039	−0.826	1.629	−0.034	0.119	−0.117	3.932	0.092	2.044	−0.252	0.577
Personal income	−0.088	−0.820	−1.535	3.290	−4.400***	17.615	−0.461	−12.861 ***	−0.246	−3.838***	−0.162	0.235
Producer price index	0.714	−4.424***	−0.723	1.434	0.220	−1.894	−0.514	−2.803***	0.061	−0.121	−0.723	1.480
Treasury balance	0.046	−1.620	−0.821	1.977	−0.803	2.967	−0.126	−0.738	−0.033	−1.503	−0.314	0.842
Announcements specific to commodity markets												
Weekly crude oil stock	−0.977	−3.473***										
Natural Gas Weekly inventory							−1.747	−7.027***				
Observations	1,193,455		1,190,001		1,138,696		1,101,836		1,180,816		749,168	
$R^2(\%)$	1.98		1.77		0.56		1.85		1.28		1.51	

This table presents estimates of the equation $R^{t+\tau}_{SPREAD_t} = \alpha + \sum_{m=1}^{22} \gamma_m D_{m,t} + \delta X_{t,i} + \sum_{m=1}^{22} \theta_m (D_{m,t} \cdot X_t) + \beta R^t_{SPREAD_{t-\tau}} + \epsilon_t$, analyzing the effects of speculative trading intensity and macroeconomic announcements on the bid-ask spread using the speculative trading intensity variable NLS_t. The period covered is from 2007-04-01 to 2024-02-11. The γ_m coefficients capture the instantaneous change in the bid-ask spread when a macroeconomic announcement occurs. The δ coefficient represents the effect of the speculative trading intensity variable NLS_t. The θ_m coefficients capture the interaction effect between macroeconomic announcements and speculative trading intensity.

Table 1.9: Effects of Macro Announcements and Speculative Trading Intensity (Money Managers only) on Futures Returns

Commodities	Crude Oil		Gold		Copper		Silver		Palladium		Natural Gas	
Announcements	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m
Macroeconomic News Announcements												
Initial jobless claims	−0.030	0.020	0.301***	−0.962***	−0.026	0.334	−0.009	−0.030	0.046**	−0.275***	0.010	0.061
ADP Employment	0.292**	−1.237*	−0.068	0.154	0.223***	−0.974**	−0.023***	0.150	0.010	−0.118	−0.022	−0.492
CB Consumer	0.115***	−0.654**	−0.048***	0.056	0.070***	−0.140**	0.000	0.002	−0.067***	0.115**	0.037**	0.122
Advance retail sales	0.180***	−0.951***	0.006	−0.228***	0.095***	−0.523***	0.013	−0.069	−0.021	0.021	−0.005	−0.244
Building permit	−0.031	0.308*	−0.014**	−0.002	0.049***	−0.122**	0.007	0.001	−0.003	−0.038	0.003	0.087
Construction spending	0.047	−0.415	−0.048***	0.171***	0.034**	−0.152*	0.005	−0.126**	0.010	−0.003	−0.011	−0.093
Consumer credit	−0.009	0.057	−0.010**	0.041*	−0.010	0.053	0.001	0.020	0.011	−0.013	−0.002	−0.048
Consumer price index	0.086***	−0.420**	−0.158***	0.472***	0.217***	−0.727***	−0.029***	−0.015	−0.136***	0.223***	−0.001	0.078
Durable goods orders	0.093***	−0.696***	−0.058***	0.157***	0.053***	−0.119	−0.015**	−0.051	−0.062***	0.177***	0.010	−0.020
Existing home sales	0.095***	−0.802***	−0.007	0.012	−0.012	0.079	−0.021**	−0.013	0.013	−0.050	0.013	0.053
Factory orders	−0.024	0.219	−0.044***	0.069	0.050**	−0.145	0.003	0.012	−0.019	0.000	0.042**	0.387**
Gross domestic product	0.044*	−0.252	−0.083***	0.023	0.128***	−0.187**	0.006	0.027	−0.060***	0.024	−0.004	−0.282*
Housing starts	−0.017	0.198	−0.028***	0.008	0.046***	−0.037	0.003	0.024	0.002	−0.041*	−0.013	0.142
Industrial production	−0.003	−0.035	−0.023**	−0.020	−0.022	−0.101	−0.007	−0.034	−0.032	0.045	−0.013	0.233
New home sales	0.040	−0.201	−0.035***	0.053	0.032**	0.061	−0.024***	−0.045	−0.012	0.032	0.004	−0.183
Non-farm employment	0.798***	−4.073***	−0.517***	2.145***	0.771***	−4.430***	−0.015*	0.613***	0.025	−0.162	−0.005	−0.956
Pending home sales	0.056	−0.312	−0.019*	0.007	−0.021	0.016	−0.016*	−0.034	−0.015	0.050	0.010	−0.270
Personal consumption	0.091**	−0.641***	−0.019**	0.043	0.005	−0.064	0.000	0.081	0.012	−0.049	−0.025	0.232
Personal income	0.058	−0.501	0.017	−0.097	0.145**	−0.834**	−0.008	−0.200	−0.017	0.009	0.004	−1.050*
Producer price index	0.039*	−0.326*	−0.055***	0.133***	−0.006	−0.122	−0.003	−0.031	−0.022	0.055	−0.026	−0.112
Trade balance	0.007	−0.108	−0.031**	0.081*	−0.009	0.040	0.004	−0.118	−0.025	0.060	0.000	−0.269
Crude Oil Weekly inventory	−0.072***	−0.610***										
Natural Gas Weekly inventory											−0.355***	−0.122
Observations	1,193,455		1,190,001		1,180,816		1,138,696		749,168		1,101,836	
R^2(%)	0.181		0.151		0.0923		0.0284		0.0235		0.113	

This table presents estimates of eq. 1.4, $R_t^{t+\tau} = \alpha + \sum_{m=1}^{22} \gamma_m S_{m,t} + \delta X_{t,i} + \sum_{m=1}^{22} \theta_m (S_{m,t} \cdot X_t) + \beta R_{t-\tau}^{t} + \epsilon_t$ using the method proposed by Kurov et al. (2019) and the money manager-only speculative trading intensity variable $NLS_{t,MM}$, calculated with the money manager positions. The period covered is from 2007-04-01 to 2024-02-11. The γ_m coefficients capture the instantaneous change in return when an announcement has just occurred. The coefficients θ_m capture the instantaneous change in return when an announcement has just occurred in conjunction with the level of speculative trading intensity.

Table 1.10: Effects of Macro Announcements and Speculative Trading Intensity (Swap Dealers only) on Futures Returns

Commodities	Crude Oil		Gold		Copper		Silver		Palladium		Natural Gas	
Announcements	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m
Macroeconomic News Announcements												
Initial jobless claims	−0.186***	−0.763***	0.361***	1.023***	0.014	−0.511	0.082**	0.448**	0.006	−0.035	−0.007	0.236
ADP Employment	0.426***	1.770***	−0.732***	−1.923***	0.143***	1.908***	0.105	−0.417	−0.084*	−0.379*	−0.070***	1.981***
CB Consumer	0.086***	0.387***	−0.027***	0.061	0.047***	0.045	−0.057***	−0.252***	−0.024**	−0.137**	0.029	0.015
Advance retail sales	0.188***	0.837***	−0.221***	−0.622***	0.023*	0.330	−0.052**	−0.344**	−0.024	0.037	−0.012	0.183
Building permit	0.012	0.034	−0.028***	−0.064**	0.039***	0.196*	0.009	−0.008	−0.015**	0.062	0.067**	−0.571***
Construction spending	0.009	0.104	−0.015*	−0.040	0.016	0.154	−0.046**	−0.199**	0.013	0.067	−0.006	−0.035
Consumer credit	−0.001	0.007	−0.003	0.002	−0.004	0.005	0.007	−0.019	0.007	0.027	−0.020	0.185
Consumer price index	−0.051***	−0.123	−0.042***	−0.116**	0.109***	0.501***	−0.080***	−0.206**	−0.054***	−0.284***	0.042	−0.404
Durable goods orders	0.082***	0.408***	−0.028***	−0.028	0.023**	0.278***	−0.001	0.047	0.009	−0.090**	0.011	−0.012
Existing home sales	0.025**	0.285***	−0.007	−0.021	−0.001	0.075	−0.060***	−0.336***	−0.006	0.051	0.054	−0.336
Factory orders	−0.003	−0.038	−0.004	0.145**	0.032***	−0.194	0.015	−0.047	−0.020	−0.032	0.096***	−0.622**
Gross domestic product	0.038**	0.177**	−0.091***	−0.062	0.100***	−0.154	−0.025	0.146**	−0.052***	−0.009	−0.032	0.302
Housing starts	0.030**	0.134*	−0.055***	−0.149***	0.051***	0.218**	−0.018	0.120*	−0.010	0.023	0.019	−0.321
Industrial production	0.024	0.199*	−0.004	0.107**	0.034***	0.147	−0.030	0.098	−0.013	−0.013	−0.029	0.090
New home sales	0.072***	0.343***	−0.040***	−0.082*	0.038***	0.090	−0.038**	−0.288***	0.000	−0.026	0.016	−0.110
Non-farm employment	1.820***	8.772***	−3.079***	−8.258***	0.226***	3.881***	−0.485***	−2.661***	−0.146*	−0.684*	−0.067**	2.423**
Pending home sales	0.042***	0.275***	−0.017	0.002	−0.017	0.075	−0.033	0.199**	0.004	−0.025	−0.004	0.252
Personal consumption	−0.002	−0.005	−0.017*	−0.027	0.005	0.321**	−0.008	0.031	−0.006	0.071	0.003	−0.383
Personal income	−0.022	0.006	−0.058*	−0.154	0.082***	1.829***	−0.091	0.437	−0.023	0.039	−0.008	0.530
Producer price index	0.008	0.075	−0.024***	−0.002	0.025***	0.089	−0.032	0.111	−0.006	−0.036	−0.036	0.151
Trade balance	−0.001	0.038	−0.016	−0.062	−0.012	−0.239	−0.045	0.144	−0.014	−0.220*	0.000	0.125
Crude Oil Weekly inventory	−0.144***	0.036										
Natural Gas Weekly inventory											−0.473***	0.908***
Observations	1,193,455		1,190,001		1,180,816		1,138,696		749,168		1,101,836	
$R^2(\%)$	0.196		0.206		0.0809		0.0337		0.0195		0.123	

This table presents estimates of eq. 1.4, $R_t^{t+\tau} = \alpha + \sum_{m=1}^{22} \gamma_m S_{m,t} + \delta X_{t,i} + \sum_{m=1}^{22} \theta_m (S_{m,t} \cdot X_t) + \beta R_{t-\tau}^t + \epsilon_t$ using the method proposed by Kurov et al. (2019) and the swap dealer-specific speculative trading intensity variable $NLS_{t,SD}$, calculated with the swap dealer positions. The period covered is from 2007-04-01 to 2024-02-11. The γ_m coefficients capture the instantaneous change in return when an announcement has just occurred. The coefficients θ_m capture the instantaneous change in return when an announcement has just occurred in conjunction with the level of speculative trading intensity.

Table 1.11: Effects of Macro Announcements and Speculative Trading Intensity (Money Managers only) on Futures Conditional Variance

Commodities	Crude Oil		Gold		Copper		Silver		Palladium		Natural Gas	
Announcements	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m
	Macroeconomic News Announcements											
Initial jobless claims	0.086***	−0.470***	0.052***	0.011	0.104***	−0.179***	0.027***	−0.048***	0.126***	−0.140***	−0.011**	−0.048
ADP Employment	0.001	−0.032	0.040***	−0.020	0.051***	−0.025	0.025***	−0.063*	−0.028*	0.016	0.012	0.057
CB Consumer	0.112***	−0.666***	0.040***	−0.020	0.048***	0.019	0.031***	−0.047	0.062***	−0.104***	−0.008	−0.043
Advance retail sales	0.142***	−0.963***	0.081***	−0.014	0.125***	−0.085	0.033***	−0.112***	0.125***	−0.210***	0.039***	0.275**
Building permit	0.103***	−0.757***	0.021***	0.028	0.046***	−0.166***	0.006	−0.038	0.085***	−0.100***	0.028**	0.077
Construction spending	0.142***	−0.784***	0.066***	−0.050*	0.072***	−0.041	0.049***	−0.174***	0.056***	−0.013	0.007	−0.303***
Consumer credit	0.039	−0.287	0.004	0.015	0.010	0.054	0.001	0.036	−0.008	0.039	0.000	0.009
Consumer price index	0.059*	0.039	0.159***	−0.210***	0.234***	−0.520***	0.104***	−0.218***	0.182***	−0.282***	0.023	0.015
Durable goods orders	0.057**	−0.210	0.024***	0.021	0.049***	−0.025	0.009	−0.032	0.077***	−0.066*	0.002	−0.158
Existing home sales	0.115***	−0.691***	0.024***	0.040	0.041***	−0.058	0.027***	−0.115***	0.051***	−0.077**	0.032***	−0.084
Factory orders	0.162***	−1.114***	0.032***	−0.070**	0.029***	−0.138**	0.032***	−0.132***	0.042***	−0.046	0.033***	−0.460***
Gross domestic product	0.081***	−0.538**	0.055***	0.044	0.091***	−0.029	0.025***	0.008	0.122***	−0.159***	−0.022*	−0.240**
Housing starts	0.112***	−0.857***	0.021***	0.028	0.049***	−0.175***	0.011*	−0.057	0.067***	−0.074**	0.017	0.095
Industrial production	0.120***	−1.146***	0.014**	0.031	0.032***	−0.096	0.004	−0.067*	0.022	−0.080**	−0.019	−0.212*
New home sales	0.078***	−0.534**	0.049***	−0.017	0.068***	−0.033	0.025***	−0.004	0.069***	−0.121***	−0.004	−0.158
Non-farm employment	0.302***	−1.113***	0.199***	−0.243***	0.348***	0.038	0.136***	−0.173***	0.174***	−0.053	0.057***	−0.324***
Pending home sales	0.124***	−0.827***	0.022***	−0.033	0.026**	0.025	0.028***	0.002	0.061***	−0.092***	0.010	−0.336***
Personal consumption	0.032	−0.201	0.066***	−0.142***	0.087***	−0.274***	0.008	−0.030	0.120***	−0.178***	0.063***	−0.192
Personal income	0.035	−0.188	0.066***	−0.115***	0.100***	−0.306***	0.016**	−0.058	0.103***	−0.133***	0.045***	−0.219*
Producer price index	0.123***	−0.729***	0.036***	0.044	0.089***	−0.130**	0.013**	−0.039	0.127***	−0.093**	0.009	0.024
Trade balance	0.110***	−0.467*	0.017**	−0.209***	0.037***	−0.315***	0.009	0.007	0.106***	−0.096***	0.007	0.037
Crude Oil Weekly inventory	0.053*	0.553**										
Natural Gas Weekly inventory											0.598***	−0.380***
Observations	1,193,455		1,190,001		1,180,816		1,138,696		749,168		1,101,836	
R^2	7.869		8.237		7.398		7.119		7.50		13.643	

This table presents estimates of eq. 1.7 using the speculative trading intensity variable $NLS_{t,MM}$, calculated with the money manager positions. The equation is $\sigma_t^2 = \alpha_0 + \alpha_1\sigma_{t-1}^2 + \alpha_2\epsilon_t^2 + \sum_{m=1}^{22} \Phi_m D_{m,t} + \beta X_{i,t} + \sum_{k=1}^{n} \phi_k I_{kt} + \sum_{h=1}^{23} \rho_h D_h$ where $I_{kt} = D_{m,t} \cdot X_{i,t}$ and $D_{m,t}$ is a macro announcement dummy variable for release m. The period covered is from 2007-04-01 to 2024-02-11. The Φ_m coefficients capture the instantaneous change in the conditional variance when an announcement has just occurred. The ϕ_m coefficients capture the conditional variance when an announcement has just occurred in conjunction with the level of speculative trading intensity.

Table 1.12: Effects of Macro Announcements and Speculative Trading Intensity (Swap Dealers only) on Futures Conditional Variance

Commodities	**Crude Oil**		**Gold**		**Copper**		**Silver**		**Palladium**		**Natural Gas**	
Announcements	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m	Φ_m	ϕ_m
					Macroeconomic News Announcements							
Initial jobless claims	0.058***	0.233***	0.038***	0.094***	0.083***	0.264***	0.016**	0.171***	0.078***	0.212***	−0.027**	0.148*
ADP Employment	0.010	0.128	0.035***	−0.008	0.045***	−0.116	−0.004	0.132**	−0.023**	−0.062	0.019	−0.085
CB Consumer	0.065***	0.242***	0.025***	0.065**	0.052***	0.094	0.060***	0.373***	0.025**	0.163***	0.003	−0.077
Advance retail sales	0.096***	0.545***	0.075***	−0.014	0.112***	−0.050	0.070***	0.413***	0.055***	0.281***	0.048**	−0.175
Building permit	0.049***	0.250***	0.034***	0.044	0.028***	0.236***	0.068***	0.320***	0.053***	0.135***	−0.020	0.375**
Construction spending	0.086***	0.289***	0.057***	0.008	0.069***	0.200**	0.077***	0.495***	0.051***	0.030	0.009	0.084
Consumer credit	0.024*	0.154*	0.010	0.017	0.016**	−0.066	0.010	−0.033	0.004	−0.084	−0.001	0.008
Consumer price index	0.039**	0.235**	0.081***	0.205***	0.166***	0.293***	0.054**	0.177**	0.088***	0.449***	−0.022	0.360*
Durable goods orders	0.064***	0.282***	0.024***	−0.027	0.048***	0.199**	0.034**	0.179***	0.054***	0.110**	−0.043**	0.430***
Existing home sales	0.032**	−0.057	0.031***	−0.010	0.035***	0.133	−0.027	0.212***	0.024**	0.168***	0.033	0.021
Factory orders	0.060***	0.192**	0.015**	0.192***	0.045***	0.233***	−0.026	0.217***	0.025**	0.088	−0.028	0.672***
Gross domestic product	0.056***	0.345***	0.022***	0.248***	0.090***	0.167*	−0.010	0.137*	0.067***	0.245***	−0.059**	0.362**
Housing starts	0.049***	0.289***	0.027***	0.003	0.028***	0.192**	0.065***	0.317***	0.043***	0.096*	−0.015	0.244
Industrial production	0.031**	0.319***	0.033***	0.077**	0.019***	0.025	−0.011	0.058	−0.004	0.133**	−0.027	0.141
New home sales	0.048***	0.279***	0.039***	−0.041	0.066***	0.186**	−0.021	0.193***	0.028***	0.207***	−0.025	0.231
Non-farm employment	0.268***	0.842***	0.242***	−0.058*	0.361***	0.481***	0.018	0.475***	0.155***	0.187***	0.062***	0.066
Pending home sales	0.060***	0.264***	0.015**	0.004	0.030***	0.036	0.037**	0.269***	0.029***	0.150***	0.014	0.076
Personal consumption	0.005	−0.044	0.031***	−0.019	0.051***	0.262***	0.015	−0.040	0.057***	0.316***	−0.045*	0.919***
Personal income	0.008	−0.053	0.062***	0.124***	0.062***	0.455***	0.018	−0.023	0.054***	0.233***	−0.065***	0.990***
Producer price index	0.062***	0.167*	0.038***	−0.046	0.071***	0.052	−0.029	0.169**	0.095***	0.187***	0.020	−0.108
Trade balance	0.075***	0.150*	0.035***	0.155***	0.067***	0.760***	0.014	−0.022	0.073***	0.123**	0.017	−0.104
Crude Oil Weekly inventory	0.116***	0.005										
Natural Gas Weekly inventory											0.484***	1.025***
Observations	1,193,455		1,190,001		1,180,816		1,138,696		749,168		1,101,836	
$R^2(\%)$	7.01		7.716		8.023		7.975		7.732		13.66	

This table presents estimates of eq. 1.7 using the speculative trading intensity variable $NLS_{t,SD}$, calculated with the swap dealer positions. The equation is $\sigma_t^2 = \alpha_0 + \alpha_1\sigma_{t-1}^2 + \alpha_2\epsilon_t^2 + \sum_{m=1}^{22}\Phi_m D_{m,t} + \beta X_{i,t} + \sum_{k=1}^{n}\phi_k I_{kt} + \sum_{h=1}^{23}\rho_h D_h$ where $I_{kt} = D_{m,t}\cdot X_{i,t}$ and $D_{m,t}$ is a macro announcement dummy variable for release m. The period covered is from 2007-04-01 to 2024-02-11. The Φ_m coefficients capture the instantaneous change in the conditional variance when an announcement has just occurred. The ϕ_m coefficients capture the conditional variance when an announcement has just occurred in conjunction with the level of speculative trading intensity.

Table 1.13: Effects of Macro Announcements and Speculative Trading Intensity (Money Managers only) on Futures Price Bid-Ask Spreads

Commodities	Crude Oil		Gold		Copper		Silver		Palladium		Natural Gas	
Announcements	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m
Macroeconomic News Announcements												
Initial jobless claims	−2.466**	9.751	0.222	−6.107	−0.414	−2.993	−1.520**	−5.809	−7.738*	10.173	−2.069**	2.892
ADP Employment	−1.105	4.874	−0.031	0.313	−0.446	−0.606	−0.464	0.041	11.325	−20.087	−1.567	9.513
CB Consumer	−0.883	−7.590	−1.073**	−21.883 ***	−0.609	−6.322	−2.270*	−5.006	−4.894	6.322	−6.110***	29.787
Advance retail sales	−1.753	6.539	−1.848	4.412	−1.507*	−4.140	−1.112	−5.342	−8.473	16.815	−2.521	−23.726
Building permit	0.048	−5.084	0.123	−2.625	−0.438	−3.660	−0.859	−0.392	−16.807 **	−9.470	−2.481	−3.851
Construction spending	0.715	−21.505	0.411	−19.636 **	−2.036**	−6.822	−0.782	1.915	−3.651	14.688	−6.182***	14.393
Consumer credit	−0.391	−2.145	−0.232	0.923	−0.634	−0.728	0.723	4.263	2.894	0.826	−2.085	−10.933
Consumer price index	−1.794	13.111	−0.328	−2.431	−0.594	−0.574	−1.526	−7.462	−28.010 **	33.089	−2.166	17.815
Durable goods orders	−2.201	5.989	0.237	−7.012	−1.029	−5.376	−2.466*	−19.430 **	−19.847 **	−52.267 ***	−3.505	22.773
Existing home sales	−0.494	−6.379	0.152	−2.614	−0.832	−5.410	−1.458	−2.821	−5.480	6.158	−5.716***	23.659
Factory orders	−2.690	15.967	−0.517	−6.286	−1.621*	−9.048	−1.962	−1.641	−0.552	28.548	−4.095**	17.201
Gross domestic product	−3.813	16.725	−1.274	−5.901	−1.184	−6.391	−1.558	−17.412 *	−10.428	13.464	−2.793	6.617
Housing starts	−0.268	−3.673	0.022	−2.479	−0.326	−3.621	−0.123	−1.595	−14.939 *	−11.913	−2.480	−3.106
Industrial production	0.999	−10.142	−0.040	8.877	−0.268	−0.374	0.521	0.478	7.052	−13.194	−3.891*	10.562
New home sales	0.517	−13.676	−0.320	−4.684	−0.542	−2.955	−0.874	−12.355	−8.912	6.444	−4.416**	11.799
Non-farm employment	−1.439	1.498	−0.839	−7.010	−3.093***	−24.194 ***	−0.262	−5.058	−19.800 **	22.461	−2.069	19.687
Pending home sales	−1.224	4.329	0.233	−3.304	−0.660	−2.140	−2.381*	−5.643	4.603	−14.701	−5.088***	10.999
Personal consumption	−1.580	8.274	−0.382	2.582	0.286	−1.307	−0.717	−10.634	−14.710 *	−0.853	−2.604	16.070
Personal income	−1.833	4.187	0.035	−0.933	−0.548	−3.162	−1.787	−3.128	−18.449 **	−7.560	−3.686	15.017
Producer price index	1.232	−25.459	0.173	−5.424	−1.585	−8.726	0.463	−4.845	−3.853	−6.534	−0.312	7.498
Trade balance	0.780	−13.301	−0.004	0.057	0.137	−10.608	−0.445	−5.855	−11.442	39.185	−0.224	5.237
Crude Oil Weekly inventory	0.494	−12.743										
Natural Gas Weekly inventory											−7.596***	−19.173 *
Observations	1,041,497		1,022,592		1,023,897		1,022,006		683,875		995,919	
$R^2(\%)$	0.0572		0.115		0.12		0.118		0.034		0.0942	

This table presents estimates of the equation $R_{\text{SPREAD}}{}_t^{t+\tau} = \alpha + \sum_{m=1}^{22} \gamma_m D_{m,t} + \delta X_{t,i} + \sum_{m=1}^{22} \theta_m (D_{m,t} \cdot X_t) + \beta R_{\text{SPREAD}}{}_{t-\tau}^{t} + \epsilon_t$, analyzing the effects of speculative trading intensity and macroeconomic announcements on the bid-ask spread using the speculative trading intensity variable $NLS_{t,MM}$, calculated with the money manager positions. The period covered is from 2007-04-01 to 2024-02-11. The γ_m coefficients capture the instantaneous change in the bid-ask spread when a macroeconomic announcement occurs. The δ coefficient represents the effect of the speculative trading intensity variable $NLS_{t,MM}$. The θ_m coefficients capture the interaction effect between macroeconomic announcements and speculative trading intensity. The β coefficient represents the effect of the lagged bid-ask spread.

Table 1.14: Effects of Macro Announcements and Speculative Trading Intensity (Swap Dealers only) on Futures Price Bid-Ask Spreads

Commodities	**Crude Oil**		**Gold**		**Copper**		**Silver**		**Palladium**		**Natural Gas**	
Announcements	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m	γ_m	θ_m
					Macroeconomic News Announcements							
Initial jobless claims	−2.762***	8.993*	−2.060	−5.311	−0.635	−2.291	1.030	−14.269	−4.552	−14.417	−0.236	−20.006
ADP Employment	−1.537	−6.199	1.566	4.657	−1.004	−6.696	−0.020	−3.164	4.428	8.987	1.484	−32.340
CB Consumer	−2.100	−2.325	−8.627**	−22.901	−1.035	−2.639	1.088	−19.132	−2.840	−14.640	1.027	76.983***
Advance retail sales	−0.398	4.053	−4.037	−14.260	−2.884**	16.595*	−1.587	4.119	−3.706	−15.433	−3.912	24.528
Building permit	−0.476	0.181	−1.729	−6.980	−0.989	−5.019	−2.651	12.393	−20.045 ***	17.715	−1.270	−10.402
Construction spending	−2.171	−3.441	−3.305	5.828	−3.515**	−15.862	1.246	−12.790	1.030	−16.524	−2.314	−39.772
Consumer credit	−1.209	−3.802	−4.210	−19.234	−1.537	−10.764	−2.643	18.441	3.189	5.418	−0.788	−12.389
Consumer price index	−0.732	−2.252	−3.398	−10.564	−1.302	−8.591	0.418	−9.966	−16.993 **	−50.191	1.324	−41.288
Durable goods orders	−1.774	−1.828	−2.583	−8.903	−1.864	−8.853	−1.158	−7.402	−3.721	56.423*	−1.139	−26.706
Existing home sales	−3.868**	17.196*	−1.306	−5.419	−1.448	−5.318	1.868	−18.968	−3.195	−16.482	−0.816	−51.321 *
Factory orders	−1.794	−5.866	−7.113*	−20.339	−3.079**	−13.686	6.848**	52.381***	9.273	−49.326	1.513	62.452**
Gross domestic product	−3.160*	−8.265	−2.815	−1.405	−2.089	−8.872	−1.765	−1.475	−5.932	−23.151	−0.138	−30.615
Housing starts	−0.585	0.775	−2.085	−8.115	−0.890	−4.663	−3.395	20.134	−18.992 ***	23.191	−0.842	−15.893
Industrial production	−0.294	−1.051	5.507	13.455	−0.043	1.716	−2.711	19.603	2.817	10.151	−0.861	−33.181
New home sales	−2.890	−11.629	−0.676	2.292	−1.332	−7.428	0.002	−5.272	−6.636	−19.194	1.597	66.714**
Non-farm employment	−2.461	−7.536	−3.234	−3.808	−5.006***	12.326	0.734	−7.126	−13.547 *	−21.500	2.513	55.393*
Pending home sales	−2.620	−11.868	1.005	1.475	−1.269	−4.640	0.428	−15.443	−0.192	5.139	−2.254	−28.309
Personal consumption	−1.677	−6.903	−9.795**	50.083**	0.158	0.206	0.923	−10.787	−14.996 *	0.800	−0.649	−19.878
Personal income	−4.570**	−20.246	−4.806	−25.185	−1.160	−5.971	1.284	−17.017	−21.247 ***	17.208	−0.657	−31.763
Producer price index	−1.370	1.601	−4.059	−13.597	−3.143**	−14.842	−2.954	17.631	−5.662	0.279	0.315	−8.871
Trade balance	−1.050	−2.682	0.702	3.237	−0.238	6.589	2.776	−18.993	−0.382	−34.845	−0.464	2.483
Crude Oil Weekly inventory	0.059	7.795										
Natural Gas Weekly inventory											−1.312	67.433***
Observations	1,041,497		1,022,592		1,023,897		1,022,006		683,875		995,919	
$R^2(\%)$	0.071		0.063		0.20		0.29		0.031		0.48	

This table presents estimates of the equation $R^{t+\tau}_{\text{SPREAD},t} = \alpha + \sum_{m=1}^{22} \gamma_m D_{m,t} + \delta X_{t,i} + \sum_{m=1}^{22} \theta_m (D_{m,t} \cdot X_t) + \beta R^{t}_{\text{SPREAD},t-\tau} + \epsilon_t$, analyzing the effects of financialization and macroeconomic announcements on the bid-ask spread using the speculative trading intensity variable $NLS_{t,SD}$, calculated with the swap dealer positions. The period covered is from 2007-04-01 to 2024-02-11. The γ_m coefficients capture the instantaneous change in the bid-ask spread when a macroeconomic announcement occurs. The δ coefficient represents the effect of the speculative trading intensity variable $NLS_{t,SD}$. The θ_m coefficients capture the interaction effect between macroeconomic announcements and speculative trading intensity. The β coefficient represents the effect of the lagged bid-ask spread.

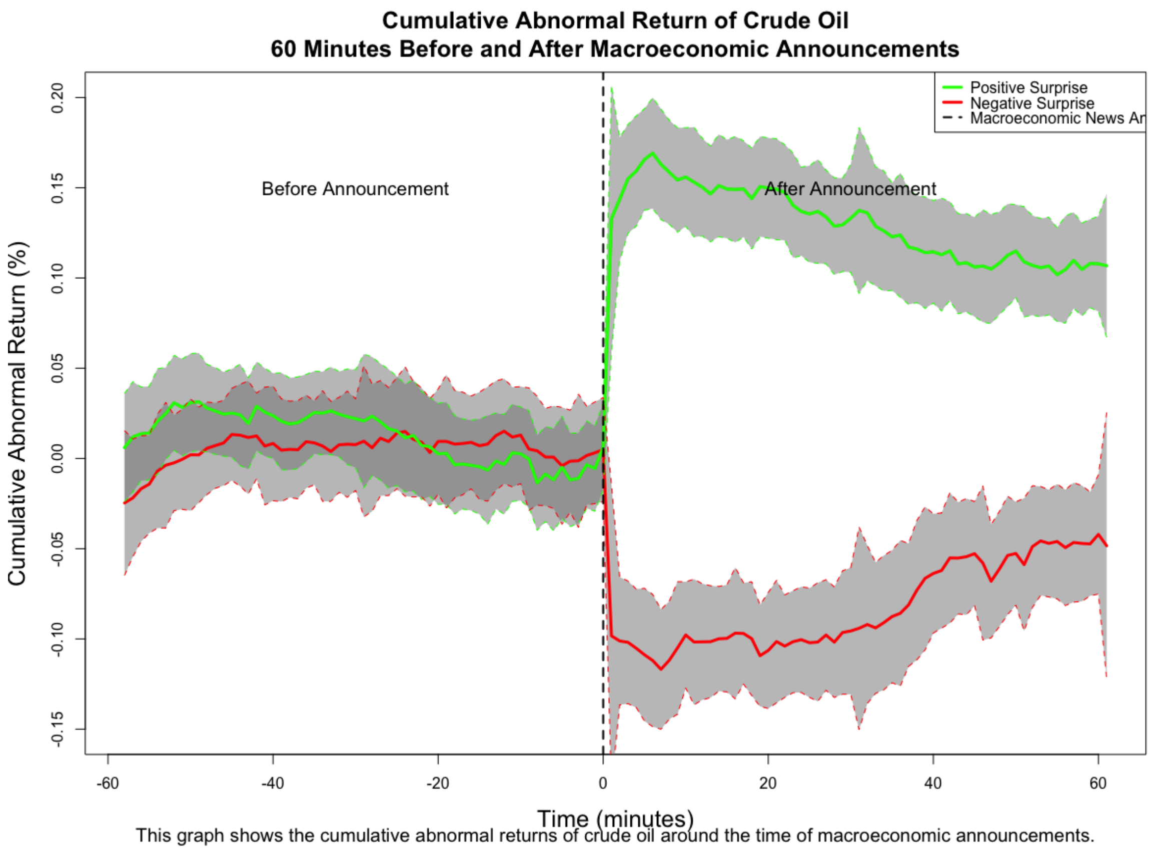


Figure 1.1: Cumulative Abnormal Return of Crude Oil 60 Minutes Before and After Macroeconomic Announcements

Figure 1.2: Cumulative Abnormal Return of Gold 60 Minutes Before and After Macroeconomic Announcements

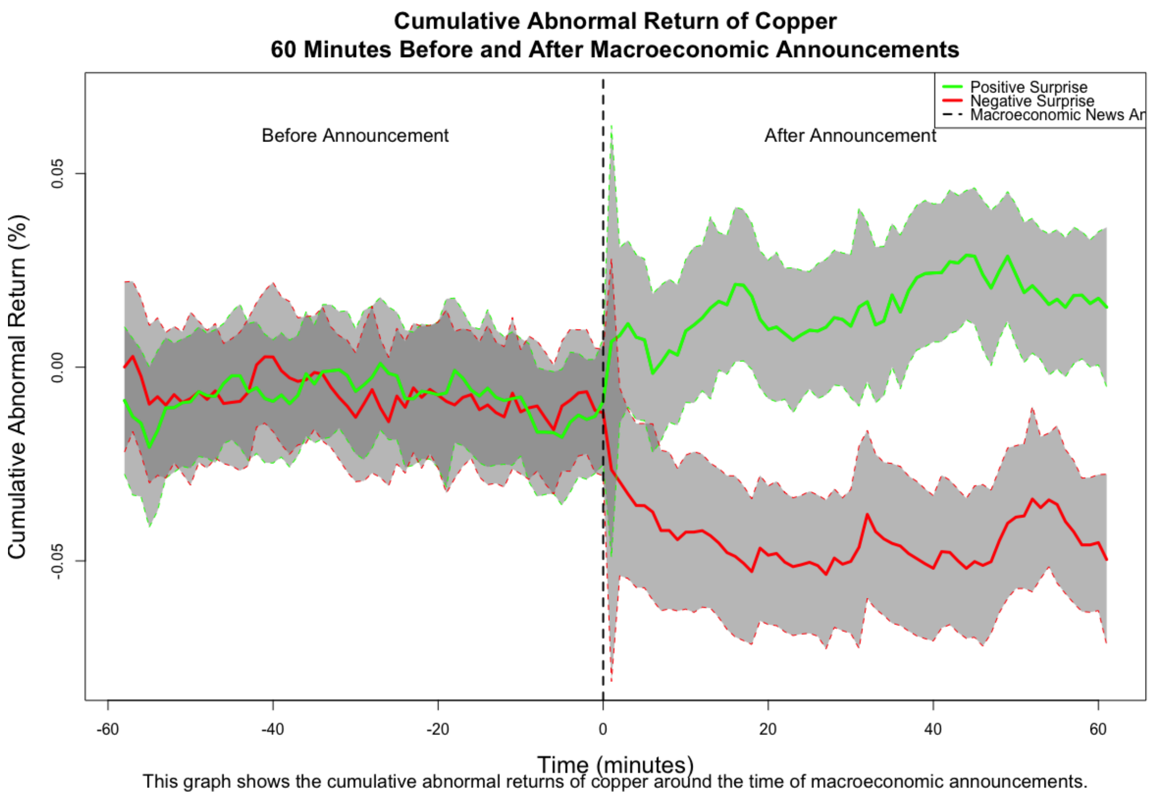


Figure 1.3: Cumulative Abnormal Return of Copper 60 Minutes Before and After Macroeconomic Announcements

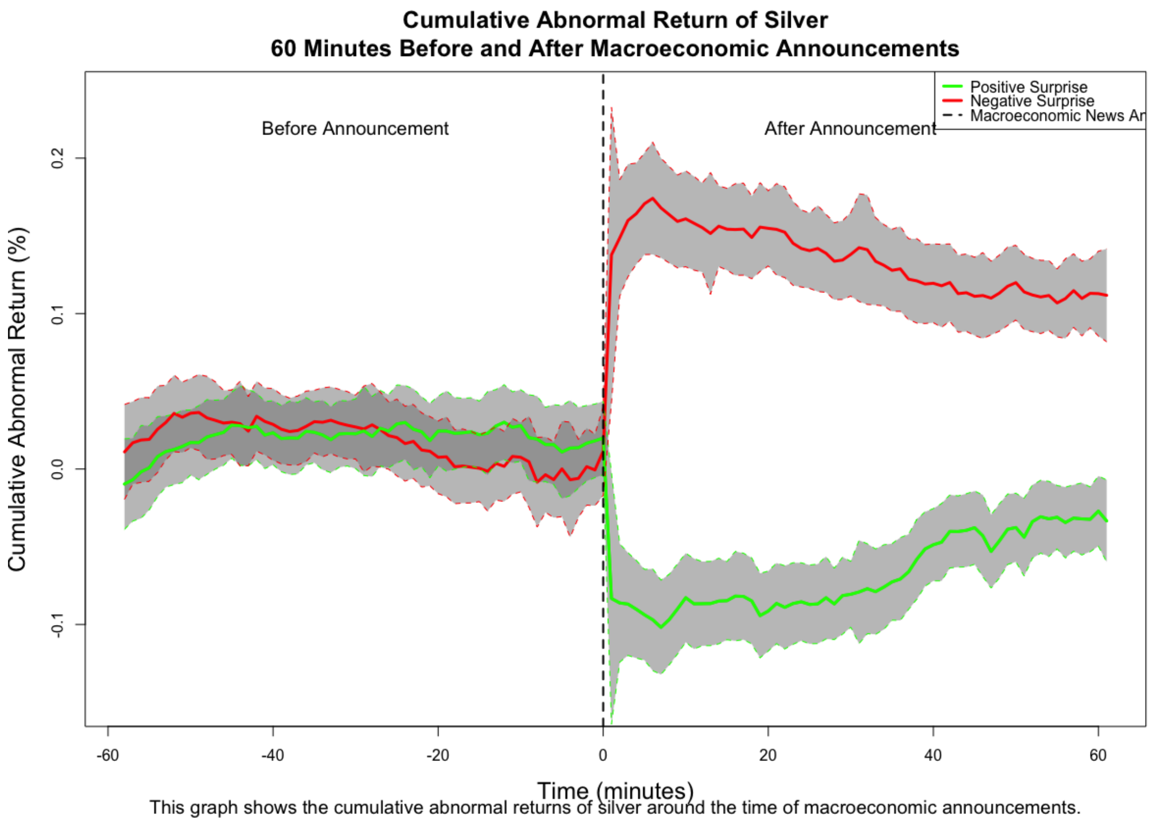
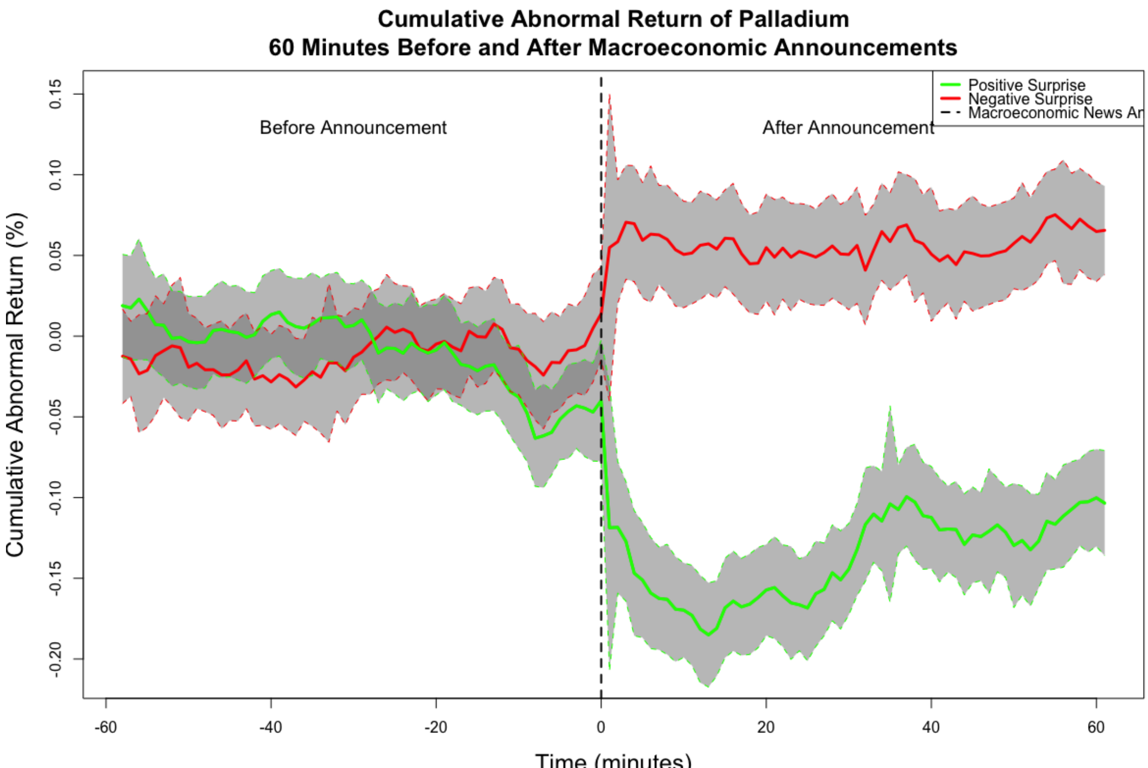
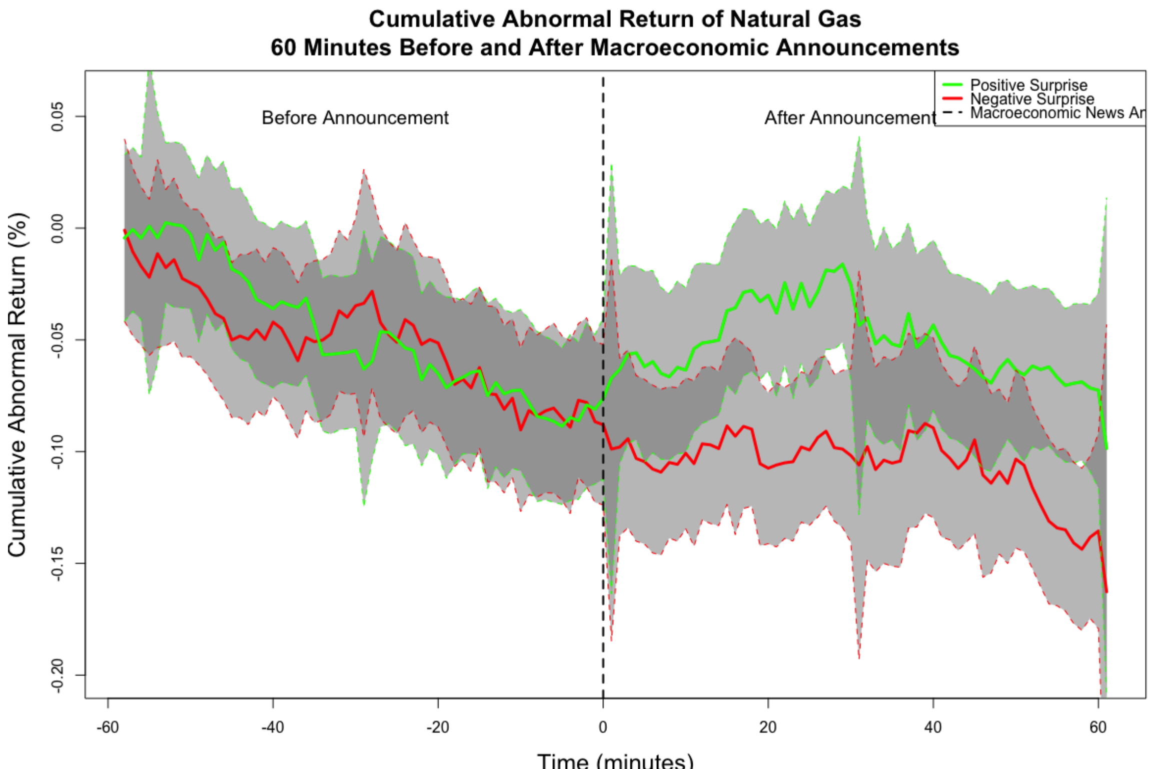


Figure 1.4: Cumulative Abnormal Return of Silver 60 Minutes Before and After Macroeconomic Announcements



This graph shows the cumulative abnormal returns of palladium around the time of macroeconomic announcements.

Figure 1.5: Cumulative Abnormal Return of Palladium 60 Minutes Before and After Macroeconomic Announcements



This graph shows the cumulative abnormal returns of natural gas around the time of macroeconomic announcements.

Figure 1.6: Cumulative Abnormal Return of Natural Gas 60 Minutes Before and After Macroeconomic Announcements

Chapter 2

Seeing Through the ETF: Indicative NAV and Commodity Volatility Transmission

2.1 Résumé

Cet article construit une base de données inédite de valeurs liquidatives indicatives (iNAV) pour les fonds négociés en bourse (FNB) de matières premières et mesure la transmission de volatilité entre les FNB et leurs paniers sous-jacents à l'aide de la variance réalisée à haute fréquence. Appliquée au pétrole brut, à l'or, à l'argent et au gaz naturel, cette approche montre que l'iNAV offre une image plus nette de la relation de volatilité. En décomposant la variance réalisée en composantes continue et de sauts, nous établissons que la transmission passe principalement par les sauts plutôt que par la diffusion, un canal que les mesures usuelles de connectivité peuvent masquer. La fréquence d'échantillonnage importe : les données à une minute révèlent une transmission quotidienne jusqu'à deux fois plus grande que les estimations à 30 minutes. Enfin, la transmission est unidirectionnelle — de l'iNAV vers le FNB — pour l'or et l'argent, mais bidirectionnelle et asymétrique pour les FNB énergétiques.

2.2 Abstract

This paper builds a novel dataset of indicative Net Asset Value (iNAV) observations for commodity ETFs and measures volatility transmission between ETFs and their underlying baskets using high-frequency realized variance. Applying this design to crude oil, gold, silver, and natural gas, we show that the iNAV provides a sharper image of the volatility relationship. Decomposing realized variance into continuous and jump components, we establish that transmission runs primarily through jumps rather than diffusion, a channel that standard volatility connectedness measures may obscure. Sampling frequency also matters: 1-minute data reveal daily volatility transmission up to twice as large as 30-minute estimates. Finally, transmission differs across commodity types. For gold and silver, it is unidirectional from iNAV to ETF, consistent with passive arbitrage. For energy ETFs, which are more liquid and hold futures as underlying assets, transmission is bidirectional and asymmetric.

Keywords: ETF, volatility, transmission, commodity markets, futures, intraday, realized variance, jump, HAR, arbitrage, microstructure.
JEL Classification: G12, G13, G14, C32

2.3 Introduction

The exchange-traded fund (ETF) market has grown rapidly over the past two decades, with global assets under management (AUM) exceeding $23 trillion as of May 2026. The growth in the ETF market has also altered the structure of underlying asset markets (Petäjistö, 2017). ETF ownership has been found to increase non-fundamental volatility in underlying equities (Ben-David et al., 2018), weaken the link between prices and fundamentals (Israeli et al., 2017), and strengthen correlations during periods of market stress (Da and Shive, 2018). The number and economic size of commodity ETFs, such as the SPDR Gold Trust (GLD), has rapidly grown in recent years. These products offer exposure to commodity markets without the need for direct futures trading or physical storage (Gorton and Rouwenhorst, 2006), thus increasing the participation of financial investors, creating new economic links between commodity and equity prices, and establishing the ETF as an alternative venue for price discovery (Basak and Pavlova, 2016; Buyuksahin and Robe, 2014).

Commodities offer a particularly interesting setting to study ETF pricing and volatility transmission. Exchange-traded funds are usually studied as near-transparent wrappers around their holdings. For example, VOO and IVV replicate the U.S. S&P 500 market index and hold the stocks that comprise this index according to their weights. For a typical equity ETF this framing is close to exact. The underlying securities trade on the same exchanges during the same hours, and creation-redemption arbitrage pins the fund's price to its net asset value quickly and in essentially one direction. Thus, for an equity ETF, asking whether the ETF leads the basket would be similar to asking whether a shadow leads the object that casts it. In contrast, commodity ETFs break this identity in a way that is not feasible in other major ETF classes. This is because their underlying may not be the physical good, but could be instead a derivative claim. For instance, if it is a futures position in crude oil or natural gas, there are economic implications for cost-of-carry, roll, and the shape of the term structure. In the case of a stored physical claim in gold and silver there is lease-and-storage economics to consider.

Thus, for commodity ETFs, the arbitrage relationship is not as simple as it is for equities or bonds, and volatility transmission becomes a genuine economic question rather than a purely mechanical one. To the point of this paper, it is important to investigate the fund's indicative net asset value (iNAV), which is the real-time fair value of that basket. Commodity ETFs also play a role in the financialization debate (Basak and Pavlova, 2016), as they are popular retail-accessible instruments. It is important for investors, traders, hedgers and policymakers to understand how volatility travels between these vehicles and their underlyings. The direction of this volatility flow is not clear a priori. Indeed, commodity ETFs differ along another significant dimension, namely that unlike equities, in commodities the ETF may well be more liquid and more continuously accessible than its underlying. This feature provides a clear reason why the ETF might contribute to price discovery rather than simply inherit it. Whether it does, however, remains an empirical question.

The theoretical foundation of ETF pricing rests on arbitrage. Authorized participants (APs) maintain price alignment through creation and redemption, buying undervalued ETFs while selling their underlying constituents, or vice versa (Ackert and Tian, 2000). Classic arbitrage theory predicts that volatility transmission would be unidirectional, flowing from the underlying constituents (NAV) to the ETF: when underlying volatility rises, arbitrageurs trade more actively to maintain alignment, which transmits volatility to the ETF. Recent evidence challenges this prediction, documenting bidirectional transmission in which ETF trading influences underlying asset volatility (Ben-David et al., 2018; Da and Shive, 2018). This outcome arises because ETFs often trade more frequently and with smaller spreads than their constituents, making the ETF a primary venue for price formation (Glosten et al., 2021; Pan and Zeng, 2016). The relative magnitude and direction of transmission therefore reveal which market dominates information discovery and whether arbitrage functions efficiently. These questions have direct implications for hedging (directional dependencies) and for regulation, e.g., as feedback effects may amplify volatility during crises (Madhavan, 2012; Petäjistö, 2017).

This paper studies how volatility is transmitted between commodity ETFs and their underlying assets using high-frequency realized variance over 2010–2023. We focus on four major single-commodity ETFs spanning two market structures: physically-backed precious metals (SPDR Gold Trust, GLD; iShares Silver Trust, SLV) and futures-based energy funds (United States Oil Fund, USO; United States Natural Gas Fund, UNG). Our central question is whether transmission is unidirectional or bidirectional, and how its direction, strength, and time horizon vary with commodity type, sampling frequency, and the continuous versus jump nature of volatility. We study volatility transmission rather than price discovery because it reveals how risk, not just information, propagates across linked markets and whether ETF arbitrage stabilizes or amplifies it—a question central to risk management and systemic-risk regulation.

The empirical literature on ETF volatility transmission is incomplete in three respects. First, it focuses on equity ETFs using daily data (Ben-David et al., 2018; Israeli et al., 2017), leaving the intraday dynamics that govern arbitrage largely unexplored. Indeed, as arbitrage operates continuously through the day, daily aggregation may obscure rapid transmission. Second, the literature has not examined how spillovers differ across commodity types, even though precious metals trade in liquid global markets with physical arbitrage while energy commodities rely on futures with rollover costs and storage constraints—features that should generate systematically different transmission. Third, it has not distinguished the roles of underlying assets versus ETFs in driving volatility, leaving open the questionwhether ETFs passively follow their constituents or actively feed back into them.

We address these gaps by constructing minute-by-minute indicative Net Asset Value (iNAV) series for the four ETFs over thirteen years spanning the European sovereign debt crisis, the 2014–2016 commodity collapse, the COVID-19 pandemic, and the subsequent inflation surge. We combine Heterogeneous Autoregressive (HAR) models that capture the long memory of realized volatility across daily, weekly, and monthly horizons (Corsi, 2009) with Bayesian Vector Autoregression (BVAR) models to accommodate time-varying dependence while avoiding overfitting (Koop, 2013). We measure realized variance from high-frequency returns (Andersen et al., 2001) and decompose it into continuous and jump components (Barndorff-Nielsen and Shephard, 2004). We separate jumps because commodity prices respond to discrete shocks—geopolitical events, supply disruptions—that may transmit through different channels than smooth price movements, with distinct consequences for tail-risk hedging. On this basis, we test four hypotheses, stated formally in Section 2.5: (i) whether transmission is uni- or bidirectional and varies by commodity type; (ii) whether high-frequency sampling reveals dynamics hidden in daily data; (iii) whether jumps or the continuous component dominate transmission; and (iv) whether transmission is stable over time.

Our results reveal substantial heterogeneity across commodity types and sampling frequencies. For precious metals (GLD, SLV), transmission is strongly unidirectional from iNAV to ETF, with spillover coefficients ranging from 0.42 (silver) to 0.63 (gold) at one-minute frequency and negligible reverse effects. Energy ETFs (USO, UNG) show bidirectional transmission, with the iNAV-to-ETF direction dominant for crude oil (about four to one) and the two directions comparable for natural gas. Sampling frequency matters: the daily iNAV-to-ETF spillover is up to roughly twice as large in one-minute as in thirty-minute data. Jump components dominate continuous transmission, especially for precious metals. The BVAR analysis confirms these patterns through impulse responses and variance decompositions in which underlying volatility rivals or exceeds ETF self-persistence in explaining ETF volatility.

The paper makes two contributions. Methodologically, we construct the first comprehensive high-frequency iNAV series for commodity ETFs over an extended period, enabling precise measurement of intraday arbitrage relationships that are unobservable with daily NAV data. Empirically, we document systematic differences in transmission across commodity categories and sampling frequencies, showing that market structure shapes information transmission and that empirical inference is affected by temporal aggregation. Thus, we extend the literature on ETF volatility transmission from equity to commodity markets (Ben-David et al., 2018; Israeli et al., 2017). We further connect it to the microstructure literature (Hasbrouck, 2003; Richie et al., 2008), and we contribute to research on commodity-ETF financialization (Todorov, 2024; Buyuksahin and Robe, 2014) by showing that the effects vary in systematic ways between physically-backed and futures-based ETFs. The remainder of the paper is as follows. Section 2.4 describes the data and the construction of high-frequency iNAV and realized-volatility series. Section 2.5 presents the econometric framework and states the hypotheses. Section 2.6 reports the empirical results. Section 2.7 concludes.

2.4 Data and Sample Construction

Our data processing approach improves on what has been previously used in the ETF volatility literature, which typically relies on daily data and may not detect high-frequency transmission channels (Ben-David et al., 2018; Israeli et al., 2017). By constructing minute-by-minute price and indicative NAV series for four commodity ETFs over thirteen years, we capture volatility dynamics at horizons not previously examined and compare volatility transmission across heterogeneous commodity market structures.

2.4.1 Sample Selection and Data Sources

We study four large single-commodity ETFs, each among the top funds in its category ranked by assets under management:

- **SPDR Gold Trust (GLD)** and **iShares Silver Trust (SLV)** are physically-backed precious-metals ETFs holding bullion in trust (over \$50 billion and \$10 billion in assets, respectively). Physical backing minimizes tracking error but creates arbitrage frictions tied to delivery, storage, and insurance.
- **United States Oil Fund (USO)** and **United States Natural Gas Fund (UNG)** are futures-based energy ETFs tracking WTI crude oil and Henry Hub natural gas through NYMEX futures. Continuous rolling of expiring contracts generates tracking errors related to contango and backwardation (Todorov, 2024), and natural gas adds extreme seasonality and storage constraints.

We select these ETFs for three reasons: high liquidity (top-five funds by assets in each category), continuous tradability over the full sample (January 2010 to January 2023) and-, most importantly, contrasting tracking mechanisms. Physically-backed ETFs (GLD, SLV) settle through physical delivery against London Bullion Market Association (LBMA) spot prices, with no rollover costs but slower arbitrage. Futures-based ETFs (USO, UNG) are arbitraged electronically against NYMEX futures, with faster settlement but involving roll costs, contango/backwardation patterns, and basis risk. Since these market structures create different frictions, transaction costs and lags, we expect them to generate different patterns of volatility transmission. We use single-commodity rather than index ETFs to avoid cross-commodity correlation effects.

Our primary data source is the Bloomberg terminal, which provides tick-by-tick trade prices, bid–ask quotes, and volume with millisecond timestamps for the ETFs and their underlying assets. We supplement this with futures contract price data from the Chicago Mercantile Exchange (CME) and spot price data from the LBMA. We begin the sample in 2010 for three reasons: By this date, the ETFs (launched 2004–2007) had plausibly reached the trading volume and market-making infrastructure needed for reliable high-frequency data and meaningful arbitrage. Moreover, starting in 2010 avoids the atypical microstructure effects during certain periods of the 2008–2009 crisis. Lastly, consistent tick-level data became available across all instruments in our sample. The resulting thirteen-year period spans the 2010–2012 European sovereign debt crisis, the 2014–2016 commodity collapse, the 2020 COVID-19 pandemic, and the 2021–2022 inflation surge, covering diverse volatility regimes.

2.4.2 High-Frequency Data Construction and Cleaning

High-frequency data require careful filtering to remove microstructure noise. Starting from raw tick data over regular U.S. market hours (9:30 AM–4:00 PM EST), we apply standard procedures adapted to the ETF market structure (Barndorff-Nielsen et al., 2009). First, we remove outliers using the Brownlees and Gallo (2006) method, deleting observations more than 10 standard deviations from a rolling 20-minute median (about 0.03% of observations). Second, we apply the duration filter suggested by Hansen and Lunde (2005), excluding trades separated by more than 30 minutes, which typically indicates closures or technical failures (less than 0.1% of observations). We then build synchronized price series at 1-, 5-, and 30-minute frequencies using previous-tick interpolation (Andersen et al., 2001). The final cleaned dataset contains roughly 45 million price observations. Data loss is minimal and concentrated in thinly traded periods.

2.4.3 Indicative Net Asset Value (iNAV) Construction

Official NAV is published only at the end of each trading day (Petäjistö, 2017). This variable cannot capture the intraday arbitrage activities that actually drive ETF pricing. We therefore construct an indicative NAV (iNAV) at a high frequency, which provides a real-time estimate of fundamental value from current underlying prices. This approach lets us measure arbitrage and transmission at the frequency where authorized participants make creation/redemption decisions, and let us separate fundamental ETF volatility (changes in underlying value) from non-fundamental volatility (liquidity shocks, inventory effects, or temporary arbitrage breakdowns).

For physically-backed ETFs (GLD, SLV), iNAV follows a composition-based approach:

$$\text{iNAV}_t = \frac{1}{N_t}\left[\text{Cash}_t + \sum_i (P_{it} \cdot f_{it} \cdot q_{it} \cdot c_{it})\right] \tag{2.1}$$

where N_t is the number of outstanding ETF shares, Cash_t is the fund's cash holdings, P_{it} is the price of underlying asset i in local currency, f_{it} the currency conversion factor, q_{it} the quantity held, and c_{it} an adjustment for accrued interest, dividends, or other cash flows. We use LBMA gold and silver prices converted to U.S. dollars at real-time exchange rates. Physical holdings are updated daily with creation/redemption activity, while intraday changes reflect only price movements.

For futures-based ETFs (USO, UNG), iNAV follows a futures-position model:

$$\text{iNAV}_t = \frac{1}{N_t}\left[\text{Cash}_t + \sum_j (F_{jt} \cdot cc_{jt} \cdot q_{jt} \cdot m_{jt})\right] \cdot FX_t \tag{2.2}$$

where F_{jt} is the price of futures contract j, cc_{jt} the contract conversion factor, q_{jt} the number of contracts held, m_{jt} the contract multiplier, and FX_t any currency conversion. We use real-time NYMEX WTI crude oil and Henry Hub natural gas futures, accounting for the funds' actual contract positions and monthly roll schedules.

A technical consideration is the mismatch between ETF trading hours and underlying market hours: gold and silver trade nearly around the clock in London and Asia, while energy futures have defined sessions. To handle this issue, we weight each market session by its share of price discovery and carry forward the most recent adjusted prices when an underlying market is closed. We validate the construction against published end-of-day NAV. We find that our iNAV has a correlation greater than 0.999 with the official NAV, and mean absolute deviations of less than 5 basis points.

2.4.4 Realized Variance Construction and Jump Detection

We measure volatility using realized variance, i.e., the sum of squared intraday returns (Andersen et al., 2001). For asset i on day t:

$$RV_{i,t} = \sum_{j=1}^{M} r_{i,t,j}^2 \tag{2.3}$$

where $r_{i,t,j} = \log(P_{i,t,j}) - \log(P_{i,t,j-1})$ and M is the number of intraday returns, giving $M = 390$ (1-minute), $M = 78$ (5-minute), and $M = 13$ (30-minute). Under standard conditions, realized variance converges to integrated variance as sampling intensifies (Barndorff-Nielsen and Shephard, 2002). The choice of frequency trades statistical efficiency against microstructure bias from bid–ask bounce (Hansen and Lunde, 2005; Liu et al., 2015), so we use three frequencies to assess robustness.

To separate continuous movements from discrete jumps, we use the bipower variation (Barndorff-Nielsen and Shephard, 2004):

$$BV_{i,t} = \mu_1^{-2} \sum_{j=2}^{M} |r_{i,t,j}| \cdot |r_{i,t,j-1}| \tag{2.4}$$

where $\mu_1 = \sqrt{2/\pi} \approx 0.798$. Bipower variation consistently estimates integrated variance even in the presence of jumps. The jump component is

$$J_{i,t} = \max(RV_{i,t} - BV_{i,t}, 0) \tag{2.5}$$

and the continuous component is $C_{i,t} = RV_{i,t} - J_{i,t}$ (Huang and Tauchen, 2005). This continuous component is reported as quadratic power variation (QPV) in the estimation tables. Separating the two matters because jumps—from supply disruptions or geopolitical events—may transmit across markets differently than smooth volatility.

2.4.5 Descriptive Statistics and Stylized Facts

Tables 2.1, 2.2, and 2.3 report descriptive statistics for realized variance, bipower variation, and jumps across the three frequencies. Several stylized facts emerge. Realized variance is heterogeneous across commodities and between ETFs and their iNAVs. Natural gas and crude oil have the highest average volatility: the 5-minute mean realized variance is 0.104%, and 0.081% for the iNAV. For crude oil, the iNAV (0.081%) is more volatile than the ETF value (0.066%), as expected from the damping effect of arbitrage. The heaviest distributional tails occur in crude oil and gold: For crude oil, the iNAV realized variance reaches a maximum above 40%. For gold, the ETF series shows occasional extreme outliers (maximum near 79%), reflecting occasional disruptions in physical gold arbitrage adjustments. Precious metals are otherwise less volatile on average than energy, though for gold the ETF mean volatility (0.034%) exceeds its iNAV mean (0.012%). Jump activity is most pronounced for crude oil, consistent with this commodity's sensitivity to geopolitical and supply shocks, and it is weakest for precious metals, for which prices move more smoothly.

Figures 2.1, 2.2, 2.3, and 2.4 show the evolution of realized variance over time. All series show strong volatility clustering, which is most pronounced during the 2014–2016 oil collapse, the 2016 Brexit referendum (precious metals), and the 2020 pandemic (for all commodities). ETF–iNAV synchronization is high for precious metals but more variable for energy commodities, where periods of close co-movement alternate with divergence. As these relationships are time-varying, a simple correlation analysis would fail to detect economically significant dynamics, which motivates the modeling framework in Section 2.5.

Finally, all realized-variance series are stationary in both levels and logs (augmented Dickey–Fuller tests), and we validate the GLD iNAV against the NYSE indicative optimized portfolio value (IOPV), obtaining correlations above 0.995. Tests for structural breaks and the full set of robustness checks are reported with the econometric framework in Section 2.5.

2.5 Econometric Methodology

We analyze volatility transmission between commodity ETFs and their iNAVs with two complementary frameworks. Heterogeneous Autoregressive (HAR) models (Corsi, 2009) capture the long memory of realized volatility across daily, weekly, and monthly horizons and, in cross-market form, yield interpretable spillover coefficients. Bayesian Vector Autoregression (BVAR) models (Koop, 2013; Carriero et al., 2009) treat ETF and iNAV volatility as jointly endogenous, using Minnesota-prior shrinkage to control parameter proliferation and providing impulse responses and variance decompositions with proper uncertainty bands. We favor realized-variance models over GARCH because realized variance is a nearly model-free volatility estimate that forecasts better (Andersen et al., 2001), and HAR over fractionally integrated (ARFIMA) specifications because it captures long memory parsimoniously through a horizon cascade (Corsi, 2009).

2.5.1 Theoretical Framework for Volatility Transmission

Under frictionless arbitrage, ETF prices equal their NAV and volatility transmission is instantaneous and bidirectional (Ackert and Tian, 2000; Petäjistö, 2017). Real-world frictions—transaction costs, inventory and funding constraints—can instead create asymmetric transmission that varies across horizons and market conditions. We study volatility transmission rather than price discovery because it reveals how risk propagates and whether arbitrage stabilizes or amplifies fluctuations, with direct implications for hedging effectiveness and systemic risk. Examining multiple horizons follows the heterogeneous market hypothesis (Müller et al., 1997), under which day traders, institutions, and longer-horizon participants generate volatility that persists over their characteristic time scales. Finally, the contrast between physically-backed ETFs (GLD, SLV), whose arbitrage requires physical delivery, and futures-based ETFs (USO, UNG), which face roll costs and basis risk (Todorov, 2024), motivates our cross-sectional comparison.

2.5.2 Heterogeneous Autoregressive (HAR) Models

The HAR model (Corsi, 2009) is the dominant framework for realized-volatility dynamics (Andersen et al., 2007). The baseline specification is

$$\log(RV_{i,t}) = \beta_0 + \beta_1 \log(RV_{i,t-1}) + \beta_2 \log(\overline{RV}_{i,t-5:t-1}) + \beta_3 \log(\overline{RV}_{i,t-22:t-1}) + \varepsilon_{i,t} \quad (2.6)$$

where $\overline{RV}_{i,t-h:t-1} = \frac{1}{h}\sum_{j=1}^{h} RV_{i,t-j}$ averages realized variance over the previous h days, capturing daily, weekly ($h = 5$), and monthly ($h = 22$) persistence. The log transform keeps fitted volatility positive and stabilizes the residual variance.

To measure transmission, we add cross-market terms. For iNAV volatility,

$$\begin{aligned}\log(RV_{t,\mathrm{NAV}}) = \beta_0 + \beta_1 \log(RV_{t-1,\mathrm{NAV}}) + \beta_2 \log(\overline{RV}_{t-5:t-1,\mathrm{NAV}}) + \beta_3 \log(\overline{RV}_{t-22:t-1,\mathrm{NAV}}) \\ + \alpha_1 \log(RV_{t-1,\mathrm{ETF}}) + \varepsilon_{t,\mathrm{NAV}}\end{aligned} \quad (2.7)$$

and symmetrically for ETF volatility,

$$\begin{aligned}\log(RV_{t,\mathrm{ETF}}) = \gamma_0 + \gamma_1 \log(RV_{t-1,\mathrm{ETF}}) + \gamma_2 \log(\overline{RV}_{t-5:t-1,\mathrm{ETF}}) + \gamma_3 \log(\overline{RV}_{t-22:t-1,\mathrm{ETF}}) \\ + \delta_1 \log(RV_{t-1,\mathrm{NAV}}) + \varepsilon_{t,\mathrm{ETF}}\end{aligned} \quad (2.8)$$

Each equation includes the own daily, weekly, and monthly terms and a single cross-market term at the daily lag, following the parsimony of the HAR cascade. The coefficient α_1 measures the daily ETF-to-iNAV spillover and δ_1 the daily iNAV-to-ETF spillover; the own weekly (β_2, γ_2) and monthly (β_3, γ_3) terms capture longer-horizon volatility persistence.

To separate the transmission of continuous and jump volatility, we estimate a HAR-CJ-X specification using the components C and J defined in Section 2.4:

$$\begin{aligned}\log(RV_{t,\mathrm{NAV}}) = \beta_0 &+ \beta_1 \log(C_{t-1,\mathrm{NAV}}) + \beta_2 \log(\overline{C}_{t-5:t-1,\mathrm{NAV}}) + \beta_3 \log(\overline{C}_{t-22:t-1,\mathrm{NAV}}) \\ &+ \beta_4 \log(1 + J_{t-1,\mathrm{NAV}}) + \beta_5 \log(1 + \overline{J}_{t-5:t-1,\mathrm{NAV}}) + \beta_6 \log(1 + \overline{J}_{t-22:t-1,\mathrm{NAV}}) \\ &+ \alpha_1 \log(C_{t-1,\mathrm{ETF}}) + \alpha_2 \log(1 + J_{t-1,\mathrm{ETF}}) + \varepsilon_{t,\mathrm{NAV}}\end{aligned} \quad (2.9)$$

with an analogous ETF equation. The own continuous and jump components enter at all three horizons through β_1–β_3 and β_4–β_6; the daily cross-market continuous and jump spillovers are α_1 and α_2. The transform $\log(1 + J_t)$ keeps jump terms defined when $J_t = 0$.

2.5.3 Bayesian Vector Autoregression (BVAR)

The HAR cascade imposes a fixed horizon structure. As a flexible complement, we estimate a VAR treating ETF and iNAV volatility as jointly endogenous:

$$\mathbf{y}_t = \mathbf{c} + \sum_{k=1}^{p} \mathbf{A}_k \mathbf{y}_{t-k} + \mathbf{u}_t \tag{2.10}$$

where $\mathbf{y}_t = [\log(RV_{t,\text{ETF}}), \log(RV_{t,\text{NAV}})]'$ and $\mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$. Information criteria (BIC) select $p = 2$. Because unrestricted VARs over-parameterize, we apply the Minnesota prior (Litterman, 1986):

$$\begin{aligned} \beta_{ij}^{(k)} &\sim \mathcal{N}(0, \lambda_1^2 \cdot k^{-\lambda_3}) \quad \text{for } i \neq j \\ \beta_{ii}^{(1)} &\sim \mathcal{N}(1, \lambda_1^2) \\ \beta_{ii}^{(k)} &\sim \mathcal{N}(0, \lambda_1^2 \cdot k^{-\lambda_3}) \quad \text{for } k > 1 \end{aligned} \tag{2.11}$$

where $\beta_{ij}^{(k)}$ is the coefficient on variable j at lag k in equation i; λ_1 controls overall tightness, λ_2 cross-variable shrinkage, and λ_3 lag decay. We set $\lambda_1 = 0.2$, $\lambda_2 = 0.5$, $\lambda_3 = 2$ and vary them in sensitivity analysis. The prior encodes that own lags matter more than others, recent lags more than distant ones, and coefficients are not extreme—assumptions well suited to persistent, mean-reverting volatility.

We estimate by Gibbs sampling, alternating between the coefficients (multivariate normal given $\mathbf{\Sigma}$) and the covariance (inverse-Wishart given the coefficients):

$$\boldsymbol{\beta}|\mathbf{\Sigma}, \mathbf{Y} \sim \mathcal{N}(\hat{\boldsymbol{\beta}}, \mathbf{\Sigma} \otimes (\mathbf{X}'\mathbf{X} + \mathbf{V}_0^{-1})^{-1}), \qquad \hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X} + \mathbf{V}_0^{-1})^{-1}(\mathbf{X}'\text{vec}(\mathbf{Y}) + \mathbf{V}_0^{-1}\boldsymbol{\beta}_0) \tag{2.12}$$

$$\mathbf{\Sigma}|\boldsymbol{\beta}, \mathbf{Y} \sim \text{IW}(\mathbf{S} + \mathbf{S}_0, T + \nu_0) \tag{2.13}$$

where $\mathbf{Y}$ and $\mathbf{X}$ are the stacked data, $\mathbf{V}_0$, $\mathbf{S}_0$, and ν_0 are prior parameters, $\mathbf{S}$ the residual sum of squares, and T the sample size. We run 50,000 iterations, discard 10,000 as burn-in, and thin every tenth draw; convergence is assessed with trace plots and the Geweke test.

From the posterior we compute orthogonalized impulse responses and forecast error variance decompositions:

$$\text{IRF}(h) = \mathbf{C}_h \mathbf{P}, \qquad \text{FEVD}_{i,j}(h) = \frac{\sum_{k=0}^{h-1} [\mathbf{C}_k \mathbf{P}]_{i,j}^2}{\sum_{k=0}^{h-1} [\mathbf{C}_k \mathbf{\Sigma} \mathbf{C}_k']_{i,i}} \tag{2.14}$$

where $\mathbf{C}_h$ is the h-step moving-average matrix and $\mathbf{P}$ the Cholesky factor of $\mathbf{\Sigma}$. We order iNAV before ETF, so iNAV innovations may affect ETF volatility contemporaneously but not the reverse, reflecting that underlying price movements lead ETF adjustments through arbitrage. The confidence bands follow Sims and Zha (1999). A high $\text{FEVD}_{\text{ETF,NAV}}(h)$ indicates that iNAV innovations explain ETF volatility, and conversely for $\text{FEVD}_{\text{NAV,ETF}}(h)$.

2.5.4 Hypotheses and Testing

We test four hypotheses, mapping each to coefficients in the equations above.

H1 (Direction and cross-commodity heterogeneity). Theory predicts unidirectional iNAV-to-ETF transmission, but a more liquid ETF can reverse the flow (Glosten et al., 2021). We expect physically-backed precious metals to show unidirectional iNAV-to-ETF transmission and futures-based energy ETFs more balanced bidirectional transmission. In equations (2.7)–(2.8), ETF-to-iNAV transmission tests $H_0 : \alpha_1 = 0$ against $H_a : \alpha_1 \neq 0$, and iNAV-to-ETF transmission tests $H_0 : \delta_1 = 0$ against $H_a : \delta_1 \neq 0$. Transmission is unidirectional when only δ_1 is significant and bidirectional when both α_1 and δ_1 are significant. Heterogeneity means these outcomes differ across commodities.

H2 (Frequency dependence). If arbitrage occurs within minutes, daily aggregation would understate transmission. Using the daily spillover coefficients δ_1 and α_1 estimated at the 1-, 5-, and 30-minute frequencies, we test $H_0 : \delta_1^{(1m)} = \delta_1^{(5m)} = \delta_1^{(30m)}$ against H_a : the coefficients differ across frequencies.

H3 (Jumps versus continuous component). Tail risk from discrete shocks is harder to hedge than smooth volatility, which is why we decompose volatility into continuous and jump parts. In the HAR-CJ-X model (2.9), we compare the daily cross-market jump and continuous spillovers, testing $H_0 : \alpha_2 = \alpha_1$ (jump and continuous transmission equal) against $H_a : \alpha_2 > \alpha_1$ (jump transmission dominates), and analogously for the ETF equation.

H4 (Stability over time). Over a data sample spanning several crises, patterns of transmission may evolve. We re-estimate the BVAR over sub-periods (2010–2014, 2015–2019, 2020–2023) and apply the structural-break test of Bai and Perron (2003), testing H_0 : transmission coefficients are constant across regimes against H_a : they change.

2.5.5 Estimation, Model Selection, and Robustness

We select lag lengths and compare nested models with the AIC and BIC (the DIC for Bayesian models, which accounts for shrinkage), and evaluate out-of-sample forecasts with rolling windows (a 1,000-day estimation window re-estimated every 250 days) at the 1-, 5-, and 22-day horizons using RMSE and MAE. Residual diagnostics include the Ljung–Box (serial correlation), Breusch–Pagan (heteroskedasticity), and Jarque–Bera (normality) tests, plus posterior predictive checks for the BVAR. We also gauge economic magnitude through the effect of one-standard-deviation shocks on forecast volatility.

We assess robustness along three dimensions, consolidating the checks noted in Section 2.4. First, *specifications*: alternative HAR lag structures, BVAR lags from one to four, level and square-root (rather than log) transforms, and time-varying-parameter versions. Second, *sample stability*: the sub-period and regime estimation underlying H4, structural-break tests (Bai and Perron, 2003)—which flag the 2014 oil collapse, the 2016 Brexit referendum, and the 2020 pandemic—and bootstrap inference. Our transmission results hold across regimes, indicating they capture general mechanisms rather than period-specific effects. Third, *data construction*: additional 15- and 60-minute frequencies, alternative estimators (truncated realized variance, realized kernels, range-based measures), and alternative iNAV constructions (currency conversion, cash treatment), including validation against the GLD IOPV. Our main findings are robust throughout.

2.6 Empirical Results

This section reports the results of our empirical analysis of volatility transmission between commodity ETFs and their underlying assets. We organize the presentation of results around the four hypotheses stated in Section 2.5, analyzing them in order: namely, the direction of transmission and its variation across commodity types (H1), the role of sampling frequency (H2), the relative importance of jump and continuous components (H3), and the stability of transmission over time (H4).

2.6.1 Direction of Transmission and Commodity-Specific Asymmetries

Our first hypothesis concerns whether transmission between an ETF and its iNAV is unidirectional or bidirectional, and whether this varies across commodities (H1). As discussed in Section 2.5, theory predicts a flow only from iNAV to ETF, but a more liquid ETF can reverse the direction of the flow. In the case of commodities, the storability and settlement mechanism is also expected to matter. We therefore expect physically-backed precious metals to show unidirectional iNAV-to-ETF transmission, and futures-based energy ETFs more balanced bidirectional transmission.

The HAR-X estimates in Tables 2.4 through 2.6 reject the null of uniform bidirectional transmission across all commodities, extending the heterogeneity documented by Gorton and Rouwenhorst (2006) and Buyuksahin and Robe (2014) to the ETF setting and consistent with Basak and Pavlova (2016) on how financialization effects vary across commodity types.

The results for precious metals show a strongly unidirectional transmission from iNAV to ETF, with little evidence of reverse transmission. Consider the iNAV-to-ETF coefficient, which is the coefficient for lagged NAV volatility in the ETF equation. For gold at a 1-minute frequency, this coefficient is 0.632 (Table 2.5), which is among the largest across commodities, while the ETF-to-iNAV effect (i.e., lagged ETF volatility in the NAV equation) is not statistically different from zero (-0.001, $p = 0.96$). The results for silver show a similar but weaker pattern: the NAV-to-ETF coefficient is 0.418 while the reverse effect is 0.004 and not significant. On the other hand, our results for energy commodities show bidirectional transmission. For crude oil at a 1-minute frequency, the iNAV-to-ETF coefficient is 0.378 and the ETF-to-iNAV coefficient is 0.089. Both are significant at the 1% level, and the coefficient for iNAV-to-ETF is roughly four times larger. The results for natural gas are the most directionally balanced: the effects in both directions are significant and of similar magnitude (iNAV-to-ETF 0.079; ETF-to-iNAV 0.127). This contrast between the strongly unidirectional results for metals and the bidirectional results for energy supports our hypothesis H1. Indeed, the evidence confirms that physical versus futures-based arbitrage, settlement mechanism, and liquidity shape the transmission of volatility between the commodity underlying and the ETF.

These patterns are confirmed by the results for the formal tests described in Section 2.5. The ETF-to-iNAV restriction $H_0 : \alpha_1 = 0$ in equation (2.7) is not rejected for gold or silver—their reverse coefficients are insignificant at every frequency—but is rejected at the 1% level for crude oil and natural gas. The iNAV-to-ETF restriction $H_0 : \delta_1 = 0$ in equation (2.8) is rejected at the 1% level for all four commodities. Transmission is therefore unidirectional (iNAV-to-ETF) for precious metals and bidirectional for energy. Moreover, the iNAV-to-ETF channel dominates for crude oil (with a magnitude of about four to one), while for natural gas the effects for the two directions are comparable.

The economic interpretation is consistent across models. In precious metals, arbitrage requires physical delivery against LBMA bullion, which is costly and slow. Authorized participants readily create or redeem ETF shares in response to underlying price moves, but cannot easily push ETF-specific shocks back into the tightly arbitraged spot market, so volatility flows essentially one way. The reverse coefficients for both metals hover near zero at all frequencies, confirming that ETF activity does not transmit volatility back to the spot market. In energy markets, by contrast, both the ETF and the underlying futures settle electronically and trade with comparable liquidity, so shocks propagate in both directions, even though fundamental supply-and-demand information still enters first through the futures-based iNAV.

Within precious metals, gold shows stronger unidirectional effects than silver at every frequency. The iNAV-to-ETF coefficients are 0.632 at a 1-minute and 0.387 at a 30-minute frequency for gold, versus 0.418 to 0.277 for silver. These results are consistent with the hypothesis that gold is a financial store of value, while silver carries additional industrial demand-related volatility. The reverse (ETF-to-iNAV) coefficients for both metals stay close to zero across frequencies: gold between -0.015 and 0.021, silver between -0.034 and 0.004, and neither is economically meaningful. In the category of energy commodities, crude oil shows a clear iNAV-to-ETF dominance across frequencies. However, the effects are more balanced for natural gas. For iNAV-to-ETF and ETF-to-iNAV, respectively, the coefficients are 0.105 and 0.126 at the 5-minute, and 0.081 and 0.074 at the 30-minute frequency. These results reflect the illiquidity, storage limits, and contango that impede arbitrage in the ETF and futures markets.

As the HAR-X specification includes cross-market terms only at the daily lag, directional transmission is identified at the daily horizon. The weekly and monthly coefficients measure each series' own persistence rather than spillovers. This own-persistence is high and, for precious metals, stable across horizons and frequencies: for instance, gold's weekly own-volatility coefficient stays near 0.33, confirming the presence of long memory in realized volatility.

2.6.2 The Relevance of Using High-Frequency Data

Our second hypothesis (H2) is that higher-frequency sampling reveals transmission that is obscured in daily data. Tables 2.4, 2.5, and 2.6 report the HAR-X estimates at 5-, 1-, and 30-minute frequencies. The daily iNAV-to-ETF transmission is markedly larger at a finer sampling frequency. For crude oil, it is measured at 0.378 at a 1-minute, 0.311 at a 5-minute, and 0.20 at a 30-minute frequency. The 1-minute estimate represents a 22% increase over the 5-minute estimate and nearly double the 30-minute estimate. Thus, much of the same-day arbitrage transmission occurs within minutes, and measurements of this mechanism would be understated at a coarser sampling, as is traditionally used in the literature.

The sensitivity of estimates according to frequency also varies by commodity. The iNAV-to-ETF transmission decreases substantially from 1-minute to 30-minute sampling in the case of crude oil (0.378 to 0.20, i.e., a 47% drop), gold (0.632 to 0.387, or 39% smaller), and silver (0.418 to 0.277, or 34% smaller). It is, however, essentially flat for natural gas (0.079 versus 0.081). The economic interpretation is that transmission in the actively and continuously arbitraged crude oil and bullion markets clears within minutes, while for natural gas this adjusts over longer horizons because storage and pipeline constraints slow down the arbitrage activities that would otherwise help equalize ETF and underlying volatility.

The reverse channel, ETF-to-iNAV transmission, is empirically weaker. Moreover, unlike the iNAV-to-ETF channel, the estimated effects do not strengthen monotonically with the sampling frequency. For crude oil, the daily ETF-to-iNAV coefficient is 0.089 at a 1-minute, 0.111 at a 5-minute, and 0.103 at a 30-minute frequency, peaking at the intermediate horizon rather than rising with the frequency. The contrast between the strongly frequency-dependent iNAV-to-ETF channel and this flatter reverse channel reinforces the directional dominance finding which we describe under hypothesis H1.

Overall, we reject the null hypothesis of frequency-invariant transmission. A Wald test of $H_0 : \delta_1^{(1\mathrm{m})} = \delta_1^{(5\mathrm{m})} = \delta_1^{(30\mathrm{m})}$ on the daily iNAV-to-ETF coefficient rejects equality at the 1% level for crude oil, gold, and silver; for natural gas, whose spillover is small and flat across frequencies, the difference is not significant. Daily data thus understate short-horizon transmission by up to roughly a factor of two for the most affected commodities, which would lead to incorrect conclusions about arbitrage effectiveness, addressing the sampling-frequency question raised by Hansen and Lunde (2005). For practitioners, this finding supports the relevance of high-frequency market monitoring to detect transmission patterns that are understated in daily analysis.

2.6.3 Jump Components and Discontinuous Volatility Transmission

Our third hypothesis (H3) concerns whether transmission is driven by the continuous (diffusion) component or by jumps. Tables 2.7, 2.8, and 2.9 report the HAR-CJ-X model estimates. Across the four commodities, the daily cross-market jump coefficient dwarfs the continuous one: discrete price moves transmit volatility, while smooth price moves essentially do not. In the dominant iNAV-to-ETF direction at a 1-minute frequency, the jump transmission (lagged NAV jump in the ETF equation) is large and positive—0.892 for gold, 0.578 for silver, 0.432 for crude oil, 0.09 for natural gas—while the corresponding continuous transmission is near zero or negative for every commodity (gold −0.123, silver −0.065, crude oil −0.019, natural gas −0.002). The effect is strongest for precious metals, where the jump component is the largest in the sample.

The same pattern holds in the reverse (ETF-to-iNAV) direction but with smaller economic magnitudes: the daily jump coefficient is 0.096 for crude oil and 0.128 for natural gas, while the continuous coefficients are near −0.01. Natural gas shows the weakest iNAV-to-ETF jump coefficient (0.090), which is consistent with natural gas volatility spikes arising from idiosyncratic, localized events—hurricanes, pipeline failures, extreme weather—that do not propagate systematically. In contrast, there is a larger value for the *own* continuous effect: i.e., the weekly own-continuous coefficient reaches 0.116 at 1-minute frequency. This result suggests a separate, within-market phenomenon rather than a cross-market transmission pattern. Testing $H_0 : \alpha_2 = \alpha_1$ (daily jump versus continuous effect) against $H_a : \alpha_2 > \alpha_1$, we find that the dominance of the jump component is confirmed at the 1% level for all four commodities. This result supports H3: transmission mainly occurs through discrete jumps rather than continuous diffusion, which has implications for tail-risk hedging. Indeed, jump-driven volatility is harder to hedge with strategies that are usually built for the assumption of continuous price processes and volatility diffusions.

2.6.4 Bayesian VAR Analysis and Stability Over Time

Our fourth hypothesis (H4) concerns whether transmission is stable over time. The BVAR results shown in Tables 2.10–2.13 characterize the joint dynamics of ETF and iNAV volatility and, through sub-period estimation, the stability of the coefficient estimates. Since the BVAR treats the two volatilities as endogenous, it further provides a check on the HAR-X spillover coefficients without imposing a cascade structure. Across the four commodities, the results for the BVAR model confirm the asymmetries documented using the HAR-X model. Moreover, the cross-market coefficients remain stable across the 2010–2014, 2015–2019, and 2020–2023 sub-periods. Thus, we do not find evidence that the transmission mechanisms meaningfully change over time. The structural break test due to Bai and Perron (2003) indicates breaks in the volatility time series during the 2014 crude oil price collapse, the 2016 Brexit referendum, and the 2020 Covid-19 pandemic. However, reestimating the BVAR model for each regime period does not change the sign and relative magnitude of the cross-market coefficients. Therefore, we fail to reject the null hypothesis (H4) of a volatility transmission model that is stable over the full sample period.

For crude oil (Table 2.10), both volatilities are strongly persistent (iNAV first-lag coefficient 0.5704, 95% credible interval [0.5237, 0.6171]; second lag 0.2606 [0.2206, 0.3005]). The ETF has only a weak effect on the iNAV (first lag 0.0735 [0.0343, 0.1135]), whereas the iNAV strongly drives the ETF: its first-lag effect (0.2861 [0.2315, 0.3433]) is comparable to ETF self-persistence (0.2941 [0.2460, 0.3418]) and its second-lag effect remains economically large (0.1934 [0.1443, 0.2410]). That the iNAV's first-lag effect rivals the ETF's own self-persistence is striking: for crude oil, underlying volatility is about as important as the ETF's recent volatility in explaining today's ETF volatility—the BVAR counterpart of the strong iNAV-to-ETF spillover found in the HAR-X estimates.

Gold (Table 2.11) shows the most asymmetric configuration. The ETF has essentially no effect on the iNAV (first lag −0.0101 [−0.0547, 0.0336], with only a marginal second-lag effect of 0.0383 [0.0001, 0.0759]), while the iNAV dominates the ETF: its first-lag effect (0.3487 [0.2848, 0.4126]) exceeds ETF self-persistence (0.1549 [0.1004, 0.2083]) by more than twofold, with a large second lag (0.2758 [0.2221, 0.3307]). Silver (Table 2.12) is intermediate: iNAV self-persistence is the highest in the sample (first lag 0.6192 [0.5631, 0.6738]), the ETF-to-iNAV effect remains negligible (−0.0446 [−0.0932, 0.0051], turning weakly positive but still economically trivial at the second lag, 0.0584 [0.0176, 0.0996]), and the iNAV-to-ETF effect is strong (first lag 0.3760 [0.3102, 0.4404] versus ETF self-persistence 0.1790 [0.1226, 0.2362]) and remains economically large at the second lag (0.2016 [0.1471, 0.2563]). Natural gas (Table 2.13) shows the clearest bidirectionality. Its iNAV is moderately persistent (first lag 0.3919 [0.3488, 0.4366], second lag 0.2313 [0.1915, 0.2715]) and, unlike the other commodities, the ETF affects the iNAV through both lags, with a larger second-lag effect (0.0799 [0.0381, 0.1214] then 0.1567 [0.1189, 0.1938]), indicating that ETF activity feeds back to the underlying market through delayed channels tied to natural-gas storage operations. The forecast error variance decompositions reinforce this interpretation. Indeed, iNAV innovations explain a large and increasing share of ETF volatility forecast errors at longer horizons, while the share of iNAV forecast-error variance attributable to ETF innovations stays small for every commodity and is near zero for gold and silver. The impulse responses examined below confirm the same asymmetry.

2.6.5 Graphical Evidence on Volatility Patterns and Dynamic Responses

The realized volatility series shown in Figures 2.1–2.4 corroborate the findings reported above. For crude oil (Figure 2.1), ETF and iNAV volatility co-move closely with synchronized peaks during stress periods, consistent with bidirectional transmission. For gold (Figure 2.2), iNAV volatility consistently precedes ETF volatility in periods of significant spikes, which is the visual counterpart of unidirectional iNAV-to-ETF transmission. Silver (Figure 2.3) shows the same pattern, but with occasional divergence, while natural gas (Figure 2.4) shows the most frequent ETF–iNAV divergences.

The impulse responses shown in Figures 2.5–2.8 tell the same story but add evidence on how the shocks evolve over time. For crude oil (Figure 2.5), iNAV shocks produce large, persistent responses in ETF volatility while ETF shocks produce small, transitory responses in the iNAV. Gold (Figure 2.6) shows the most pronounced asymmetry, with negligible responses of the iNAV to ETF shocks. Silver (Figure 2.7) is similar but noisier, while natural gas (Figure 2.8) shows sizable responses in both directions with delayed peaks, confirming bidirectional transmission.

2.6.6 Summary of the Evidence for the Hypotheses

Taken together, the estimates from the HAR-X, HAR-CJ-X and BVAR models support our four research hypotheses, and the formal tests defined in Section 2.5 reject each null hypothesis. We can summarize as follows.

H1. We find that transmission is heterogeneous: the reverse (ETF-to-iNAV) coefficient is not significant for gold and silver but it is significant for crude oil and natural gas, while the forward (iNAV-to-ETF) coefficient is significant everywhere. Transmission is therefore unidirectional from iNAV to ETF for precious metals (a magnitude about four to one greater) and bidirectional for energy commodities (the coefficients are similar in magnitude). Thus, we reject the null hypothesis of uniform bidirectional transmission and we extend the evidence from Gorton and Rouwenhorst (2006) to setting of commodity volatility. **H2.** Transmission is frequency-dependent: the daily iNAV-to-ETF spillover is up to roughly twice as large at the 1-minute as the 30-minute frequency. It is flat only for natural gas. These results confirm that daily data understate these volatility dynamics (Hansen and Lunde, 2005; Liu et al., 2015). **H3.** The daily cross-market jump component of transmission dwarfs the continuous part for all four commodities—most sharply for precious metals (i.e., the jump coefficient for gold iNAV-to-ETF jump is 0.892 versus -0.123 for the continuous component)—so transmission occurs mainly through discrete jumps. **H4.** The cross-market coefficients are stable across sub-periods and are robust to astructural-break test, so the documented mechanisms hold in general and are not specific to sub-periods.

These findings build around a single mechanism, namely that the market structure and the arbitrage mechanism that links an ETF to its underlying asset—physical delivery for precious metals versus electronic futures settlement for energy—strongly shapes the characteristics of volatility transmission (e.g., direction, magnitude, horizon, and importance of jump vs diffusion parts), consistent with the arguments of Petäjistö (2017) and Basak and Pavlova (2016). Finally, we explainthe implications for investors, market makers, and regulators in the conclusion.

2.7 Conclusion

Using high-frequency realized variance and a combination of HAR and Bayesian VAR models, this paper examines volatility transmission between four commodity ETFs and their underlying assets over the period 2010–2023. We find that volatility transmission is unidirectional from iNAV to ETF for physically-backed precious metals (GLD, SLV), while it is bidirectional (though still dominated by iNAV-to-ETF effects) for futures-based energy commodity funds (USO, UNG). Short-horizon transmission is far stronger in high-frequency data than in daily data, while longer-horizon effects are frequency-invariant; jump components transmit more strongly than continuous ones across all commodities; and these patterns are stable across sub-periods.

These results show how arbitrage operates in practice and how its effectiveness depends on market structure, extending the limits-to-arbitrage frameworks of Ackert and Tian (2000), Pontiff (1996), and Gromb and Vayanos (2010). The unidirectional transmission which we document for precious metals is consistent with physical-delivery frictions that limit ETF activity from influencing underlying prices. In contrast, the more balanced volatility transmission in energy commodity markets fits the concept of futures-based arbitrage with electronic settlement, as argued by Basak and Pavlova (2016). The frequency-specific results suggest that arbitrage operates primarily through high-frequency channels. The evidence of longer-horizon persistence reflects separate fundamental forces, building on the intraday microstructure analysis of Hasbrouck (2003) and Richie et al. (2008). The dominance of jumps in the transmission of volatility, which we identify using the bipower variation decomposition due to Barndorff-Nielsen and Shephard (2004) and Huang and Tauchen (2005), indicates that models used for ETF pricing should pay separate attention to the continuous and discontinuous components of volatility.

In terms of methodology, our high-frequency iNAV series is the first that are built specifically for commodity ETFs over an extended sample period. These new series allow us to measure intraday arbitrage relationships that are unobservable when one uses end-of-day NAV (Petäjistö, 2017). Moreover, our comparison of HAR and BVAR models, together with a thorough analysis by sampling frequency, shows that temporal aggregation strongly affects the reliability of the empirical conclusions that can be drawn (Corsi, 2009; Koop, 2013; Andersen et al., 2007).

The findings in this paper carry practical implications for different market participants and extend the risk management discussion of Madhavan (2012) and Staer (2017). For investors and risk managers, we show that precious metals ETF volatility can be forecast from underlying volatility alone, while energy ETF models must account for bidirectional feedback. Both benefit from high-frequency information for short horizons. For market makers and authorized participants, arbitrage in precious metals flows mainly from underlying markets to ETFs, while energy markets contain more balanced bidirectional opportunities (Hendershott and Riordan, 2013). For regulators, the asymmetries that we document suggest monitoring that is tailored to the commodity type: precious metals markets show little ETF feedback and thus have a lower destabilization risk, while energy markets exhibit stronger bidirectional links warranting closer monitoring during stress periods (O'Hara and Zhou, 2021; Dannhauser, 2017). The strength of the jump transmission further suggests that stress testing should account for discontinuous shock scenarios.

Finally, for future work, we note that the high-frequency iNAV methodology shown in this paper could be applied to equity, international, and fixed-income ETFs where similar arbitrage mechanisms operate. Future research could also investigate the relationship between jumps, which we find drive volatility transmission, to specific news events, order-flow imbalances, and market-maker inventory constraints.

2.8 Tables

Table 2.1: Descriptive Statistics for Realized Volatility, Quadratic Power Variation, and Jump Variables

Commodity	Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A: Realized Volatility						
Crude Oil	$RV_{t,NAV}$	3,935	0.081	0.717	0.002	42.082
	$RV_{t,ETF}$	3,935	0.066	0.204	0.001	10.335
Gold	$RV_{t,NAV}$	3,935	0.012	0.017	0.001	0.327
	$RV_{t,ETF}$	3,935	0.034	1.259	0.001	78.988
Silver	$RV_{t,NAV}$	3,935	0.043	0.063	0.002	1.120
	$RV_{t,ETF}$	3,935	0.042	0.067	0.004	1.286
Natural Gas	$RV_{t,NAV}$	3,935	0.104	0.173	0.010	6.635
	$RV_{t,ETF}$	3,935	0.103	0.116	0.004	1.651
Panel B: Quadratic Power Variation						
Crude Oil	$QPV_{t,NAV}$	3,935	0.110	6.797	0.000	426.363
	$QPV_{t,ETF}$	3,935	0.003	0.121	0.000	7.295
Gold	$QPV_{t,NAV}$	3,935	0.000	0.000	0.000	0.007
	$QPV_{t,ETF}$	3,935	0.000	0.000	0.000	0.027
Silver	$QPV_{t,NAV}$	3,935	0.000	0.003	0.000	0.139
	$QPV_{t,ETF}$	3,935	0.000	0.001	0.000	0.043
Natural Gas	$QPV_{t,NAV}$	3,935	0.003	0.152	0.000	9.555
	$QPV_{t,ETF}$	3,935	0.001	0.030	0.000	1.881
Panel C: Jump Component						
Crude Oil	$J_{t,NAV}$	3,935	−0.028	6.132	0.000	8.941
	$J_{t,ETF}$	3,935	0.063	0.124	0.000	3.040
Gold	$J_{t,NAV}$	3,935	0.012	0.017	0.001	0.324
	$J_{t,ETF}$	3,935	0.034	1.259	0.001	78.981
Silver	$J_{t,NAV}$	3,935	0.042	0.061	0.002	1.094
	$J_{t,ETF}$	3,935	0.042	0.066	0.004	1.243
Natural Gas	$J_{t,NAV}$	3,935	0.101	0.209	0.000	6.634
	$J_{t,ETF}$	3,935	0.102	0.113	0.000	1.651

This table presents descriptive statistics for realized volatility (RV), quadratic power variation (QPV), and jump component (J) variables constructed using 5-minute price data. All values are expressed in percentages. NAV refers to net asset value prices, and ETF refers to exchange-traded fund prices. The sample period includes 3,935 daily observations for each commodity.

Table 2.2: Descriptive Statistics for Realized Volatility, Quadratic Power Variation, and Jump Variables (1-minute data)

Commodity	Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A: Realized Volatility						
Crude Oil	$RV_{t,NAV}$	3,935	0.083	0.959	0.001	58.167
	$RV_{t,ETF}$	3,935	0.058	0.250	0.001	14.051
Gold	$RV_{t,NAV}$	3,935	0.012	0.018	0.001	0.302
	$RV_{t,ETF}$	3,935	0.031	1.259	0.001	78.986
Silver	$RV_{t,NAV}$	3,935	0.040	0.064	0.001	1.347
	$RV_{t,ETF}$	3,935	0.036	0.061	0.002	1.367
Natural Gas	$RV_{t,NAV}$	3,935	0.099	0.174	0.007	6.699
	$RV_{t,ETF}$	3,935	0.086	0.096	0.004	1.634
Panel B: Quadratic Power Variation						
Crude Oil	$QPV_{t,NAV}$	3,935	0.019	1.152	0.000	72.288
	$QPV_{t,ETF}$	3,935	0.004	0.191	0.000	11.936
Gold	$QPV_{t,NAV}$	3,935	0.000	0.000	0.000	0.004
	$QPV_{t,ETF}$	3,935	0.000	0.000	0.000	0.016
Silver	$QPV_{t,NAV}$	3,935	0.000	0.001	0.000	0.060
	$QPV_{t,ETF}$	3,935	0.000	0.001	0.000	0.025
Natural Gas	$QPV_{t,NAV}$	3,935	0.000	0.004	0.000	0.232
	$QPV_{t,ETF}$	3,935	0.000	0.001	0.000	0.039
Panel C: Jump Component						
Crude Oil	$J_{t,NAV}$	3,935	0.064	0.330	0.000	8.933
	$J_{t,ETF}$	3,935	0.055	0.112	0.001	3.172
Gold	$J_{t,NAV}$	3,935	0.012	0.018	0.001	0.300
	$J_{t,ETF}$	3,935	0.031	1.259	0.001	78.969
Silver	$J_{t,NAV}$	3,935	0.040	0.063	0.001	1.328
	$J_{t,ETF}$	3,935	0.036	0.060	0.002	1.344
Natural Gas	$J_{t,NAV}$	3,935	0.098	0.173	0.007	6.697
	$J_{t,ETF}$	3,935	0.086	0.095	0.004	1.634

This table presents descriptive statistics for realized volatility (RV), quadratic power variation (QPV), and jump component (J) variables constructed using 1-minute price data. All values are expressed in percentages. NAV refers to net asset value prices, and ETF refers to exchange-traded fund prices. The sample period includes 3,935 daily observations for each commodity.

Table 2.3: Descriptive Statistics for Realized Volatility, Quadratic Power Variation, and Jump Variables (30-minute data)

Commodity	Variable	Obs	Mean	Std. Dev.	Min	Max
Panel A: Realized Volatility						
Crude Oil	$RV_{t,NAV}$	3,935	0.073	0.459	0.001	23.817
	$RV_{t,ETF}$	3,935	0.052	0.176	0.000	8.995
Gold	$RV_{t,NAV}$	3,935	0.011	0.019	0.001	0.407
	$RV_{t,ETF}$	3,935	0.030	1.257	0.000	78.854
Silver	$RV_{t,NAV}$	3,935	0.039	0.069	0.001	1.701
	$RV_{t,ETF}$	3,935	0.033	0.060	0.001	1.618
Natural Gas	$RV_{t,NAV}$	3,935	0.093	0.176	0.004	6.829
	$RV_{t,ETF}$	3,935	0.077	0.093	0.002	1.668
Panel B: Quadratic Power Variation						
Crude Oil	$QPV_{t,NAV}$	3,935	0.002	0.067	0.000	4.036
	$QPV_{t,ETF}$	3,935	0.000	0.013	0.000	0.831
Gold	$QPV_{t,NAV}$	3,935	0.000	0.000	0.000	0.002
	$QPV_{t,ETF}$	3,935	0.000	0.000	0.000	0.002
Silver	$QPV_{t,NAV}$	3,935	0.000	0.001	0.000	0.050
	$QPV_{t,ETF}$	3,935	0.000	0.001	0.000	0.037
Natural Gas	$QPV_{t,NAV}$	3,935	0.000	0.001	0.000	0.044
	$QPV_{t,ETF}$	3,935	0.000	0.001	0.000	0.035
Panel C: Jump Component						
Crude Oil	$J_{t,NAV}$	3,935	0.072	0.401	0.001	19.781
	$J_{t,ETF}$	3,935	0.051	0.165	0.000	8.164
Gold	$J_{t,NAV}$	3,935	0.011	0.018	0.001	0.404
	$J_{t,ETF}$	3,935	0.030	1.257	0.000	78.852
Silver	$J_{t,NAV}$	3,935	0.038	0.068	0.001	1.665
	$J_{t,ETF}$	3,935	0.033	0.059	0.001	1.583
Natural Gas	$J_{t,NAV}$	3,935	0.093	0.175	0.004	6.823
	$J_{t,ETF}$	3,935	0.077	0.093	0.002	1.633

This table presents descriptive statistics for realized volatility (RV), quadratic power variation (QPV), and jump component (J) variables constructed using 30-minute price data. All values are expressed in percentages. NAV refers to net asset value prices, and ETF refers to exchange-traded fund prices. The sample period includes 3,935 daily observations for each commodity.

Table 2.4: HAR-X Model Estimates with 5-minute Realized Variance

	Crude Oil		**Gold**		**Silver**		**Natural Gas**	
	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$
$RV_{t-1,NAV}$	0.288***	0.311***	0.253***	0.515***	0.345***	0.383***	0.092**	0.105***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\overline{RV}_{t-5,NAV}$	0.408***		0.338***		0.318***		0.393***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$\overline{RV}_{t-22,NAV}$	0.144***		0.321***		0.271***		0.313***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$RV_{t-1,ETF}$	0.111***	0.098***	0.021	−0.084*	−0.009	−0.015	0.126***	0.084***
	(0.000)	(0.000)	(0.158)	(0.062)	(0.612)	(0.464)	(0.000)	(0.000)
$\overline{RV}_{t-5,ETF}$		0.297***		0.270***		0.209***		0.431***
		(0.000)		(0.000)		(0.000)		(0.000)
$\overline{RV}_{t-22,ETF}$		0.227***		0.164***		0.342***		0.332***
		(0.000)		(0.000)		(0.000)		(0.000)
Observations	3,935	3,935	3,935	3,935	3,935	3,935	3,935	3,935

This table presents estimation results for the HAR-X model using 5-minute realized variance data. The dependent variables are the realized variances for NAV and ETF prices of each commodity. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. RV_{t-1} represents the lagged daily realized variance, $\overline{RV}_{t-5}$ is the average of the past 5 days' realized variances, and $\overline{RV}_{t-22}$ is the average of the past 22 days' realized variances.

Table 2.5: HAR-X Model Estimates with 1-minute Realized Variance

	Crude Oil		**Gold**		**Silver**		**Natural Gas**	
	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$
$RV_{t-1,NAV}$	0.371***	0.378***	0.331***	0.632***	0.404***	0.418***	0.121***	0.079***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
$\overline{RV}_{t-5,NAV}$	0.387***		0.329***		0.297***		0.386***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$\overline{RV}_{t-22,NAV}$	0.109***		0.280***		0.226***		0.285***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$RV_{t-1,ETF}$	0.089***	0.097***	−0.001	−0.059*	0.004	0.015	0.127***	0.124***
	(0.000)	(0.000)	(0.961)	(0.065)	(0.804)	(0.410)	(0.000)	(0.000)
$\overline{RV}_{t-5,ETF}$		0.272***		0.196***		0.206***		0.438***
		(0.000)		(0.000)		(0.000)		(0.000)
$\overline{RV}_{t-22,ETF}$		0.200***		0.113***		0.276***		0.318***
		(0.000)		(0.000)		(0.000)		(0.000)
Observations	3,935	3,935	3,935	3,935	3,935	3,935	3,935	3,935

This table presents estimation results for the HAR-X model using 1-minute realized variance data. The dependent variables are the realized variances for NAV and ETF prices of each commodity. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. RV_{t-1} represents the lagged daily realized variance, $\overline{RV}_{t-5}$ is the average of the past 5 days' realized variances, and $\overline{RV}_{t-22}$ is the average of the past 22 days' realized variances.

Table 2.6: HAR-X Model Estimates with 30-minute Realized Variance

	Crude Oil		**Gold**		**Silver**		**Natural Gas**	
	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$
$RV_{t-1,NAV}$	0.133***	0.200***	0.150***	0.387***	0.226***	0.277***	0.043	0.081***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.121)	(0.000)
$\overline{RV}_{t-5,NAV}$	0.461***		0.335***		0.319***		0.399***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$\overline{RV}_{t-22,NAV}$	0.243***		0.445***		0.394***		0.393***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$RV_{t-1,ETF}$	0.103***	0.049	−0.015	−0.128**	−0.034	−0.075**	0.074**	0.012
	(0.000)	(0.149)	(0.476)	(0.019)	(0.140)	(0.013)	(0.030)	(0.662)
$\overline{RV}_{t-5,ETF}$		0.363***		0.332***		0.238***		0.405***
		(0.000)		(0.000)		(0.000)		(0.000)
$\overline{RV}_{t-22,ETF}$		0.319***		0.212***		0.460***		0.447***
		(0.000)		(0.000)		(0.000)		(0.000)
Observations	3,935	3,935	3,935	3,935	3,935	3,935	3,935	3,935

This table presents estimation results for the HAR-X model using 30-minute realized variance data. The dependent variables are the realized variances for NAV and ETF prices of each commodity. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. RV_{t-1} represents the lagged daily realized variance, $\overline{RV}_{t-5}$ is the average of the past 5 days' realized variances, and $\overline{RV}_{t-22}$ is the average of the past 22 days' realized variances.

Table 2.7: HAR-CJ-X Model Estimates with 5-minute Realized Variance

	Crude Oil		**Gold**		**Silver**		**Natural Gas**	
	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$
Panel A: Quadratic Power Variation								
$QPV_{t-1,NAV}$	0.049**	0.008	0.026	−0.001	−0.015	−0.026	0.065***	0.016
	(0.025)	(0.661)	(0.239)	(0.957)	(0.467)	(0.273)	(0.001)	(0.400)
$\overline{QPV}_{t-5,NAV}$	0.032		−0.027		−0.030		0.110***	
	(0.167)		(0.317)		(0.236)		(0.000)	
$\overline{QPV}_{t-22,NAV}$	−0.046***		−0.021		−0.005		0.044**	
	(0.007)		(0.341)		(0.793)		(0.018)	
$QPV_{t-1,ETF}$	0.001	0.043**	−0.015	−0.030	0.006	0.010	0.011	0.028*
	(0.969)	(0.038)	(0.545)	(0.248)	(0.779)	(0.686)	(0.620)	(0.088)
$\overline{QPV}_{t-5,ETF}$		0.017		−0.008		−0.003		0.062***
		(0.454)		(0.737)		(0.876)		(0.001)
$\overline{QPV}_{t-22,ETF}$		−0.040**		0.133***		0.006		0.038
		(0.043)		(0.000)		(0.763)		(0.137)
Panel B: Jump Component								
$J_{t-1,NAV}$	0.207***	0.280***	0.212***	0.491***	0.375***	0.431***	−0.003	0.057*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.923)	(0.082)
$\overline{J}_{t-5,NAV}$	0.338***		0.407***		0.389***		0.220***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$\overline{J}_{t-22,NAV}$	0.267***		0.354***		0.272***		0.213***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$J_{t-1,ETF}$	0.094***	0.019	0.041	−0.023	−0.016	−0.030	0.044	0.039
	(0.000)	(0.473)	(0.156)	(0.497)	(0.527)	(0.291)	(0.190)	(0.195)
$\overline{J}_{t-5,ETF}$		0.264***		0.246***		0.214***		0.307***
		(0.000)		(0.000)		(0.000)		(0.000)
$\overline{J}_{t-22,ETF}$		0.329***		−0.014		0.327***		0.250***
		(0.000)		(0.631)		(0.000)		(0.000)
Observations	3,935	3,935	3,935	3,935	3,935	3,935	3,935	3,935

This table presents estimation results for the HAR-CJ-X model using 5-minute realized variance data. The model incorporates both continuous (quadratic power variation, QPV) and jump (J) components. The dependent variables are the realized variances for NAV and ETF prices of each commodity. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. QPV_{t-1} and J_{t-1} represent the lagged daily components, while $\overline{QPV}_{t-5}$, $\overline{J}_{t-5}$, $\overline{QPV}_{t-22}$, and $\overline{J}_{t-22}$ are the corresponding weekly and monthly averages.

Table 2.8: HAR-CJ-X Model Estimates with 1-minute Realized Variance

	Crude Oil		**Gold**		**Silver**		**Natural Gas**	
	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$
Panel A: Quadratic Power Variation								
$QPV_{t-1,NAV}$	0.036	−0.019	−0.033*	−0.123***	−0.041**	−0.065***	0.008	−0.002
	(0.246)	(0.525)	(0.096)	(0.000)	(0.039)	(0.002)	(0.773)	(0.920)
$\overline{QPV}_{t-5,NAV}$	−0.021		−0.076***		−0.068***		0.116***	
	(0.444)		(0.004)		(0.009)		(0.000)	
$\overline{QPV}_{t-22,NAV}$	−0.031		−0.001		0.009		−0.040**	
	(0.309)		(0.976)		(0.716)		(0.046)	
$QPV_{t-1,ETF}$	−0.011	0.009	−0.020*	−0.032*	−0.011	−0.014	−0.011	−0.021*
	(0.665)	(0.698)	(0.062)	(0.089)	(0.617)	(0.514)	(0.671)	(0.089)
$\overline{QPV}_{t-5,ETF}$		−0.014		0.016		0.003		0.016
		(0.591)		(0.566)		(0.895)		(0.554)
$\overline{QPV}_{t-22,ETF}$		−0.024**		0.060***		−0.017		−0.014
		(0.036)		(0.003)		(0.463)		(0.617)
Panel B: Jump Component								
$J_{t-1,NAV}$	0.306***	0.432***	0.426***	0.892***	0.510***	0.578***	0.092*	0.090***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.075)	(0.000)
$\overline{J}_{t-5,NAV}$	0.435***		0.509***		0.468***		0.219***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$\overline{J}_{t-22,NAV}$	0.190***		0.234***		0.161***		0.346***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$J_{t-1,ETF}$	0.096***	0.077**	0.031	0.011	0.028	0.051	0.128***	0.166***
	(0.000)	(0.019)	(0.290)	(0.689)	(0.313)	(0.106)	(0.000)	(0.000)
$\overline{J}_{t-5,ETF}$		0.300***		0.132**		0.184***		0.401***
		(0.000)		(0.023)		(0.000)		(0.000)
$\overline{J}_{t-22,ETF}$		0.250***		0.028		0.295***		0.350***
		(0.000)		(0.455)		(0.000)		(0.000)
Observations	3,935	3,935	3,935	3,935	3,935	3,935	3,935	3,935

This table presents estimation results for the HAR-CJ-X model using 1-minute realized variance data. The model incorporates both continuous (quadratic power variation, QPV) and jump (J) components. The dependent variables are the realized variances for NAV and ETF prices of each commodity. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. QPV_{t-1} and J_{t-1} represent the lagged daily components, while $\overline{QPV}_{t-5}$, $\overline{J}_{t-5}$, $\overline{QPV}_{t-22}$, and $\overline{J}_{t-22}$ are the corresponding weekly and monthly averages.

Table 2.9: HAR-CJ-X Model Estimates with 30-minute Realized Variance

	Crude Oil		**Gold**		**Silver**		**Natural Gas**	
	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$	$RV_{t,NAV}$	$RV_{t,ETF}$
Panel A: Quadratic Power Variation								
$QPV_{t-1,NAV}$	0.019	−0.026	0.058***	0.067***	0.013	0.033	0.057***	0.027
	(0.535)	(0.421)	(0.005)	(0.003)	(0.596)	(0.235)	(0.008)	(0.278)
$\overline{QPV}_{t-5,NAV}$	0.029		−0.051**		−0.008		0.114***	
	(0.375)		(0.044)		(0.779)		(0.000)	
$\overline{QPV}_{t-22,NAV}$	−0.001		0.001		−0.045*		0.156***	
	(0.975)		(0.977)		(0.098)		(0.000)	
$QPV_{t-1,ETF}$	−0.010	0.015	−0.018	−0.027	−0.006	0.001	−0.014	0.001
	(0.683)	(0.562)	(0.489)	(0.313)	(0.799)	(0.984)	(0.575)	(0.972)
$\overline{QPV}_{t-5,ETF}$		0.055**		−0.017		−0.001		0.079***
		(0.033)		(0.517)		(0.966)		(0.001)
$\overline{QPV}_{t-22,ETF}$		0.016		0.234***		−0.035		0.168***
		(0.517)		(0.000)		(0.191)		(0.000)
Panel B: Jump Component								
$J_{t-1,NAV}$	0.110**	0.221***	0.054	0.206***	0.201***	0.216***	−0.006	0.027
	(0.019)	(0.000)	(0.190)	(0.000)	(0.000)	(0.000)	(0.870)	(0.381)
$\overline{J}_{t-5,NAV}$	0.403***		0.436***		0.339***		0.204***	
	(0.000)		(0.000)		(0.000)		(0.000)	
$\overline{J}_{t-22,NAV}$	0.248***		0.440***		0.489***		0.089	
	(0.000)		(0.000)		(0.000)		(0.123)	
$J_{t-1,ETF}$	0.110***	0.043	0.009	−0.059	−0.023	−0.080*	0.016	0.007
	(0.000)	(0.181)	(0.772)	(0.183)	(0.428)	(0.077)	(0.608)	(0.823)
$\overline{J}_{t-5,ETF}$		0.259***		0.274***		0.241***		0.243***
		(0.000)		(0.000)		(0.000)		(0.000)
$\overline{J}_{t-22,ETF}$		0.283***		−0.082*		0.532***		0.088
		(0.000)		(0.058)		(0.000)		(0.224)
Observations	3,935	3,935	3,935	3,935	3,935	3,935	3,935	3,935

This table presents estimation results for the HAR-CJ-X model using 30-minute realized variance data. The model incorporates both continuous (quadratic power variation, QPV) and jump (J) components. The dependent variables are the realized variances for NAV and ETF prices of each commodity. Standard errors are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. QPV_{t-1} and J_{t-1} represent the lagged daily components, while $\overline{QPV}_{t-5}$, $\overline{J}_{t-5}$, $\overline{QPV}_{t-22}$, and $\overline{J}_{t-22}$ are the corresponding weekly and monthly averages.

Table 2.10: Bayesian Vector Autoregression Results: USO ETF and Net Asset Value

Variable	Mean	Std. Dev.	MCSE	Median	95% Credible Interval	
Panel A: $\log(RV_{t,NAV})$					Lower	Upper
$\log(RV_{t-1,NAV})$	0.570	0.024	0.000	0.570	0.524	0.617
$\log(RV_{t-2,NAV})$	0.261	0.021	0.000	0.261	0.221	0.300
$\log(RV_{t-1,ETF})$	0.074	0.020	0.000	0.074	0.034	0.113
$\log(RV_{t-2,ETF})$	0.016	0.018	0.000	0.016	-0.019	0.051
Panel B: $\log(RV_{t,ETF})$						
Variable	Mean	Std. Dev.	MCSE	Median	95% Credible Interval	
					Lower	Upper
$\log(RV_{t-1,NAV})$	0.286	0.028	0.000	0.286	0.231	0.343
$\log(RV_{t-2,NAV})$	0.193	0.025	0.000	0.193	0.144	0.241
$\log(RV_{t-1,ETF})$	0.294	0.024	0.000	0.294	0.246	0.342
$\log(RV_{t-2,ETF})$	0.106	0.022	0.000	0.106	0.064	0.148

This table presents posterior statistics from a Bayesian Vector Autoregression (VAR) model analyzing the relationship between the USO ETF and its Net Asset Value (NAV). Panel A shows results for the NAV equation, while Panel B shows results for the ETF equation. The model includes two lags of both log realized variances. Mean represents the posterior mean, Std. Dev. is the posterior standard deviation, MCSE is the Monte Carlo standard error, Median is the posterior median, and the 95% credible interval provides the range of plausible parameter values. All realized variance measures are constructed using 5-minute price data.

Table 2.11: Bayesian Vector Autoregression Results: GLD ETF and Net Asset Value

	Panel A: $\log(RV_{t,NAV})$					
Variable	Mean	Std. Dev.	MCSE	Median	95% Credible Interval	
					Lower	Upper
$\log(RV_{t-1,NAV})$	0.551	0.027	0.000	0.551	0.498	0.603
$\log(RV_{t-2,NAV})$	0.248	0.023	0.000	0.248	0.203	0.293
$\log(RV_{t-1,ETF})$	-0.010	0.023	0.000	-0.010	-0.055	0.034
$\log(RV_{t-2,ETF})$	0.038	0.019	0.000	0.038	0.000	0.076

	Panel B: $\log(RV_{t,ETF})$					
Variable	Mean	Std. Dev.	MCSE	Median	95% Credible Interval	
					Lower	Upper
$\log(RV_{t-1,NAV})$	0.349	0.033	0.000	0.348	0.285	0.413
$\log(RV_{t-2,NAV})$	0.276	0.028	0.000	0.276	0.222	0.331
$\log(RV_{t-1,ETF})$	0.155	0.028	0.000	0.155	0.100	0.208
$\log(RV_{t-2,ETF})$	0.048	0.023	0.000	0.048	0.003	0.095

This table presents posterior statistics from a Bayesian Vector Autoregression (VAR) model analyzing the relationship between the GLD ETF and its Net Asset Value (NAV). Panel A shows results for the NAV equation, while Panel B shows results for the ETF equation. The model includes two lags of both log realized variances. Mean represents the posterior mean, Std. Dev. is the posterior standard deviation, MCSE is the Monte Carlo standard error, Median is the posterior median, and the 95% credible interval provides the range of plausible parameter values. All realized variance measures are constructed using 5-minute price data.

Table 2.12: Bayesian Vector Autoregression Results: SLV ETF and Net Asset Value

Variable	Mean	Std. Dev.	MCSE	Median	95% Credible Interval	
					Lower	Upper
	Panel A: $\log(RV_{t,NAV})$					
$\log(RV_{t-1,NAV})$	0.619	0.028	0.000	0.619	0.563	0.674
$\log(RV_{t-2,NAV})$	0.214	0.024	0.000	0.214	0.168	0.261
$\log(RV_{t-1,ETF})$	−0.045	0.025	0.000	−0.045	−0.093	0.005
$\log(RV_{t-2,ETF})$	0.058	0.021	0.000	0.058	0.018	0.100
	Panel B: $\log(RV_{t,ETF})$					
Variable	Mean	Std. Dev.	MCSE	Median	95% Credible Interval	
					Lower	Upper
$\log(RV_{t-1,NAV})$	0.376	0.033	0.000	0.377	0.310	0.440
$\log(RV_{t-2,NAV})$	0.202	0.028	0.000	0.202	0.147	0.256
$\log(RV_{t-1,ETF})$	0.179	0.029	0.000	0.179	0.123	0.236
$\log(RV_{t-2,ETF})$	0.078	0.024	0.000	0.078	0.031	0.126

This table presents posterior statistics from a Bayesian Vector Autoregression (VAR) model analyzing the relationship between the SLV ETF and its Net Asset Value (NAV). Panel A shows results for the NAV equation, while Panel B shows results for the ETF equation. The model includes two lags of both log realized variances. Mean represents the posterior mean, Std. Dev. is the posterior standard deviation, MCSE is the Monte Carlo standard error, Median is the posterior median, and the 95% credible interval provides the range of plausible parameter values. All realized variance measures are constructed using 5-minute price data.

Table 2.13: Bayesian Vector Autoregression Results: UNG ETF and Net Asset Value

Variable	Panel A: $\log(RV_{t,NAV})$					
	Mean	Std. Dev.	MCSE	Median	95% Credible Interval	
					Lower	Upper
$\log(RV_{t-1,NAV})$	0.392	0.022	0.000	0.392	0.349	0.437
$\log(RV_{t-2,NAV})$	0.231	0.020	0.000	0.231	0.192	0.272
$\log(RV_{t-1,ETF})$	0.080	0.021	0.000	0.080	0.038	0.121
$\log(RV_{t-2,ETF})$	0.157	0.019	0.000	0.157	0.119	0.194
	Panel B: $\log(RV_{t,ETF})$					
Variable	Mean	Std. Dev.	MCSE	Median	95% Credible Interval	
					Lower	Upper
$\log(RV_{t-1,NAV})$	0.130	0.024	0.000	0.130	0.083	0.177
$\log(RV_{t-2,NAV})$	0.165	0.022	0.000	0.165	0.124	0.208
$\log(RV_{t-1,ETF})$	0.339	0.023	0.000	0.339	0.294	0.383
$\log(RV_{t-2,ETF})$	0.240	0.020	0.000	0.240	0.200	0.280

This table presents posterior statistics from a Bayesian Vector Autoregression (VAR) model analyzing the relationship between the UNG ETF and its Net Asset Value (NAV). Panel A shows results for the NAV equation, while Panel B shows results for the ETF equation. The model includes two lags of both log realized variances. Mean represents the posterior mean, Std. Dev. is the posterior standard deviation, MCSE is the Monte Carlo standard error, Median is the posterior median, and the 95% credible interval provides the range of plausible parameter values. All realized variance measures are constructed using 5-minute price data.

2.9 Figures

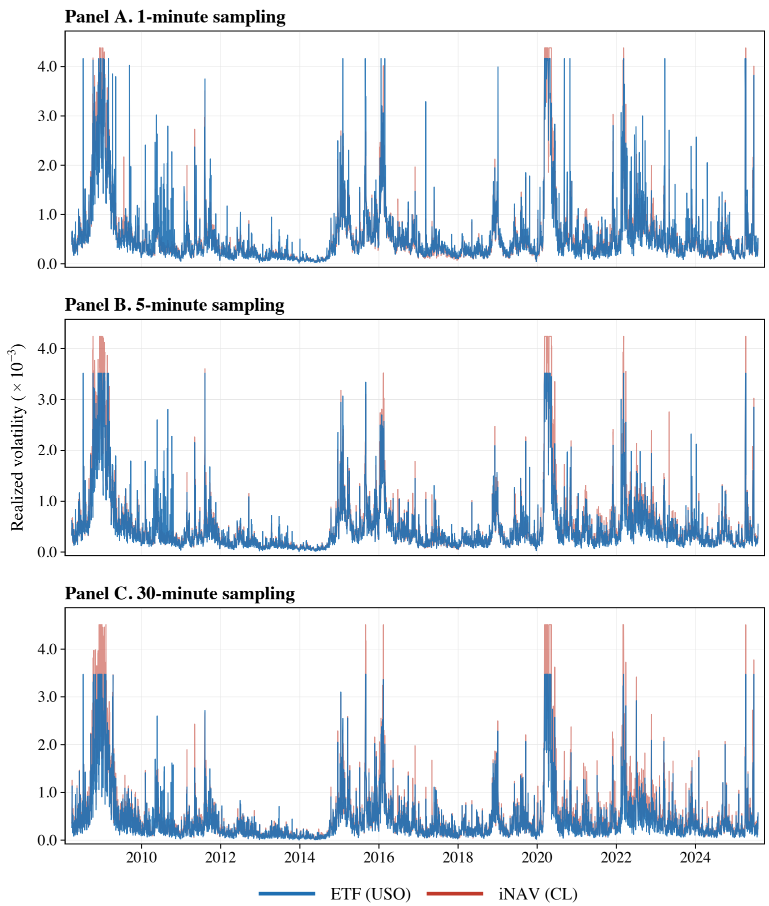


Figure 2.1: Realized volatility of the Crude Oil ETF and its iNAV. The figure plots daily realized volatility (sum of squared intraday log returns, $\times 10^{-3}$) for the USO ETF (blue) and its iNAV, proxied by the CL futures contract (red), from 2008-03-27 to 2025-08-01. Panels A, B, and C use returns sampled at the 1-minute, 5-minute, and 30-minute frequencies, respectively. For readability, each series is winsorized at its 99th percentile within each panel.

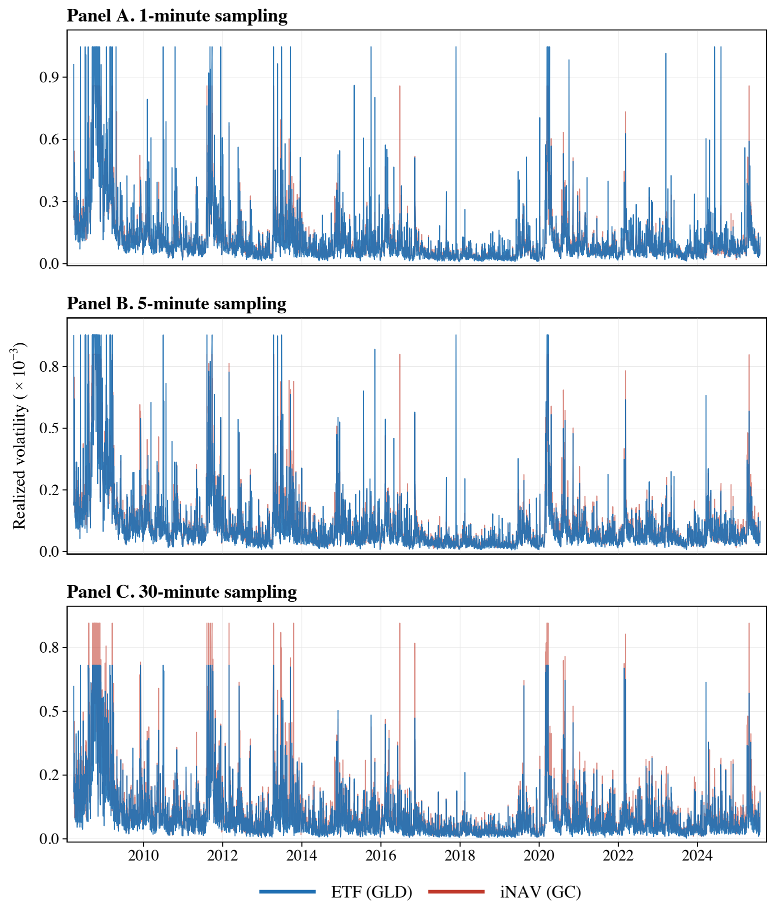


Figure 2.2: Realized volatility of the Gold ETF and its iNAV. The figure plots daily realized volatility (sum of squared intraday log returns, $\times 10^{-3}$) for the GLD ETF (blue) and its iNAV, proxied by the GC futures contract (red), from 2008-03-27 to 2025-08-01. Panels A, B, and C use returns sampled at the 1-minute, 5-minute, and 30-minute frequencies, respectively. For readability, each series is winsorized at its 99th percentile within each panel.

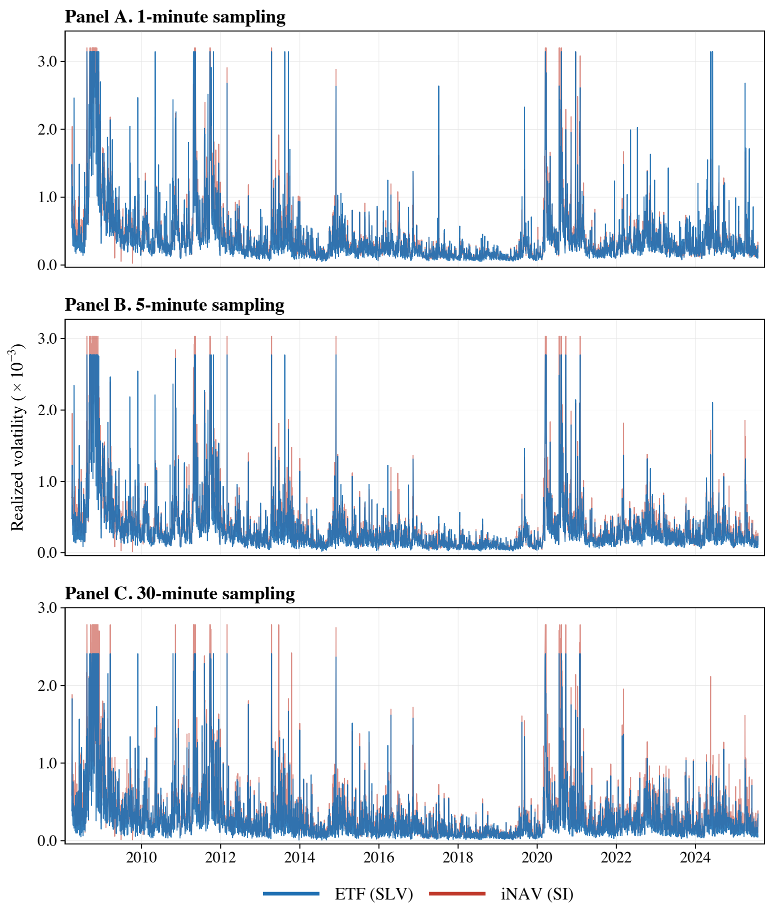


Figure 2.3: Realized volatility of the Silver ETF and its iNAV. The figure plots daily realized volatility (sum of squared intraday log returns, $\times 10^{-3}$) for the SLV ETF (blue) and its iNAV, proxied by the SI futures contract (red), from 2008-03-27 to 2025-08-01. Panels A, B, and C use returns sampled at the 1-minute, 5-minute, and 30-minute frequencies, respectively. For readability, each series is winsorized at its 99th percentile within each panel.

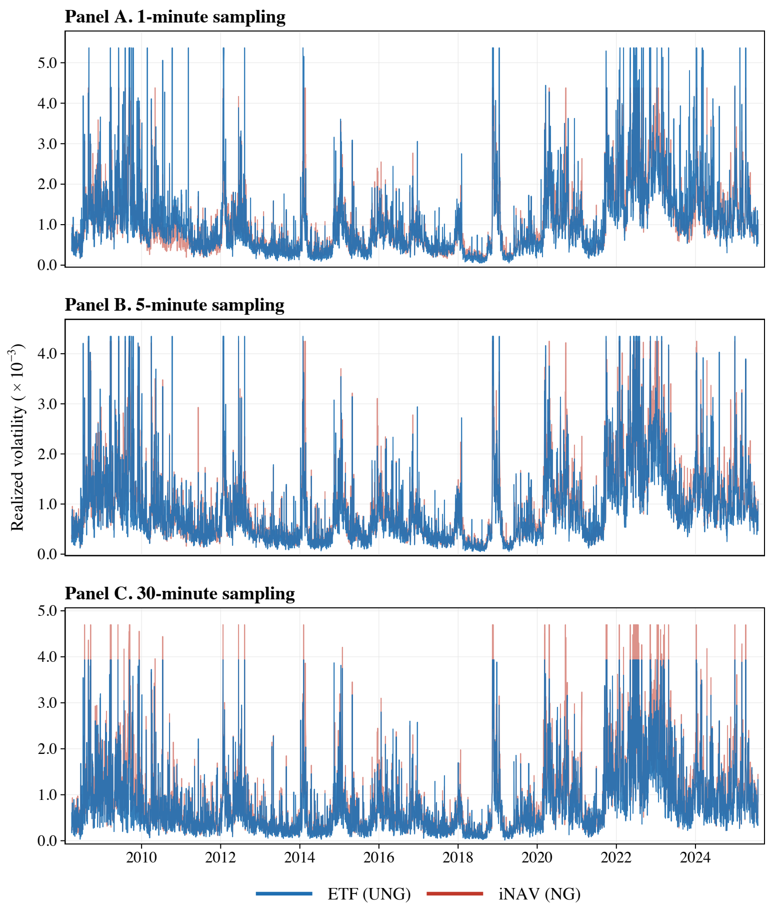
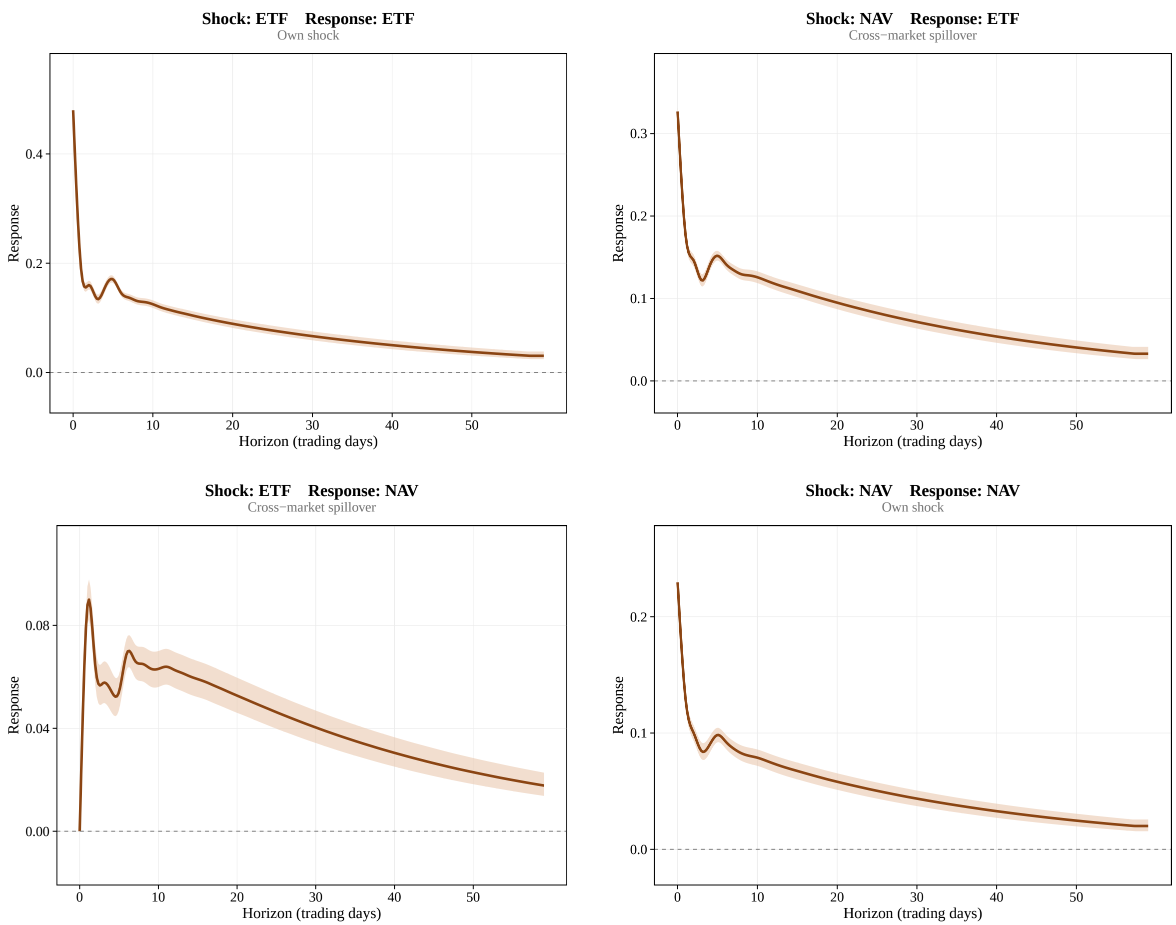
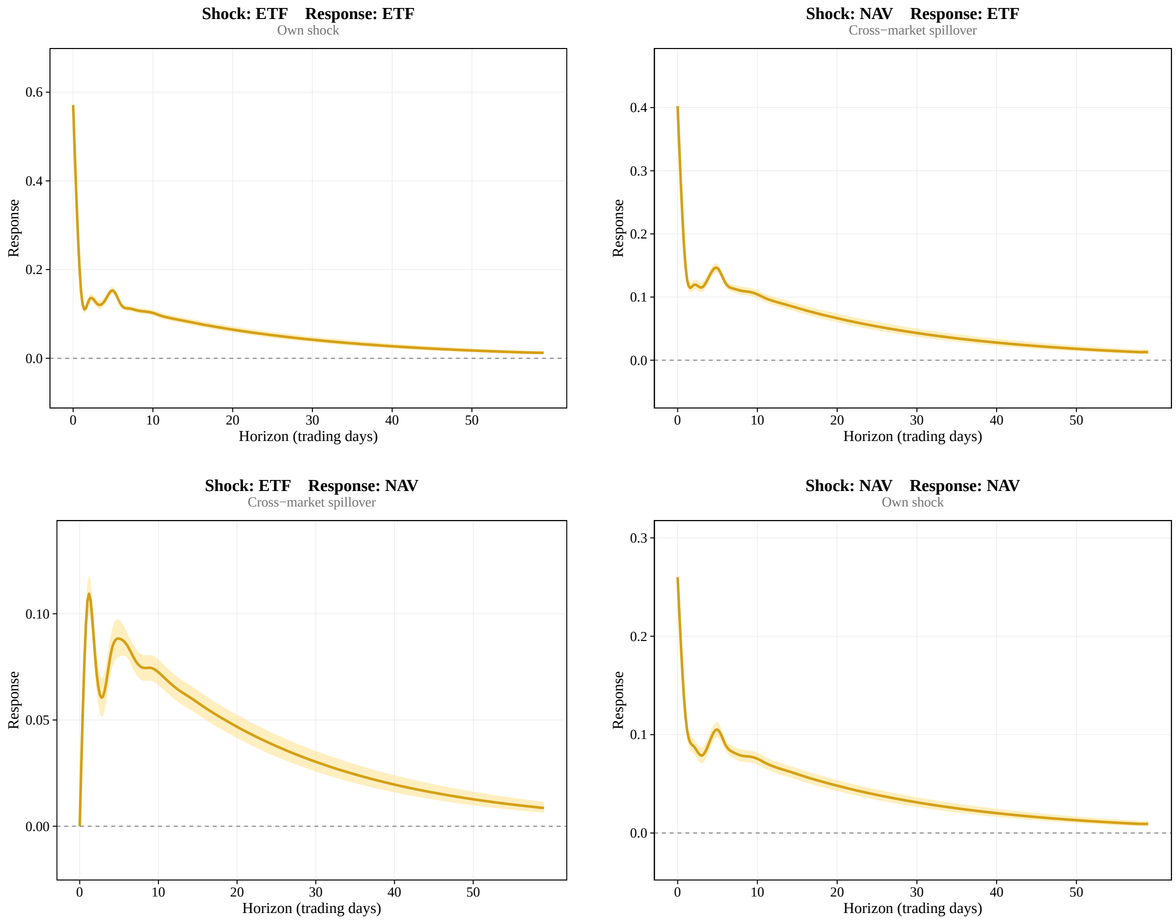


Figure 2.4: Realized volatility of the Natural Gas ETF and its iNAV. The figure plots daily realized volatility (sum of squared intraday log returns, $\times 10^{-3}$) for the UNG ETF (blue) and its iNAV, proxied by the NG futures contract (red), from 2008-03-27 to 2025-08-01. Panels A, B, and C use returns sampled at the 1-minute, 5-minute, and 30-minute frequencies, respectively. For readability, each series is winsorized at its 99th percentile within each panel.



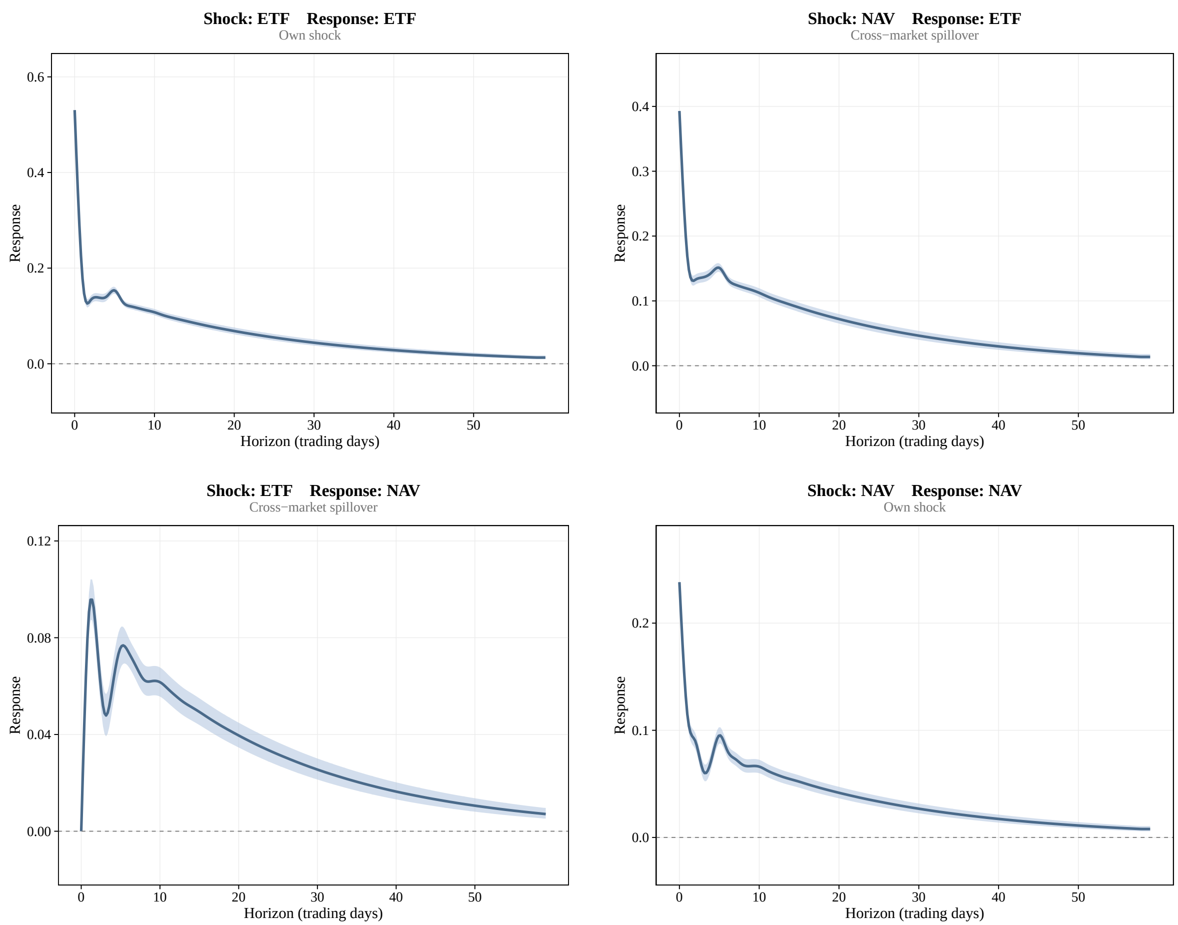
Cholesky ordering: [ETF, NAV]. ETF restricted at h=0. NAV (futures) can lead. 68% credible intervals.

Figure 2.5: Impulse Response Functions (IRF) for USO ETF and Crude Oil (CL) iNAV Volatility. Bayesian VAR impulse responses showing fast shock absorption with systematic asymmetry: iNAV shocks cause large and persistent responses in ETF volatility while ETF shocks generate smaller and more transitory effects on iNAV, validating bidirectional but asymmetric transmission in crude oil markets.



Cholesky ordering: [ETF, NAV]. ETF restricted at h=0. NAV (futures) can lead. 68% credible intervals.

Figure 2.6: Impulse Response Functions (IRF) for GLD ETF and Gold (GC) iNAV Volatility. Bayesian VAR impulse responses exhibiting the most pronounced asymmetries: iNAV shocks create large and persistent responses in ETF volatility that decay smoothly, while ETF shocks generate negligible responses in iNAV, providing strong dynamic evidence for unidirectional transmission in gold markets.



Cholesky ordering: [ETF, NAV]. ETF restricted at h=0. NAV (futures) can lead. 68% credible intervals.

Figure 2.7: Impulse Response Functions (IRF) for SLV ETF and Silver (SI) iNAV Volatility. Bayesian VAR impulse responses preserving the gold market's asymmetric structure but with qualitatively richer adjustment dynamics and greater variability, reflecting silver's dual function as both an industrial and a precious metal.

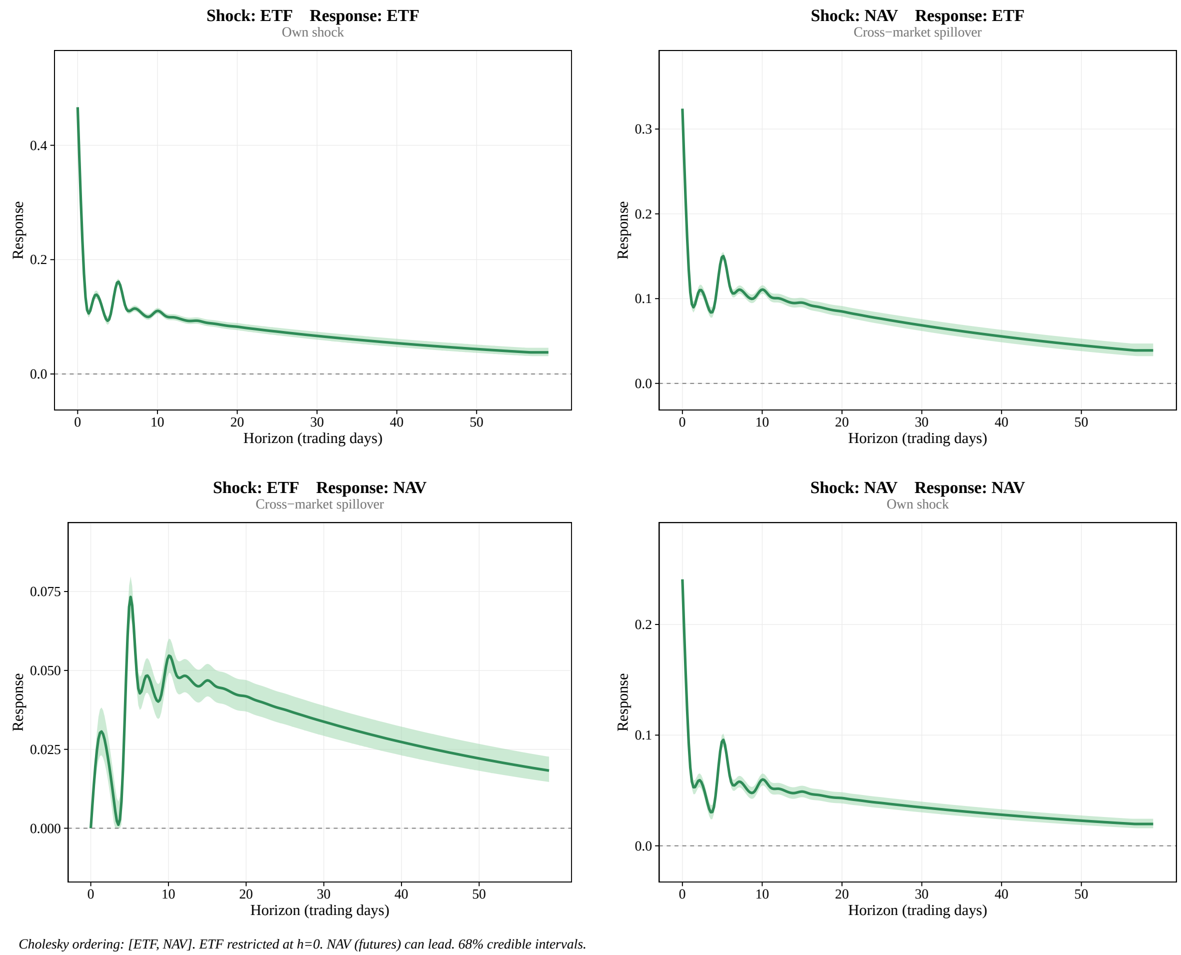


Figure 2.8: Impulse Response Functions (IRF) for UNG ETF and Natural Gas (NG) iNAV Volatility. Bayesian VAR impulse responses exhibiting the most elaborate structures with large responses in both directions, delayed peaks, and periodic behavior, substantiating bidirectional transmission involving numerous channels and horizons in natural gas markets.

Chapter 3

Returns and Volatility Around FOMC Announcements: A High-Frequency Analysis of Policy Tone and Novelty

3.1 Résumé

Nous décomposons les communiqués du FOMC en ton de politique monétaire (*hawkish*/*dovish*) et en nouveauté informationnelle (écart par rapport au communiqué précédent), puis estimons leurs effets sur les rendements et la volatilité à haute fréquence. À partir de données à une minute pour sept contrats à terme et 148 annonces du FOMC (2008–2025), nous montrons que le ton prédit les rendements directionnels — un virage accommodant d'un écart-type est associé à des gains boursiers atteignant environ 12 points de base en deux heures — tandis que la nouveauté prédit les variations de volatilité (l'interaction ton–nouveauté sur le VIX persiste de 5 à 120 minutes, $t = -5{,}06$). Le ton est associé à la volatilité réalisée pour six contrats sur sept ($p < 0{,}01$). Des tests placebo pré-annonce et cinq méthodes d'inférence indépendantes valident ces résultats, fondés sur un ensemble bi-modèle (MiniLM et BERT) avec sélection des références par analyse en composantes principales.

3.2 Abstract

We decompose FOMC statements into policy tone (hawkish/dovish) and informational novelty (departure from previous messaging) and estimate their effects on high-frequency asset returns and volatility. Using 1-minute data for 7 futures contracts across 148 FOMC events (2008–2025), we find that tone predicts directional returns (a one-standard-deviation dovish shift is associated with equity gains that build to about 12 basis points within two hours) while novelty predicts volatility changes (the stance–novelty interaction on VIX persists from 5 to 120 minutes, $t = -5.06$). Policy stance is associated with realized volatility changes in 6 of 7 contracts ($p < 0.01$). Pre-announcement placebo tests and five independent inference methods validate these results. To construct our measures, we train a dual-model ensemble (MiniLM and BERT) on FOMC communications with data-driven PCA-based reference selection.

Keywords: Monetary policy, FOMC announcements, high-frequency data, textual analysis, market volatility.
JEL Classification: E52, E58, G12, G14

3.3 Introduction

We study how the textual content of Federal Reserve communications affects asset returns and market volatility at high frequency. The Federal Open Market Committee (FOMC)—composed of the seven members of the Board of Governors, the president of the Federal Reserve Bank of New York, and four of the remaining eleven Reserve Bank presidents on a rotating basis—meets eight times per year to set the target federal funds rate. Since 1994, the Committee has released a public statement after each meeting; since 2011, the Chair has also held post-meeting press conferences. Other communications include meeting minutes (released three weeks later) and the Summary of Economic Projections with individual rate forecasts (the “dot plot”). Our analysis focuses exclusively on post-meeting statements, which are the first piece of information available to market participants at a precise, scheduled time. These statements affect U.S. asset prices, global capital flows, and exchange rates (Blinder et al., 2008; Campbell et al., 2012; Wongswan, 2009; Ehrmann et al., 2011).

Our central research question is: *Do different semantic dimensions of FOMC communications—specifically policy tone and informational novelty—affect financial market returns and volatility through distinct economic channels?* If tone captures the directional policy signal (hawkish vs. dovish) while novelty captures how much genuinely new information the statement contains, standard asset pricing theory predicts that these dimensions should have different effects. In the standard framework, asset prices equal expected future cash flows discounted at a rate that reflects policy expectations (Fama, 1970; Bernanke and Kuttner, 2005), so the directional content of the statement—tone—should move expected cash flows and discount rates, and hence returns. Volatility, by contrast, reflects uncertainty and disagreement about how to interpret new information (Veldkamp, 2011; Patton and Verardo, 2012), so the amount of new language—novelty—should affect information processing complexity and interpretive uncertainty, and hence volatility. We test this prediction using 1-minute price data for 7 futures contracts—the E-mini S&P 500 (ES), futures on the Chicago Board Options Exchange Volatility Index (VIX futures, VX), the 10-year (ZN) and 5-year (ZF) Treasury notes, the U.S. Dollar Index (DX), crude oil (CL), and gold (GC)—across 148 FOMC events from 2008 to 2025, estimating panel minute-level regressions, event-level regressions, and local projection impulse response functions (Jordà, 2005). Our identification strategy exploits the exogenous timing of announcements and is validated by pre-announcement placebo tests (Andersen et al., 2003, 2007).

To measure these dimensions, we construct two primary variables: (1) a tone measure capturing the hawkish-dovish spectrum, and (2) a novelty measure quantifying the semantic distance between consecutive FOMC communications. Concretely, we proceed in three steps. First, we convert each statement into a numerical vector (an *embedding*) using two language models of different sizes—MiniLM (which produces 384-dimensional vectors) and BERT (768-dimensional vectors)—and combine their outputs; using two architecturally distinct models guards against findings that are artifacts of a single model. Second, because generic language models are not trained on central bank language, we adapt both models to the corpus of FOMC statements in two stages: an unsupervised stage based on the Transformer-based Sequential Denoising Auto-Encoder (TSDAE) of Wang et al. (2021), which teaches the models Fed-specific vocabulary, followed by a supervised contrastive learning stage on pairs of statements with known relationships. Third, we locate each statement on the hawkish-dovish axis by measuring its proximity to reference statements at the two extremes of the spectrum. Whereas previous studies select these reference statements by hand—introducing researcher degrees of freedom, since results may depend on which dates the researcher chooses—we select them algorithmically using principal component analysis (PCA), making the construction of the semantic axes fully reproducible. Novelty is then measured as the distance between the embeddings of consecutive statements.

FOMC announcements are widely seen as the most important scheduled monetary policy events in global financial markets, as they contribute substantially to shaping expectations about the future path of the U.S. economy. Because markets are forward-looking, surprising announcements are typically followed by sizable adjustments across several markets (Bernanke and Kuttner, 2005; Blinder et al., 2008). Savor and Wilson (2014) show that a disproportionate share of the equity risk premium is earned on macroeconomic announcement days, with FOMC days being particularly important. Brusa et al. (2015, 2019) find that average stock returns and Sharpe ratios on FOMC days are 20–40 times higher than on non-announcement days. Lucca and Moench (2015) document a systematic pre-FOMC drift in equity prices. Nakamura and Steinsson (2018) and Jaroćinski and Karádi (2020) decompose FOMC surprises into policy shocks and information shocks and show that these have different effects on asset prices.

While earlier research focuses exclusively on the measurable policy surprise content of FOMC announcements, typically extracted from federal funds futures (Kuttner, 2001; Bernanke and Kuttner, 2005), a more recent strand of literature has begun to apply textual analysis to extract subtler information from the language itself (Hansen et al., 2018; Shapiro et al., 2022; Schmeling and Wagner, 2019; Gorodnichenko et al., 2023). This is an important step forward, given that the early contributions identify the effect of unexpected rate changes but ignore the qualitative content of the accompanying statement—forward guidance, risk assessments, and descriptions of economic conditions—which Gürkaynak et al. (2005) show can move long-term yields even when the rate decision is fully anticipated. Within this new strand of literature, however, most studies use daily data (Shapiro et al., 2022; Schmeling and Wagner, 2019; Eklund and Kim, 2024), which cannot separate the immediate market reaction to the statement from confounding information that arrives later in the day. Furthermore, the standard approach treats the statement as a one-dimensional object (hawkish vs. dovish), ignoring other dimensions such as how much the statement departs from prior language. In this paper, we address both limitations by using 1-minute data and by decomposing statements into two distinct dimensions: directional tone and informational novelty.

Textual analysis of financial communications has progressed from keyword counting (Bligh and Hess, 2008) and dictionary-based sentiment scoring (Loughran and McDonald, 2011) to transformer-based language models that capture context and word order (Gentzkow et al., 2019; Devlin et al., 2019; Araci, 2019). Within central bank communication, Hayo and Neuenkirch (2010); Apel and Grimaldi (2012) develop systematic measures of FOMC tone and show that these predict monetary policy expectations. Shapiro et al. (2022); Hansen et al. (2018) use natural language processing (NLP) classifiers to place FOMC statements on the hawkish-dovish spectrum, and Eklund and Kim (2024) find that hawkish sentiment measures predict subsequent realizations of the consumer price index (CPI). However, this literature treats FOMC communication as essentially one-dimensional—hawkish versus dovish—and has not examined whether other dimensions of the text, such as how much the language departs from the prior statement, have independent effects on asset prices.

We define informational novelty as the degree to which a new FOMC statement departs from the previous one, measured as the cosine distance between their sentence-transformer embeddings. It is useful to distinguish this concept from three related constructs. *Tone* or *sentiment* refers to the directional content of text (hawkish vs. dovish) and has been studied using dictionaries (Loughran and McDonald, 2011) and classifiers (Shapiro et al., 2022). *Economic policy uncertainty* (EPU), as measured by Baker et al. (2016), is a macro-level index based on newspaper coverage, tax code provisions, and forecaster disagreement. *Monetary policy surprises*, in the tradition of Kuttner (2001), are the unexpected component of the rate decision itself. Our novelty measure differs from all three: it captures changes in *how* information is presented between consecutive statements, independent of both directional stance and the rate decision. A statement can be highly novel yet neutral in tone (e.g., introducing new forward guidance language), or low in novelty yet accompanied by a large rate surprise.

Why should novelty matter for asset prices? Under rational expectations (Muth, 1961; Fama, 1970), only genuinely new information should move prices; repetitive content should already be priced in. But theory also suggests that novel and familiar information may affect returns and volatility differently. Returns respond to revisions in expected cash flows and discount rates—that is, to directional content. Volatility may respond more to disagreement among market participants about the interpretation of new language or to the processing cost of unfamiliar information (Patton and Verardo, 2012; Hu et al., 2013; Veldkamp, 2011). Consistent with this distinction, Manela and Moreira (2017) show that their news-based implied volatility index (NVIX) predicts future market volatility.

Our principal finding is that tone is associated with directional asset returns while novelty is associated with volatility, consistent with two distinct transmission channels. We make two main contributions. First, we develop a decomposition framework for FOMC statements based on a dual-model ensemble (MiniLM and BERT) fine-tuned on the FOMC corpus with data-driven PCA-based reference selection. Second, we provide the first high-frequency evidence on the differential effects of tone versus novelty, using local projection impulse response functions across 7 futures contracts and 148 FOMC events. In addition, we show that these two dimensions operate through distinct channels: tone is associated with fundamental valuations (equity returns building to about 12 basis points per standard deviation within two hours), while novelty is associated with uncertainty resolution (the stance–novelty interaction on VIX futures has a t-statistic of -5.06 and persists from 5 to 120 minutes). Our results are robust to five independent inference methods (Newey–West, wild bootstrap, clustered standard errors, quantile regression, and permutation tests) and are validated through pre-announcement placebo tests.

The remainder of this paper proceeds as follows. Section 3.4 describes data construction and processing. Section 3.5 presents our theoretical framework, propositions, and empirical methodology, including our identification strategy and econometric specifications. Section 3.6 reports our main empirical findings. Section 3.7 concludes with implications for monetary policy transmission and central bank communication strategy.

3.4 Data

3.4.1 FOMC Statement Collection

We collect 451 FOMC communications from the Federal Reserve Board website, spanning 2000–2025. After filtering to policy statements only, our textual corpus contains $N = 217$ FOMC statements, each averaging 3,840 characters. We use the broader 2000–2025 textual sample to train the embedding models and construct the semantic axes, which requires examples spanning multiple policy regimes. For the high-frequency event study, we focus on 148 FOMC events from 2008 to 2025. The 2008 start date reflects two considerations. First, reliable 1-minute futures data became consistently available following the shift to predominantly electronic trading on CME Globex in 2006–2007; pre-2008 high-frequency data for several of our contracts contains gaps and timing errors that would compromise our minute-level identification. Second, the 2008–2025 window encompasses sufficient variation in monetary policy regimes—the zero lower bound (2008–2015), quantitative easing programs, post-crisis normalization (2016–2019), the pandemic response (2020–2021), and the inflation-driven tightening cycle (2022–2025)—across three Fed Chairs (Bernanke, Yellen, Powell) to identify communication effects without sacrificing data quality (Swanson and Williams, 2014; Gürtler and Gürtler, 2010; Blinder et al., 2008). The broader 2000–2025 textual sample used for model fine-tuning starts in 2000 because the FOMC adopted its current practice of issuing post-meeting statements with substantive policy language beginning in 1999–2000; earlier statements were shorter and formulaic, providing too little variation for our NLP models to learn from.

All statements are obtained directly from Federal Reserve official releases archived on the Board of Governors website. We focus exclusively on post-meeting statements rather than meeting minutes, transcripts, or other Fed communications because these statements represent the information that is immediately available to market participants at precise announcement times (Gürkaynak et al., 2005). This temporal precision is necessary for high-frequency identification (Andersen et al., 2003; Rosa, 2013).

We preprocess statements following Gentzkow et al. (2019): we remove headers, footers, voting records, and administrative content, keeping only policy-relevant text. We account for structural breaks in statement format (e.g., the lengthening of statements post-2008) to avoid introducing spurious variation in our semantic measures (Hansen et al., 2018). We cross-reference all statements with the Fed's public archives to verify completeness.

3.4.2 Dual-Model Embedding Generation

We convert each FOMC statement into a numerical vector using two sentence transformer models, adapted to the FOMC corpus through the training procedure described in Section 3.5.1.

Model Architectures

1. **MiniLM** (`all-MiniLM-L6-v2`): 33M parameters, 384-dimensional embeddings.
2. **BERT** (`bert-base-uncased` with mean pooling): 110M parameters, 768-dimensional embeddings. A larger model that captures finer contextual distinctions in policy language (Devlin et al., 2019).

We train both models on the FOMC corpus using the two-stage procedure (TSDAE domain adaptation, then MNRL contrastive learning) described in Section 3.5.1. The training hyperparameters are: batch size 32 (MiniLM) / 16 (BERT), learning rate 1×10^{-5}, 5 (MiniLM) / 4 (BERT) supervised epochs with cosine warmup schedule (10% warmup ratio), and approximately 4,000 training examples per model from six pairing strategies.

Construction of Training Pairs

We construct training pairs using six strategies: (1) overlapping text segments from the same document; (2) consecutive sentence pairs within a document; (3) hawk–dove contrastive pairs classified through keyword banks; (4) temporal proximity pairs from consecutive FOMC meetings; (5) topic-based pairs sharing the same Fed topic (monetary policy, inflation, employment, etc.); and (6) key phrase paraphrases containing identical policy phrases (*"maintain the target range"*, *"decided to raise"*, etc.). The diversity of pairing strategies prevents the trained models from learning only a single sentiment dimension.

Data-Driven Reference Selection

Previous studies select reference dates manually or construct counterfactual statements by hand. We instead use the hybrid PCA–percentile approach described in Definition 3.4 to identify semantically extreme statements algorithmically. This procedure eliminates researcher degrees of freedom and produces more stable centroids (5 reference statements per pole versus 2 in typical manual selection). All semantic axes exceed the 0.04 minimum separation threshold. Ensemble separations are: total policy stance (0.302), risk assessment (0.274), policy communication (0.248), and economic priority (0.252).

We validate the embeddings in three ways. First, we compare embedding-based tone rankings with expert classifications and verify that known dovish and hawkish statements sort correctly. Second, we measure inter-model agreement: mean ensemble confidence is 0.837 for novelty and 0.562–0.806 for tone across axes. Third, we verify that the PCA axis assignments produce intuitive orderings.

3.4.3 High-Frequency Financial Data Construction

We use 1-minute OHLCV data for seven futures contracts spanning the major asset classes relevant to monetary policy transmission:

- **ES**: E-mini S&P 500 (equity benchmark)—responds to policy communications through discount rate effects, growth expectations, and risk premium adjustments.
- **VX**: VIX Futures (volatility)—captures implied volatility and uncertainty resolution, providing the most direct measure of information processing effects.
- **ZN**: 10-Year Treasury Note—reflects the interaction between policy expectations and term premium effects.
- **ZF**: 5-Year Treasury Note—primarily captures expectations about near-term policy rate changes.

- **DX**: Dollar Index—reflects relative monetary policy stances and international spillover effects.
- **CL**: Crude Oil WTI—reflects both inflation expectations and growth concerns influenced by monetary policy.
- **GC**: Gold—a safe-haven asset whose demand falls when accommodative communication triggers a rotation toward risk assets, and rises with policy-induced uncertainty.

The data span 148 FOMC events (approximately 8 per year from 2008 to 2025). For each event, we extract a ±120 minute window around the 14:00 ET announcement time. We interpolate the raw data to a regular 1-minute grid with forward-filling of short gaps (≤ 5 minutes) and returns recomputed on the regularized grid. We require at least 80% valid observations per rolling window.

For each instrument, we construct minute-by-minute log returns as:

$$r_{i,m}^{(j)} = \log\left(P_{i,m}^{(j)}\right) - \log\left(P_{i,m-1}^{(j)}\right) \tag{3.1}$$

where $P_{i,m}^{(j)}$ is the price of asset i at minute m relative to FOMC event j. We align timestamps across instruments and exchanges, handle market closures and trading halts, and filter outliers following standard microstructure procedures.

3.4.4 Event Window Specification and Temporal Alignment

Our primary analysis employs a 45-minute event window following each FOMC announcement, spanning from the announcement time ($t = 0$) to 45 minutes after ($t = +45$). We choose this window length for two reasons.

First, most of the announcement effect occurs within 15 minutes of the release, but equity responses to tone continue to develop through 45 minutes post-announcement. Second, press conferences typically begin 30 minutes after the statement, so a 45-minute window includes only the first 15 minutes of the press conference, limiting contamination.

We also implement several alternative window specifications as robustness checks. A narrow 15-minute post-announcement window provides very clean identification but with reduced statistical power. An extended 60-minute post-announcement window captures longer adjustment dynamics but with increased contamination risk, particularly from press conferences that typically begin 30 minutes after the statement release.

FOMC statements are typically released at 2:00 PM Eastern Time. We verify actual release times using Federal Reserve timestamps, news services, and market data providers, and correct any timing discrepancies before aligning with price data.

3.5 Methodology

This section describes our NLP framework for measuring tone and novelty (Section 3.5.1), presents the econometric specifications we use to estimate the effects of these measures on asset returns and volatility (Section 3.5.2), and states the testable hypotheses together with the specific coefficient restrictions that each one implies (Section 3.5.3).

3.5.1 Theoretical Framework

Conceptual Foundation

We decompose each FOMC statement into two dimensions: policy tone (hawkish vs. dovish) and informational novelty (distance from the previous statement). To illustrate why context-aware methods are needed for this decomposition, consider the difference between dictionary-based and machine learning approaches. A dictionary approach aggregates sentiment as:

$$S_{\text{dict}} = \frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{[\text{word}_i \in D_{\text{pos}}]} - \mathbb{1}_{[\text{word}_i \in D_{\text{neg}}]} \tag{3.2}$$

where D_{pos} and D_{neg} are predetermined word lists. The dictionary approach counts words one at a time, without regard to their position or neighbors. A transformer-based model instead produces $S_{\text{ML}} = f_\theta(x_1, x_2, \ldots, x_N; C)$, where f_θ is a neural network with learned parameters θ, and x_i are token representations that depend on surrounding words and document-level context C. To see why context matters, note that "the policy remains accommodative" and "the policy is no longer accommodative" contain the same keyword ("accommodative") but have opposite meanings; a dictionary assigns similar scores to both, while a transformer assigns opposite scores.

Our approach builds on three literatures. The information processing literature predicts that markets react more strongly to genuinely new information than to repetition (Grossman and Stiglitz, 1980; Fama, 1970). The central bank communication literature shows that policy statements move asset prices independently of rate decisions (Gürkaynak et al., 2005; Campbell et al., 2012). The market microstructure literature demonstrates that high-frequency data can isolate the real-time incorporation of new information into prices (Andersen et al., 2003; Rosa, 2013).

Definition 3.1 (FOMC Statement Representation)**.** Each FOMC statement F_t released at time t is mapped to a vector $\mathbf{f}_t \in \mathbb{R}^d$ in d-dimensional semantic space through domain-specific transformer embeddings. We employ two architecturally distinct models: MiniLM ($d = 384$) and BERT ($d = 768$), combined into a separation-weighted ensemble.

Each model maps a statement to a real-valued vector in which semantically similar texts (e.g., "accommodative" and "easy policy") are represented by nearby points, enabling cosine similarity comparisons and PCA. Unlike bag-of-words representations, where each coordinate is a word count and most are zero, these vectors are dense: every coordinate carries information. The output dimensionality ($d = 384$ for MiniLM, $d = 768$ for BERT) is fixed by the model architecture—MiniLM-L6-v2 has 6 layers and 384 hidden units per layer; BERT-base has 12 layers and 768 hidden units (Devlin et al., 2019)—and is not a hyperparameter we select. We adapt both architectures to the FOMC domain using a two-stage training procedure.

Stage 1: TSDAE Unsupervised Pre-Training. We first apply Transformer-based Sequential Denoising Auto-Encoder (TSDAE) (Wang et al., 2021) to adapt each model to the FOMC domain vocabulary:

$$\mathcal{L}_{\text{TSDAE}} = -\sum_{i=1}^{N} \log P\big(x_i \mid \tilde{x}_i; \theta_{\text{enc}}, \theta_{\text{dec}}\big), \tag{3.3}$$

where $\tilde{x}_i$ is a corrupted version of sentence x_i (token deletion with probability 0.6). This unsupervised stage teaches domain-specific vocabulary and structure without requiring labeled pairs, training for 3 epochs (MiniLM) or 2 epochs (BERT) with learning rate 3×10^{-5}.

Stage 2: Supervised Contrastive Learning with MNRL. We then fine-tune using Multiple Negatives Ranking Loss (MNRL):

$$\mathcal{L}_{\text{MNRL}} = -\frac{1}{B}\sum_{i=1}^{B} \log \frac{\exp\big(\text{sim}(a_i, p_i)/\tau\big)}{\sum_{j=1}^{B} \exp\big(\text{sim}(a_i, p_j)/\tau\big)}, \tag{3.4}$$

where (a_i, p_i) are positive pairs, $\text{sim}(\cdot, \cdot)$ denotes cosine similarity, τ is the temperature parameter, and B is the batch size. In-batch negatives provide contrastive signal without explicit negative sampling. We construct approximately 4,000 training pairs per model using the six pairing strategies described in Section 3.4.

Dual-Model Ensemble. The two models are combined using separation-weighted averaging:

$$\widehat{\text{Tone}}^{\text{ens}}_{t,k} = w_k^{\text{MiniLM}} \cdot \text{Tone}^{\text{MiniLM}}_{t,k} + w_k^{\text{BERT}} \cdot \text{Tone}^{\text{BERT}}_{t,k}, \tag{3.5}$$

where the weights are proportional to axis separation quality:

$$w_k^m = \frac{\text{Sep}_k^m}{\text{Sep}_k^{\text{MiniLM}} + \text{Sep}_k^{\text{BERT}}}, \quad m \in \{\text{MiniLM}, \text{BERT}\}. \tag{3.6}$$

For each statement, we measure inter-model confidence as:

$$\text{Conf}_{t,k} = \frac{1}{2}\big[\mathbb{1}[\text{sign}(\text{Tone}^{\text{M}}_{t,k}) = \text{sign}(\text{Tone}^{\text{B}}_{t,k})] + (1 - |\text{Tone}^{\text{M}}_{t,k} - \text{Tone}^{\text{B}}_{t,k}| / \max_j |\text{Tone}^{\text{M}}_{j,k} - \text{Tone}^{\text{B}}_{j,k}|)\big]. \tag{3.7}$$

Mean ensemble confidence is 0.837 for novelty and ranges from 0.562 to 0.806 for tone across axes. This dual-model approach reduces model-specific biases: MiniLM (33M parameters, 384 dimensions) provides computational efficiency and strong separation on policy stance, while BERT (110M parameters, 768 dimensions) captures richer contextual relationships. The ensemble achieves balanced performance across all semantic dimensions.

Semantic Dimensions

Definition 3.2 (Semantic Similarity Measure)**.** The semantic similarity between text vectors $\mathbf{u}$ and $\mathbf{v}$ is quantified using cosine similarity:

$$\text{sim}(\mathbf{u}, \mathbf{v}) = \frac{\mathbf{u} \cdot \mathbf{v}}{||\mathbf{u}||_2 \cdot ||\mathbf{v}||_2} = \frac{\sum_{i=1}^{d} u_i v_i}{\sqrt{\sum_{i=1}^{d} u_i^2} \cdot \sqrt{\sum_{i=1}^{d} v_i^2}} \tag{3.8}$$

The cosine similarity measure ranges from -1 to 1, where values approaching 1 indicate high semantic similarity, values near 0 indicate orthogonal or unrelated content, and values approaching -1 indicate semantic opposition. This metric is particularly well-suited for high-dimensional text embeddings because it captures angular distance between vectors while being invariant to vector magnitude differences that might arise from statement length variations.

Definition 3.3 (Statement Novelty)**.** The informational novelty of statement F_t relative to the immediately preceding statement is defined as:

$$\text{Novelty}_t = 1 - \text{sim}(\mathbf{f}_t, \mathbf{f}_{t-1}) \tag{3.9}$$

This novelty measure captures the degree to which current policy communications represent genuine departures from recent messaging patterns. Higher novelty values indicate statements containing substantively new informational content relative to recent communications, while lower values suggest continuity with established messaging. The measure is bounded between 0 and 2, where 0 indicates identical statements and 2 indicates maximally opposed statements.

To construct our policy tone measure, we develop a data-driven approach based on PCA-based semantic axis construction that replaces manually selected reference dates with an algorithmic procedure, eliminating researcher degrees of freedom.

Definition 3.4 (Data-Driven Semantic Axis Construction)**.** Let $\mathbf{E} \in \mathbb{R}^{N \times d}$ be the L2-normalized embedding matrix. We compute the first $K = 8$ principal components:

$$\mathbf{E} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\top + \boldsymbol{\epsilon}. \tag{3.10}$$

For each semantic dimension $k \in \{1, \ldots, 4\}$ (policy stance, risk assessment, policy communication, economic priority), we define keyword banks $\mathcal{K}_k^+$ (positive pole) and $\mathcal{K}_k^-$ (negative pole), and compute a keyword differential score:

$$\Delta_i^{(k)} = \sum_{w \in \mathcal{K}_k^+} \nVdash[w \in \text{text}_i] - \sum_{w \in \mathcal{K}_k^-} \nVdash[w \in \text{text}_i]. \tag{3.11}$$

We match each semantic axis to the PC maximizing the absolute Pearson correlation:

$$j^*(k) = \arg \max_{j \in \{1, \ldots, K\}} \left|\text{Corr}(\mathbf{v}_j, \boldsymbol{\Delta}^{(k)})\right|. \tag{3.12}$$

The reference selection uses a hybrid seed-and-percentile approach: (1) use 3 expert-selected dates per pole to define an initial axis direction $\hat{\mathbf{d}}_k = \bar{\mathbf{e}}_{\text{pos}} - \bar{\mathbf{e}}_{\text{neg}}$; (2) project all embeddings onto this direction: $s_i = \mathbf{e}_i^\top \hat{\mathbf{d}}_k / \|\hat{\mathbf{d}}_k\|$; (3) select $M = 5$ statements with highest and lowest projection scores; (4) iteratively refine by recomputing the axis direction from selected references and re-selecting. This yields 5 reference statements per pole (compared to 2 in manual selection), producing more stable centroids.

For each axis k, we measure discrimination quality using the separation metric:

$$\text{Sep}_k = 1 - \cos(\bar{\mathbf{c}}_k^+, \bar{\mathbf{c}}_k^-), \tag{3.13}$$

where $\bar{\mathbf{c}}_k^+, \bar{\mathbf{c}}_k^-$ are the L2-normalized positive and negative centroids. All axes achieve separations well above the 0.04 minimum threshold, with ensemble separations ranging from 0.248 to 0.302.

Definition 3.5 (Policy Tone)**.** The directional tone of statement F_t on semantic axis k is measured as:

$$\text{Tone}_{t,k} = \frac{\cos(\mathbf{f}_t, \bar{\mathbf{c}}_k^+) - \cos(\mathbf{f}_t, \bar{\mathbf{c}}_k^-)}{\text{Sep}_k}, \quad \text{clipped to } [-1, +1]. \tag{3.14}$$

This tone measure captures the relative proximity of actual statements to the positive (hawkish) and negative (dovish) reference centroids for each semantic axis, normalized by the axis separation. The primary variable used in regressions is the *total policy stance tone*. In the regression tables, the z-scored stance variable is oriented so that positive values indicate more accommodative (dovish) communication; positive coefficients therefore measure the response to a one-standard-deviation dovish shift. The composite MPS measure defined below retains the hawkish-positive orientation. Relative to counterfactual-based methods, this approach is reproducible, does not depend on particular date choices, and extends to semantic dimensions other than the hawkish-dovish spectrum.

Integrated Policy Stance Measurement

Definition 3.6 (Monetary Policy Stance)**.** The overall monetary policy stance is defined as:

$$\text{Stance}_t = \text{Novelty}_t \times \text{Tone}_t \tag{3.15}$$

Note on originality: Definitions 1–11 and Proposition 3.1 are our own contributions. They apply the standard sentence-embedding framework from the NLP literature to monetary policy communications with a new decomposition structure.

The multiplicative form means that stance is large only when a statement is both novel *and* directionally clear. We adopt a multiplicative rather than additive form for two reasons. First, it ensures that Stance is zero whenever either component is zero: repetitive content has no impact regardless of its tone, and novel but directionally neutral content has none either.

Second, it captures the reinforcing interaction between the two dimensions: novel information amplifies the impact of tone, and clear directional tone amplifies the impact of novelty.

In the following two definitions, $\mathbf{f}_t^D$ and $\mathbf{f}_t^H$ denote the embeddings of counterfactual statements located at the dovish and hawkish reference centroids of the policy stance axis, $\bar{\mathbf{c}}^-$ and $\bar{\mathbf{c}}^+$; they represent what the Fed would have released had it issued a maximally dovish or maximally hawkish statement at meeting t.

Definition 3.7 (Pure Dovish and Hawkish Stances)**.** The extreme stance scenarios are defined as:

$$\text{Stance}_t^{dove} = -\left(1 - \text{sim}(\mathbf{f}_t^D, \mathbf{f}_{t-1})\right) \tag{3.16}$$

$$\text{Stance}_t^{hawk} = 1 - \text{sim}(\mathbf{f}_t^H, \mathbf{f}_{t-1}) \tag{3.17}$$

Definition 3.8 (Weighted Policy Stance Representation)**.** The overall stance can be expressed as a convex combination of extreme stances:

$$\text{Stance}_t = w_t \cdot \text{Stance}_t^{dove} + (1 - w_t) \cdot \text{Stance}_t^{hawk} \tag{3.18}$$

where the dovish weight parameter w_t is computed (not estimated) as a deterministic function of the observed similarities:

$$w_t = \frac{1 - \text{sim}(\mathbf{f}_t^H, \mathbf{f}_{t-1}) - (1 - \text{sim}(\mathbf{f}_t, \mathbf{f}_{t-1})) \cdot \text{Tone}_t}{2 - \text{sim}(\mathbf{f}_t^D, \mathbf{f}_{t-1}) - \text{sim}(\mathbf{f}_t^H, \mathbf{f}_{t-1})} \tag{3.19}$$

This weighted representation provides an alternative interpretation of policy stance as a position along the spectrum between extreme dovish and hawkish alternatives. The weight parameter w_t is not a free parameter to be estimated; rather, it is algebraically derived from the observed textual similarities.

Market Surprise Identification

Definition 3.9 (Novelty Decomposition)**.** The total novelty can be decomposed into predictable and unpredictable components:

$$\text{Novelty}_t = \overline{\text{Novelty}}_{t|t-\Delta} + \varepsilon_t \tag{3.20}$$

where $\overline{\text{Novelty}}_{t|t-\Delta} \equiv \mathbb{E}_{t-\Delta}[\text{Novelty}_t]$ represents the expected novelty based on information available Δ periods before the announcement, and ε_t represents the novelty surprise component with $\mathbb{E}_{t-\Delta}[\varepsilon_t] = 0$ and $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2 < \infty$. We assume ε_t is covariance-stationary with $\text{Cov}(\varepsilon_t, \varepsilon_{t-j}) \to 0$ as $j \to \infty$.

Definition 3.10 (Expected Policy Stance)**.** Under rational expectations, the expected policy stance is:

$$\mathbb{E}_{t-\Delta}[\text{Stance}_t] = \mathbb{E}_{t-\Delta}\left[\text{Novelty}_t \times \text{Tone}_t\right] \tag{3.21}$$

Assumption 3.1 (Conditional Independence or Zero Covariance)**.** We assume either (i) conditional independence: $\text{Novelty}_t \perp \text{Tone}_t | \mathcal{I}_{t-\Delta}$, where $\mathcal{I}_{t-\Delta}$ is the information set at time $t - \Delta$, or (ii) the weaker condition of zero conditional covariance:

$$\text{Cov}_{t-\Delta}(\text{Novelty}_t, \text{Tone}_t) = 0 \tag{3.22}$$

This assumption is plausible because novelty captures whether the Fed changes its messaging structure (how information is presented), while tone captures the directional policy stance (hawkish/dovish). Empirically, the correlation between novelty and tone in our sample is 0.12 (Figure 3.2), supporting the assumption that these are largely independent dimensions. Under this assumption:

$$\mathbb{E}_{t-\Delta}[\text{Stance}_t] = \mathbb{E}_{t-\Delta}[\text{Novelty}_t] \times \mathbb{E}_{t-\Delta}[\text{Tone}_t] \tag{3.23}$$

Further assuming, for tractability, that tone is binary ($\text{Tone}_t \in \{-1, +1\}$, taking the dovish value -1 with probability $p_{t-\Delta}$), so that $\mathbb{E}_{t-\Delta}[\text{Tone}_t] = 1 - 2p_{t-\Delta}$:

$$\mathbb{E}_{t-\Delta}[\text{Stance}_t] = (1 - 2p_{t-\Delta}) \cdot \overline{\text{Novelty}}_{t|t-\Delta} \tag{3.24}$$

Definition 3.11 (Policy Stance Surprise)**.** The communication-based policy stance surprise is defined as:

$$\text{MPS}_t = \text{Stance}_t - \mathbb{E}_{t-\Delta}[\text{Stance}_t] \tag{3.25}$$

We use the abbreviation MPS (monetary policy surprise) for brevity, but emphasize that this measure is conceptually distinct from the traditional rate-based monetary policy surprise of Kuttner (2001), which measures the unexpected component of the federal funds rate decision using futures prices. Our MPS is a *communication* surprise—it captures unexpected variation in the textual content of FOMC statements, not in the quantitative rate decision. The two measures are complementary and potentially orthogonal: a meeting can produce a large rate surprise with a routine statement, or a novel and directional statement with a fully anticipated rate decision. Our measure captures the latter dimension, which has become increasingly important as forward guidance and qualitative communication have gained prominence in the monetary policy toolkit.

Proposition 3.1 (MPS Decomposition Structure)**.** *The monetary policy surprise admits the following decomposition:*

$$MPS_t = \overline{Novelty}_{t|t-\Delta}(Tone_t + 2p_{t-\Delta} - 1) + Tone_t \cdot \varepsilon_t \tag{3.26}$$

This decomposition reveals that policy surprises consist of two distinct components: (1) surprises arising from unexpected tone conditional on expected novelty, and (2) surprises arising from unexpected novelty, weighted by the actual tone of the communication. The proof of Proposition 3.1, together with the derivations underlying Definitions 3.7 and 3.8, is in Appendix A.

3.5.2 Econometric Framework

Panel Data Structure

The data are organized as a three-dimensional panel structure with observations indexed along three dimensions:

- Asset identifier $i \in \{1, 2, \ldots, I\}$ where I represents the total number of financial instruments
- Event identifier $j \in \{1, 2, \ldots, J\}$ corresponding to FOMC announcement dates where $J = 148$
- Relative minute $m \in \{-120, \ldots, +120\}$ measured from announcement time, with the primary post-announcement window covering $m \in \{0, \ldots, +45\}$

Identification Strategy

Our identification relies on the fact that FOMC statements are drafted and finalized before the announcement and released at a scheduled time (typically 2:00 PM ET). The text at $t = 0$ therefore cannot be influenced by market reactions at $t > 0$, satisfying the timing restriction for causal interpretation.

Our identification relies on three core conditions:

Identification Condition 1 (Exogenous Timing). The precise minute-by-minute timing of FOMC announcements is predetermined and exogenous to short-term market movements within our event windows.

Identification Condition 2 (Information Concentration). Within our 45-minute post-announcement event windows, FOMC statements represent the dominant source of monetary policy-relevant information, with minimal contamination from other systematic news sources.

Identification Condition 3 (Market Efficiency). Financial markets rapidly incorporate new information from FOMC statements into prices, enabling causal interpretation of immediate price movements following announcements.

We acknowledge two potential threats to this identification strategy. First, the Federal Reserve may adjust its communication in response to broader financial conditions observed before the meeting. We mitigate this concern by including event fixed effects δ_j that absorb all meeting-level heterogeneity. Second, FOMC announcements may coincide with other information releases (e.g., Summary of Economic Projections, press conference expectations). Our focus on the immediate post-announcement window limits contamination from subsequent press conference content, which typically begins 30 minutes after the statement release.

Dynamic Response Function Estimation

Our primary empirical specification estimates dynamic impulse response functions:

$$r_{i,m}^{(j)} = \alpha_i + \delta_j + \sum_{k=0}^{+45} \beta_k \cdot \nVdash_{[m=k]} \cdot \mathrm{MPS}_j + \varepsilon_{i,m,j} \tag{3.27}$$

where $r_{i,m}^{(j)}$ denotes the log return of asset i at minute m relative to monetary policy announcement j, α_i represents asset fixed effects, δ_j captures event fixed effects, and β_k measures the marginal impact of monetary policy surprises occurring exactly k minutes relative to the announcement.

Semantic Decomposition Specification

To examine the differential effects of policy tone versus informational novelty, we estimate:

$$r_{i,m}^{(j)} = \alpha_i + \delta_j + \sum_{k=0}^{+45} \left[\beta_k^{(T)} \cdot \nVdash_{[m=k]} \cdot \mathrm{Tone}_j + \beta_k^{(N)} \cdot \nVdash_{[m=k]} \cdot \mathrm{Novelty}_j\right] + \varepsilon_{i,m,j} \tag{3.28}$$

The tone coefficients $\{\beta_k^{(T)}\}$ measure the dynamic response to hawkish versus dovish content, while the novelty coefficients $\{\beta_k^{(N)}\}$ capture the response to informational content, independent of directional bias.

Rolling Realized Measures

In addition to return effects, we examine how policy communications affect market volatility using rolling realized measures computed from 1-minute log returns.

Realized Variance. We construct NA-tolerant rolling realized variance in K-minute windows:

$$\mathrm{RV}_K(t) = \frac{K}{n_{\text{valid}}} \sum_{s=t-K+1}^{t} r_s^2 \cdot \nVdash[r_s \text{ valid}], \tag{3.29}$$

where n_{valid} is the number of non-missing observations in the window. We require at least 80% valid observations.

Realized Beta. We compute rolling beta relative to the ES (S&P 500) benchmark:

$$\hat{\beta}_K(t) = \frac{\sum_{s=t-K+1}^{t} r_{i,s} \cdot r_{m,s}}{\sum_{s=t-K+1}^{t} r_{m,s}^2}, \tag{3.30}$$

where $r_{m,s}$ denotes the ES benchmark return. We use $K = 5$ for panel minute-level regressions and $K = 30$ for event-level analysis.

Panel Minute-Level Regressions

For each ticker, we estimate:

$$\log(\mathrm{RV}_{i,t}) = \alpha + \beta \cdot (\mathrm{Post}_t \times \mathrm{Semantic}_i) + \varepsilon_{i,t}, \tag{3.31}$$

where $\mathrm{Post}_t = \mathbb{1}[t > 0]$ indicates post-announcement minutes, and $\mathrm{Semantic}_i$ is either *Stance* or *Novelty* (z-scored). Standard errors are clustered by event date.

Event-Level Regressions

For each event, we compute pre/post changes:

$$\Delta \log(\mathrm{RV})_i = \overline{\log(\mathrm{RV})}_{\mathrm{post}} - \overline{\log(\mathrm{RV})}_{\mathrm{pre}}, \tag{3.32}$$

and regress on semantic measures:

$$\Delta Y_i = \alpha + \beta_1 \mathrm{Stance}_i + \beta_2 \mathrm{Novelty}_i + \beta_3 (\mathrm{Stance}_i \times \mathrm{Novelty}_i) + \varepsilon_i, \tag{3.33}$$

with Newey–West HAC standard errors. We estimate this specification for five dependent variables: $\Delta\mathrm{RV}$, $\Delta \log(\mathrm{RV})$, RV ratio, log(RV ratio), and $\Delta\beta$.

Local Projection Impulse Response Functions

Following Jordà (2005), for each horizon $h \in \{1, 2, 3, 5, 10, 15, 20, 30, 45, 60, 90, 120\}$ minutes post-announcement:

$$\mathrm{CumRet}_i(0 \to h) = \alpha + \beta_1^{(h)} \mathrm{Stance}_i + \beta_2^{(h)} \mathrm{Novelty}_i + \beta_3^{(h)} (\mathrm{Stance}_i \times \mathrm{Novelty}_i) + \varepsilon_i^{(h)}. \tag{3.34}$$

Abnormal cumulative returns are computed as $\mathrm{CAR}_i(h) = \mathrm{CumRet}_i(h) - \mathrm{CumRet}_i^{\mathrm{ES}}(h)$. This local projection approach is robust to misspecification of the data-generating process and allows horizon-specific inference without imposing a parametric impulse response shape. Because the post-meeting press conference typically begins 30 minutes after the statement release, we interpret responses at horizons beyond $h = 30$ as the joint effect of the statement and early press-conference communication.

Pre-Announcement Placebo Test

We estimate the same specification (3.34) for h minutes *before* the announcement:

$$\mathrm{CumRet}_i(-h \to 0) = \alpha + \beta_1^{(-h)} \mathrm{Stance}_i + \beta_2^{(-h)} \mathrm{Novelty}_i + \beta_3^{(-h)} (\mathrm{Stance}_i \times \mathrm{Novelty}_i) + \varepsilon_i^{(-h)}. \tag{3.35}$$

Under the null that announcement content is not anticipated, all $\beta^{(-h)}$ should be zero. This placebo test provides a direct validation of our event study design.

Statistical Inference and Robustness

We subject all event-level results to five alternative inference methods:

1. **Newey–West HAC standard errors**: Robust to heteroskedasticity and autocorrelation.
2. **Wild bootstrap** ($B = 1{,}999$, Rademacher weights): Robust to heteroskedasticity with improved finite-sample properties.
3. **Clustered standard errors** (HC1, by event date): Accounts for cross-sectional dependence within events.
4. **Quantile regression** ($\tau = 0.5$): Robust to outliers and heavy tails in the dependent variable.
5. **Permutation test** ($B = 4{,}999$): Gold standard—shuffles semantic labels across events preserving panel structure, providing exact p-values under the null.

All p-values in panel and IRF regressions are adjusted for multiple testing using the Benjamini–Hochberg (BH) procedure within each family of tests (per dependent variable and model type) (Benjamini and Hochberg, 1995). We use only standard significance levels: 1%, 5%, and 10%.

Sub-Period Stability. We examine coefficient stability across six Fed policy regimes: crisis and recovery (2008–2012), post-crisis normalization (2013–2015), pre-pandemic tightening (2016–2019), pandemic response (2020–2021), and inflation tightening (2022–2025).

Alternative Semantic Measures. The dual-model ensemble itself provides a built-in robustness check: by comparing results from MiniLM-only, BERT-only, and ensemble measures, we verify that findings are not driven by model-specific artifacts.

3.5.3 Hypotheses and Testable Predictions

We now state our hypotheses and, for each one, the coefficient restriction that operationalizes it in the specifications of Section 3.5.2. Recall the sign conventions: the z-scored stance regressor in equations (3.28)–(3.34) is dovish-positive, while the composite MPS in equation (3.27) is hawkish-positive. Table 3.1 summarizes the mapping from hypotheses to coefficients, equations, and the tables in which each test is reported.

Hypothesis 1 (Tone Effects on Returns).

- **H1a.** Dovish tone increases returns on risk assets (equities, commodities). *Test:* $\beta_k^{(T)} > 0$ in the semantic decomposition (3.28) and $\beta_1^{(h)} > 0$ in the local projection (3.34) for ES and CL.

- **H1b.** Dovish tone decreases returns on safe-haven assets, as accommodative policy triggers portfolio rebalancing away from safe havens toward risk assets. *Test:* $\beta_1^{(h)} < 0$ in (3.34) for GC, ZN, and ZF.

- **H1c.** Tone effects strengthen with the horizon as markets progressively process policy implications. *Test:* $|\beta_1^{(h)}|$ increasing in h in (3.34) for ES.

Hypothesis 2 (Novelty Effects on Volatility).

- **H2a.** High novelty increases market volatility through information processing complexity. *Test:* $\beta_2 > 0$ in the event-level regression (3.33) with ΔRV as the dependent variable, and $\beta > 0$ in the panel regression (3.31) with Semantic = Novelty, most directly for VX.

- **H2b.** Novelty effects dissipate rapidly as markets resolve uncertainty. *Test:* $|\beta_2^{(h)}|$ decreasing in h in (3.34) for VX.

- **H2c.** Novelty has minimal effects on directional returns. *Test:* $\beta_2^{(h)} \approx 0$ (statistically indistinguishable from zero) in (3.34) for the return contracts (ES, CL, GC, DX).

Hypothesis 3 (Policy Surprise Effects).

- **H3a.** Hawkish policy surprises reduce equity returns. *Test:* $\beta_0 < 0$ in the dynamic response function (3.27) for ES.

- **H3b.** Policy surprises increase implied volatility. *Test:* $\beta_0 > 0$ in (3.27) for VX.

- **H3c.** Surprise effects on volatility are most pronounced immediately after the announcement. *Test:* $|\beta_k|$ in (3.27) largest at $k = 0$ and decaying in k for VX.

Hypothesis 4 (Cross-Asset Patterns).

- **H4a.** Risk assets respond to tone in the same direction (risk-on/risk-off). *Test:* $\text{sign}\big(\beta_1^{(h)}\big)$ equal across ES and CL in (3.34), opposite for GC, ZN, ZF.

- **H4b.** Volatility responses to novelty are asset-specific. *Test:* β_2 and β_3 in (3.33) differ in magnitude and significance across the seven contracts.

Finally, the validity of the design itself is testable: under no anticipation of statement content, all pre-announcement coefficients in the placebo specification (3.35) should be zero, $\beta_1^{(-h)} = \beta_2^{(-h)} = \beta_3^{(-h)} = 0$ for all h.

Table 3.1: Mapping of Hypotheses to Coefficients, Equations, and Evidence

Hyp.	Prediction	Coefficient restriction (equation)	Evidence
H1a	Dovish tone ⇒ risk-asset returns ↑	$\beta_1^{(h)} > 0$ in (3.34), ES/CL	Table 3.9
H1b	Dovish tone ⇒ safe-haven returns ↓	$\beta_1^{(h)} < 0$ in (3.34), GC/ZN/ZF	Table 3.9
H1c	Tone effects build with horizon	$\lvert\beta_1^{(h)}\rvert$ increasing in h, ES	Table 3.9; Fig. 3.5
H2a	Novelty ⇒ volatility ↑	$\beta_2 > 0$ in (3.33), ΔRV	Tables 3.5, 3.6
H2b	Novelty effects decay quickly	$\lvert\beta_2^{(h)}\rvert$ decreasing in h, VX	Table 3.10
H2c	Novelty does not move returns	$\beta_2^{(h)} \approx 0$ in (3.34)	Table 3.10; Fig. B.7
H3a	Hawkish surprise ⇒ equity returns ↓	$\beta_0 < 0$ in (3.27), ES	Section 3.6.2
H3b	Surprise ⇒ implied volatility ↑	$\beta_0 > 0$ in (3.27), VX	Section 3.6.2
H3c	Surprise effects peak on impact	$\lvert\beta_k\rvert$ max at $k = 0$ in (3.27), VX	Section 3.6.2
H4a	Homogeneous return response, risk assets	$\mathrm{sign}(\beta_1^{(h)})$ equal, ES/CL in (3.34)	Table 3.9
H4b	Heterogeneous volatility responses	β_2, β_3 vary across contracts in (3.33)	Tables 3.6, 3.15
Placebo	No pre-announcement effects	$\beta^{(-h)} = 0$ in (3.35), all contracts	Fig. 3.8, B.9

3.6 Results

We organize our results as follows. We first describe the semantic variables (Section 3.6.1), then report baseline responses to the composite MPS (Section 3.6.2), the tone-novelty decomposition (Section 3.6.3), and the temporal dynamics of these effects (Section 3.6.4). Throughout, each hypothesis is evaluated against the coefficient restriction assigned to it in Table 3.1. The main finding is that tone and novelty load on different dependent variables: tone predicts directional returns, while novelty predicts volatility changes. The VIX model has the highest R^2 (0.78); equity and commodity models follow. Treasury results are weaker.

R^2 values for the return models are low (0.01–0.15), as is typical when predicting minute-level returns where microstructure noise dominates (Barndorff-Nielsen and Shephard, 2002). Low R^2 does not bias the coefficient estimates, but it means that FOMC communication accounts for a small share of total intra-day return variation. The VIX model is the exception: semantic variables explain up to 78% of announcement-window variation in VIX futures, consistent with the hypothesis that these communications operate primarily on uncertainty rather than on the level of returns.

3.6.1 Descriptive statistics

Table 3.2 reports descriptive statistics for 1-minute log returns on FOMC days. Table 3.3 reports summary statistics for the rolling realized measures. Figure 3.1 plots the average intraday volatility pattern on FOMC days, which spikes sharply at the 14:00 ET release.

Novelty has a mean of 0.065. The scale runs from 0 (a statement identical to the previous one) to 2 (maximally opposite), so a mean of 0.065 corresponds to an average cosine similarity of 0.935 between consecutive statements—most meetings produce only small changes to the prior template. The median is lower (0.025), so the distribution is right-skewed: most meetings produce minor textual changes, but a few produce large departures. The maximum (0.494) occurs during the 2008 crisis, when the Fed introduced zero-lower-bound language (Figure 3.3). Skewness is 2.45 and kurtosis is 8.16.

The raw tone measure (hawkish-positive) has a mean of 0.309 on a theoretical $[-1, +1]$ scale, where -1 is the dovish centroid, 0 is equidistant, and $+1$ is the hawkish centroid. The positive mean indicates that the average FOMC statement over 2008–2025 lies closer to the hawkish than to the dovish centroid. Because cosine similarities to both centroids are high, realized values occupy a narrow band within the theoretical scale; we therefore z-score tone in all regressions. Skewness is -0.78 and kurtosis is 6.96, driven by the most dovish statements in the sample, which mark its lower extremes (2008: -0.081; 2020: -0.062).

The MPS measure has mean 0.004 and kurtosis 12.18—the fat tails reflect occasional large surprises during economic stress. Jarque-Bera tests reject normality for all three measures ($p < 0.01$), and ARCH-LM tests indicate time-varying variance ($p < 0.05$), which motivates our use of robust standard errors and multiple inference methods throughout.

3.6.2 Baseline Asset Price Responses to Monetary Policy Surprises

We begin with the composite MPS, testing H3. The dependent variable is the 1-minute log return for each contract; the regressor is the z-scored MPS. A positive coefficient means that hawkish communication surprises are associated with higher returns.

Immediate Market Reactions

The equity market response to hawkish surprises is consistent with H3a (negative equity effect): the ES contract exhibits a -1.2 basis point response ($p < 0.05$). To put this magnitude in perspective, a one-standard-deviation hawkish surprise generates an immediate equity decline roughly equivalent to the average hourly return on a non-announcement day, concentrated in a single minute.

Among commodities, crude oil (CL) falls by 1.0 bps ($p < 0.05$), consistent with tighter policy reducing growth expectations and energy demand. Gold (GC) falls by 0.6 bps, as an inflation hedge becomes less attractive when policy tightens. The Dollar Index (DX) rises by 0.8 bps, consistent with higher expected rate differentials increasing dollar demand.

The VIX response is +11.2 bps ($p < 0.05$, H3b): a one-standard-deviation hawkish surprise is associated with an 11.2 bps increase in VIX futures, about 0.5% of the average VIX level. This is the largest immediate response in our sample. Treasury securities (ZN, ZF) show small, insignificant responses, possibly because our textual MPS is largely orthogonal to the rate expectations already embedded in Treasury futures.

Dynamic Response Evolution

The 45-minute post-announcement window shows that different asset classes adjust at different speeds:

Equities (ES). The initial decline (-1.2 bps at $t = 0$) deepens to -3.8 bps by $t = 15$ and -6.8 bps by $t = 45$. The cumulative response is 5.7 times the instantaneous reaction, so the first-minute price change captures only a fraction of the total adjustment.

VIX futures. The initial spike ($+11.2$ bps at $t = 0$) loses roughly half its magnitude by $t = 20$ and returns near its pre-announcement level by $t = 45$.

Treasuries (ZN, ZF). Responses become marginally significant only at $t = 30$–45, consistent with slower transmission to term premiums.

Commodities (CL, GC). Crude oil effects continue to grow through $t = 45$, consistent with the slower adjustment typical of physical commodity markets.

These patterns support H3c for volatility—the VIX response peaks on impact and decays—but equity and commodity responses continue to build, foreshadowing the gradual tone effects documented below. The speed of adjustment varies across assets, which motivates the asset-specific decomposition.

3.6.3 Semantic Decomposition: Tone vs. Novelty Effects

We now separate tone from novelty. The dependent variables are log returns (testing H1) and realized volatility changes (testing H2). The regressors are the z-scored tone and novelty measures, entered separately and jointly.

Tone Effects on Asset Returns (H1)

A one-standard-deviation dovish shift in tone is associated with ES gains that build steadily over the post-announcement window: 3.8 bps by $h = 30$, 6.8 bps by $h = 60$, and 12.0 bps by $h = 120$ (Table 3.9, H1a)—several days of average equity returns compressed into two hours. This is consistent with lower discount rates, higher growth expectations, and greater risk appetite.

Safe-haven assets move in the opposite direction: at $h = 15$, gold falls by 3.1 bps and 10-year Treasuries (ZN) by 1.1 bps following dovish tone (H1b), consistent with a "risk-on" rotation out of safe havens. The Dollar Index response is small and changes sign across horizons ($+1.4$ bps at $h = 15$, -4.0 bps at $h = 120$).

Tone effects grow stronger, not weaker, over the post-announcement window: the ES coefficient rises monotonically from 0.2 bps at $h = 5$ to 3.8 bps at $h = 30$ and 12.0 bps at $h = 120$ (Table 3.9, H1c). This gradual amplification—rather than the immediate level shift that a frictionless model would imply—is consistent with sequential portfolio adjustment by heterogeneous participants (algorithms, institutions, retail).

Novelty Effects on Volatility (H2)

Novelty loads on volatility, not on returns. The VIX response to a one-standard-deviation increase in novelty is -5.57 bps ($p < 0.05$), meaning that genuinely new language is associated with *lower* implied volatility. This sign is opposite to our ex ante prediction (H2a) but has a straightforward interpretation: a statement that departs substantially from the prior meeting's language sends a clearer signal about the Fed's current assessment, helping investors narrow the range of possible outcomes. A repetitive statement, by contrast, leaves open the question of whether the unchanged wording reflects genuine stability or simply a failure to update the language.

The VIX novelty coefficient is largest in the first 5 minutes and falls to near zero by $t = 30$ (H2b). Novelty coefficients on directional returns (ES, CL, GC, DX) are small and insignificant (H2c). The combination—novelty predicts volatility but not returns—separates our measure from standard "surprise" variables, which conflate direction and information content.

Panel Minute-Level Results

Tables 3.4–3.5 report panel regressions of log(RV) on the interaction of post-announcement indicators with semantic measures. The coefficient on Post×Stance is negative and significant at 1% for 6 of 7 contracts: ES (-5.31), VX (-3.28), ZN (-5.25), ZF (-4.16), CL (-5.87), and GC (-6.03). Only DX is insignificant ($+0.15$). The negative sign means that dovish stance is associated with lower post-announcement realized volatility—accommodative statements calm markets—while hawkish statements are followed by larger volatility increases, consistent with contractionary signals generating more repricing. By contrast, the Post×Novelty coefficients (Table 3.5) are smaller and significant only for the Treasury contracts (ZN: -1.81, $p < 0.01$; ZF: -1.22, $p < 0.10$).

Event-Level Results

Tables 3.6–3.8 report event-level regressions with the interaction specification $\Delta Y_i = \alpha + \beta_1 \text{Stance} + \beta_2 \text{Novelty} + \beta_3 (\text{Stance} \times \text{Novelty}) + \varepsilon_i$. The stance×novelty interaction is negative and significant for VIX (-7.96, $t = -3.12$, $p < 0.01$), ZF (-2.48, $t = -2.22$), ZN (-1.66, $t = -1.99$), and CL (-8.61, $p < 0.10$); Figure 3.4 summarizes the coefficient estimates as a heatmap. The interaction indicates that stance and novelty reinforce each other: the volatility decline associated with a dovish statement is larger when the statement also departs substantially from prior language, whereas a statement that repeats the previous meeting's wording has a smaller volatility effect regardless of its tone.

R^2 values are much higher for volatility than for returns: ΔRV regressions achieve 0.286 (ES), 0.268 (CL), and 0.245 (ZF), versus below 0.05 for most return regressions. The semantic variables explain more of the variation in volatility (which depends on information content) than in returns (which depend on the direction of the surprise in a noisy environment).

Local Projection Impulse Response Functions

Tables 3.9–3.14 and Figures 3.6–3.7 report local projection estimates. The VIX response to stance is negative at every horizon, statistically strongest between 5 and 30 minutes (-14.74 bps at $h = 5$, -22.03 at $h = 30$), and attenuates thereafter (-14.27 at $h = 120$, with much wider standard errors); the profile is hump-shaped, peaking near 30 minutes. The stance$\times$novelty interaction on abnormal VIX returns (Figure 3.7) is significant from $h = 5$ through $h = 120$, with the t-statistic reaching -5.06.

The ES coefficient on stance grows monotonically from near zero at $h = 5$ to 12.0 bps at $h = 120$ (standard errors also grow). A frictionless model would predict an immediate level shift; the gradual increase we observe is more consistent with sequential adjustment by heterogeneous investors.

Cross-Asset Patterns (H4)

Risk assets (ES, CL) both rise after dovish tone, with similar timing (H4a). Safe havens (GC, ZN, ZF) fall, consistent with risk-on/risk-off. But the volatility responses differ across assets (H4b): VIX responds most to the stance$\times$novelty interaction, Treasury volatility responds mainly to stance alone, and commodity volatility responds to both. The return homogeneity and volatility heterogeneity suggest that the two channels operate with different relative strength across asset classes.

Pre-Announcement Placebo Tests

Figures 3.8–B.9 report pre-announcement placebo tests: we estimate the same IRF specification for cumulative returns from $-h$ to 0. If statement content is not anticipated, all pre-announcement coefficients should be zero. None is significant at 10% across any ticker, horizon, or semantic measure. This rules out information leakage and pre-existing trends.

Multi-Method Robustness

Tables 3.15–3.16 report five alternative inference methods for every event-level coefficient. We call a result "robust" if at least 3 of 5 methods give $p < 0.10$. The most robust results are: stance on $\Delta \log \mathrm{RV}$ for ES, CL, GC, ZF, and ZN (each #Sig = 3); stance on ΔRV for ES, CL, and DX (each #Sig = 3); and novelty on ΔRV for ES and CL (#Sig = 4). Figure 3.9 displays these counts. Sub-period analysis (Figure B.10) shows consistent coefficient signs across six Fed regimes, though magnitudes vary.

With 7 assets and multiple dependent variables, we test many coefficients; results with #Sig = 1 should be treated with caution. We rely on the Benjamini–Hochberg correction and the robustness counts to limit false discovery.

3.6.4 Economic Interpretation: Temporal Dynamics and Half-Lives

We summarize the speed of adjustment using half-lives: the time for an initial effect to fall to half its peak. For exponential decay, $t_{1/2} = \ln(2)/\lambda$.

The VIX response (+11.2 bps on impact) falls to half its peak within 15–20 minutes; reading the decay as approximately exponential gives a half-life of about 17.5 minutes. Once the statement text is known, uncertainty about the Fed's message dissipates quickly.

Equity returns show the opposite pattern. The ES response to dovish tone *grows* steadily over the post-announcement window, roughly tripling between $h = 30$ (3.8 bps) and $h = 120$ (12.0 bps). Fundamental repricing—revising expected cash flows and discount rates—takes longer than uncertainty resolution, consistent with gradual diffusion across heterogeneous investors.

The two patterns together—fast volatility decay and slow return amplification—are hard to reconcile with a single channel. The informational dimension (is the statement new?) resolves quickly; the directional dimension (is it hawkish or dovish?) takes longer to price in.

3.7 Conclusion

3.7.1 Summary of Findings

We summarize our findings under two headings: what we learn about the economics of monetary policy transmission, and what we contribute methodologically.

Tone predicts directional returns. Dovish tone is associated with higher equity returns that build to about 12 bps over two hours, and with lower safe-haven returns (gold: −3.1 bps; Treasuries: −1.1 bps at 15 minutes), supporting H1a–b. This is consistent with the standard transmission mechanism in Bernanke and Kuttner (2005) and Gürkaynak et al. (2005)—accommodative expectations lower discount rates and shift portfolios toward risk assets—but we show that the channel operates through the *qualitative* content of the statement, not only through the rate decision.

Tone effects grow stronger over time: the ES coefficient rises from 0.2 bps at $h = 5$ to 12.0 bps at $h = 120$ (H1c). Rosa (2013) documents continued price adjustment over hours after FOMC releases; our minute-level data show that this amplification is consistent with heterogeneous processing speeds across market participants (Veldkamp, 2011).

Novelty predicts lower volatility, not higher. Novelty is associated with *reduced* implied volatility (VIX: -5.57 bps, $p < 0.05$; interaction $t = -5.06$), which reverses the sign we predicted under H2a. Prior work on disagreement-driven volatility (Patton and Verardo, 2012) and news-based uncertainty (Manela and Moreira, 2017) would predict the opposite. The difference, we believe, is that in the specific setting of FOMC statements, a departure from prior language acts as a *signal of clarity*: the Fed is actively updating its message, which narrows the set of plausible policy paths. A repetitive statement, conversely, leaves open whether the Fed's views have changed but the language has not been updated. Novelty effects dissipate in under 10 minutes (H2b) and do not predict returns (H2c), confirming that this channel operates on volatility rather than on prices.

The composite surprise replicates known patterns. Hawkish surprises are associated with 1–7 bps lower equity returns (H3a) and 11.2 bps higher VIX (H3b), with volatility effects concentrated in the first minutes (H3c). That our textual MPS produces results consistent with the rate-surprise literature (Kuttner, 2001; Bernanke and Kuttner, 2005; Gürkaynak et al., 2005; Savor and Wilson, 2014) reassures us that the NLP pipeline captures economically relevant variation.

Cross-asset patterns differ between returns and volatility. Risk assets (ES, CL) respond to tone with the same sign and similar timing (H4a). Safe havens (GC, ZN, ZF) move in the opposite direction. But volatility responses vary: VIX loads on the interaction, Treasury volatility on stance alone, commodity volatility on both. This pattern is consistent with Fleming and Remolona (1999) and Balduzzi et al. (2001), who document heterogeneous announcement effects across asset classes.

Volatility resolves fast; returns adjust slowly. The VIX half-life is about 17.5 minutes; the ES tone effect continues to build for two hours after the announcement. A single-channel model cannot produce both patterns simultaneously. The fast channel (novelty $\rightarrow$ volatility) reflects the resolution of uncertainty about what the Fed will say. The slow channel (tone $\rightarrow$ returns) reflects the time needed to revise discount rates, growth expectations, and portfolio allocations across a heterogeneous investor base.

3.7.2 Methodological Contributions

On the measurement side, we make three contributions. First, using two architecturally different models (MiniLM, 33M parameters; BERT, 110M parameters) and comparing their outputs reduces the risk that any finding is an artifact of one particular architecture. The inter-model confidence (mean 0.837 for novelty) flags statements where the two models disagree. Second, two-stage domain adaptation (TSDAE then MNRL) yields axis separations of 0.248–0.302, well above the 0.04 minimum. Third, PCA-based reference selection replaces subjective date choices with an algorithm, removing a source of researcher discretion.

Working at the minute level (148 events $\times$ $\pm$120 minutes) matters for two reasons: it allows causal identification from the pre-determined release time (validated by the placebo tests), and it reveals the fast-volatility / slow-return asymmetry that daily data would average away.

3.7.3 Implications for Policy and Communication Strategy

Three findings have implications for how the Fed drafts its statements.

First, novelty reduces volatility. The conventional view is that central banks should change their language gradually to avoid "surprising" markets. Our estimates say the opposite: substantive changes in wording are associated with *lower* implied volatility, presumably because a clear departure from the prior statement narrows the range of plausible interpretations. When the Fed needs to signal a regime change—a new framework, a pivot from tightening to easing—a decisive rewrite of the statement may be less destabilizing than incremental edits.

Second, tone has persistent valuation effects. Because the equity response to tone grows over 2 hours rather than being absorbed immediately, the hawkish-dovish framing of the statement has real wealth consequences: a one-standard-deviation dovish shift is associated with up to 12 bps of equity returns, which, on a market capitalization of roughly \$40 trillion, is economically meaningful.

Third, the interaction matters. The stance$\times$novelty interaction on VIX ($t = -5.06$, persistent from 5 to 120 minutes) implies that a directionally clear statement packaged in new language reduces uncertainty more than the same message delivered in boilerplate.

3.7.4 Limitations

We note several limitations.

Sample period. Our 2008–2025 sample is dominated by extraordinary monetary policy regimes: the zero lower bound (2008–2015), quantitative easing, and the post-pandemic tightening cycle. While our sub-period stability analysis (Figure B.10) shows that coefficient signs are generally consistent across six Fed policy regimes, we cannot rule out that the magnitude of communication effects differs in more "normal" policy environments. In particular, the ZLB period may overstate the importance of qualitative communication (since rate decisions were constrained, statements became the primary policy instrument), while the recent tightening cycle may understate novelty effects (since rate hikes were widely anticipated, reducing the scope for genuine communication surprises).

Measurement error. Our NLP-based tone and novelty measures are subject to measurement error whose properties are difficult to characterize fully. Classical measurement error would attenuate our coefficient estimates toward zero (creating an "errors-in-variables" bias), making our significant findings conservative. However, if measurement error is correlated with meeting characteristics (e.g., if our models systematically misclassify statements from certain Fed Chairs or policy regimes), the bias could go in either direction. The inter-model agreement diagnostics (mean confidence 0.837 for novelty, 0.562–0.806 for tone) help identify statements where measurement may be less reliable, but they do not fully resolve this concern.

Multiple testing. With 7 assets, 5 dependent variables, and multiple inference methods, we test a large number of coefficients. We address this through Benjamini–Hochberg FDR correction within each family of tests and by reporting multi-method robustness counts. Nevertheless, some individually significant results—particularly those with #Sig = 1 in the robustness tables—should be interpreted cautiously as potentially reflecting false discovery.

Identification threats. Our identification strategy exploits the pre-determined timing and content of FOMC statements, validated by pre-announcement placebo tests. However, we cannot fully rule out confounding from simultaneous information releases. FOMC announcements sometimes coincide with the Summary of Economic Projections ("dot plot"), which provides quantitative rate path forecasts that may interact with our textual measures. Similarly, market expectations about the subsequent press conference (beginning 30 minutes after the statement) may influence post-announcement price dynamics within our event window. These concurrent information flows are not fully separable from the statement text itself.

Generalizability. We study U.S. futures markets around FOMC statements only. Whether the same tone-return and novelty-volatility patterns hold for other central banks (ECB, BOJ, BOE), other Fed communications (minutes, speeches), or other asset classes (corporate bonds, emerging-market equities) is an open question.

3.7.5 Future Directions

Several extensions seem natural.

More semantic dimensions. We measure tone and novelty; FOMC statements also vary in uncertainty language, temporal focus (forward- vs. backward-looking), and internal consensus (voting dissents). The PCA axis construction can accommodate additional dimensions.

International markets. FOMC announcements move global asset prices (Wongswan, 2009; Ehrmann et al., 2011). Applying the tone-novelty decomposition to international data would show whether both channels transmit across borders or whether one dominates.

Structural models. Our evidence is reduced-form. A model with heterogeneous agents and differential processing of directional versus informational content could rationalize the fast-volatility / slow-return asymmetry and generate further predictions.

Joint analysis with rate surprises. Our textual MPS and the rate-based surprise of Kuttner (2001) measure different things. Estimating both jointly would reveal whether they contain complementary information and whether the information-versus-policy decomposition of Jaroćinski and Karádi (2020); Nakamura and Steinsson (2018) maps onto our tone-versus-novelty decomposition.

3.8 Tables

Table 3.2: Descriptive Statistics: 1-Minute Log Returns on FOMC Days

Ticker	N	Mean	SD	Skew	Kurt	JB	Min	Median	Max	AC(1)
E-mini S&P 500	42,180	-0.030	3.70	0.96	208.0	73858777***	-118.7	0.000	148.0	-0.022
VIX Futures	37,050	0.072	14.12	0.02	77.9	8665376***	-466.1	0.000	419.8	-0.077
10Y T-Note	42,180	0.009	1.07	1.21	44.4	3027292***	-13.8	0.000	30.8	-0.103
5Y T-Note	42,180	0.004	0.68	-4.70	444.1	342093523***	-43.8	0.000	17.6	-0.059
Dollar Index	42,180	-0.000	1.69	-1.74	1508.6	3984214777***	-119.2	0.000	115.4	0.194
Crude Oil WTI	42,180	-0.006	5.14	-0.64	166.7	47098526***	-191.8	0.000	143.9	-0.062
Gold	41,895	0.001	2.91	0.24	121.3	24424589***	-91.5	0.000	108.8	-0.017

Notes: All statistics in basis points except N, AC(1), and JB. Sample includes all 1-minute observations on FOMC event days ± 1 day with grid regularization and forward-filling of short gaps. Skewness and excess kurtosis are sample moments. JB is the Jarque–Bera test statistic. AC(1) denotes the first-order autocorrelation coefficient. Negative AC(1) values reflect bid–ask bounce effects typical of high-frequency data (Andersen and Bollerslev, 1997).
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 3.3: Descriptive Statistics: Rolling Realized Measures (5-Minute Window)

	N	Realized Volatility (bps)					log(RV)		Realized Beta vs. ES		
Contract		Mean	SD	P5	P50	P95	Mean	SD	Mean	SD	P50
E-mini S&P 500	35,668	3.43	5.68	0.00	2.41	10.38	-11.78	6.55	1.000	0.000	1.000
10Y T-Note	35,668	1.41	1.63	0.00	1.31	3.81	-14.11	7.09	-0.018	0.395	0.000
5Y T-Note	35,668	0.79	1.15	0.00	0.66	2.36	-14.77	6.86	-0.007	0.225	0.000
Dollar Index	35,668	0.78	3.45	0.00	0.00	3.53	-19.32	6.34	-0.010	0.482	0.000
Crude Oil WTI	35,668	5.86	9.47	0.00	3.97	17.90	-11.24	6.67	0.184	1.472	0.006
Gold	35,427	3.64	4.64	0.00	2.73	10.51	-11.29	6.30	0.044	0.958	0.000
VIX Futures	31,330	11.16	23.33	0.00	0.00	52.40	-18.31	7.74	-0.714	4.451	0.000

Notes: Statistics computed across all minute-level observations within ±30 minutes of FOMC announcements ($K = 5$ min rolling window, NA-tolerant with up to 20% missing data). RV is the NA-tolerant rolling realized variance defined in Section 3.5, displayed in square-root (volatility) units in basis points; log(RV) is computed on the realized variance in raw decimal-return units, so its level is not directly comparable to the bps columns. $\beta_t^{\text{real}} = \text{RCov}(r_i, r_{\text{ES}})/\text{RVar}(r_{\text{ES}})$. ES beta is 1.000 by construction. log(RV) is near-Gaussian, validating its use as a regression dependent variable.

Table 3.4: Panel Regressions: log(RV) on Post × Stance (5-min, ±30 min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-14.8474***	-19.9268***	-17.7088***	-17.9767***	-22.6514***	-14.9591***	-15.0663***
	(0.4794)	(0.4462)	(0.4275)	(0.4466)	(0.1175)	(0.4704)	(0.4845)
Post × Stance	-5.3084***	-3.2805***	-5.2530***	-4.1583***	0.1481	-5.8682***	-6.0348***
	(0.8271)	(0.7803)	(0.6490)	(0.7052)	(0.2836)	(0.8526)	(0.8609)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.134	0.047	0.133	0.091	0.001	0.149	0.166
Adj. R^2	0.134	0.047	0.132	0.091	0.001	0.149	0.166

Notes: Dependent variable: $\log(\mathrm{RV}_t(5))$. Post×Stance = $\mathbf{1}[t > 0] \times$ Stance (z-scored). Post×Novelty = $\mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ±30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 3.5: Panel Regressions: log(RV) on Post × Novelty (5-min, ±30 min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-13.4894***	-18.9261***	-16.4259***	-16.9510***	-22.6505***	-13.4467***	-13.4834***
	(0.2892)	(0.3813)	(0.3396)	(0.3423)	(0.1015)	(0.2887)	(0.2946)
Post × Novelty	-0.5169	0.3020	-1.8142***	-1.2188*	0.8439	-0.3328	-0.4821
	(0.5996)	(0.8881)	(0.5683)	(0.5493)	(0.4881)	(0.6298)	(0.6316)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.003	0.001	0.031	0.015	0.069	0.001	0.002
Adj. R^2	0.002	0.001	0.031	0.015	0.068	0.001	0.002

Notes: Dependent variable: $\log(\mathrm{RV}_t(5))$. Post×Stance $= \mathbf{1}[t > 0] \times$ Stance (z-scored). Post×Novelty $= \mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ±30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 3.6: Rolling Delta RV Regressions (30-min Window, Interaction Model)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	5.54***	26.00***	2.23***	1.65***	-0.40	12.32***	9.96***
	(1.67)	(6.51)	(0.64)	(0.49)	(0.62)	(2.02)	(1.51)
Stance	-4.94**	-2.80	-1.50*	-1.30*	-1.65	-7.78**	-1.05
	(1.53)	(6.03)	(0.71)	(0.64)	(0.90)	(2.65)	(1.92)
Novelty	6.82*	3.42	0.70	0.94	1.15	8.20*	3.34*
	(3.42)	(5.98)	(0.81)	(0.83)	(1.09)	(3.63)	(1.35)
Stance $\times$ Novelty	-5.32	-7.96**	-1.66*	-2.48*	-1.20	-8.61*	-4.68
	(3.15)	(2.55)	(0.83)	(1.12)	(1.39)	(3.60)	(3.13)
N	148	130	148	148	148	148	147
R^2	0.286	0.054	0.131	0.245	0.137	0.268	0.128
Adj. R^2	0.271	0.032	0.113	0.229	0.119	0.253	0.109

Notes: Dependent variable: Delta RV (bps) (Rolling 30-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Coefficients $\times 10,000$. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 3.7: Rolling Delta log(RV) Regressions (30-min Window, Interaction Model)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	6.1744***	6.8238***	7.1287***	6.9380***	-0.3664	8.4210***	8.4805***
	(1.0374)	(1.6303)	(1.0670)	(1.0305)	(0.3671)	(0.8868)	(0.8682)
Stance	-2.9143	-0.8191	-2.5521	-2.3930	-0.9082	-2.1849	-1.7794
	(1.0198)	(1.8314)	(1.1059)	(1.0661)	(0.5066)	(1.0687)	(1.0675)
Novelty	0.3441	-0.2304	-0.7730	-0.6468	0.5076	0.8902	0.7810
	(0.6037)	(0.9547)	(0.5429)	(0.5328)	(0.5352)	(0.4145)	(0.3970)
Stance × Novelty	-0.6767	-1.5107	-1.0145	-1.0850	-0.3540	-0.7406	-0.6972
	(0.6652)	(0.7161)	(0.5278)	(0.5025)	(0.7771)	(0.6279)	(0.6369)
N	148	130	148	148	148	148	147
R^2	0.064	0.018	0.075	0.071	0.055	0.052	0.041
Adj. R^2	0.044	-0.005	0.056	0.051	0.036	0.033	0.021

Notes: Dependent variable: $\Delta \log(\text{RV})$ (Rolling 30-min). Coefficients are in log-units; a coefficient of -2.91 on Stance for ES means that a one-standard-deviation hawkish shift is associated with a 2.91 log-unit decrease in realized volatility. All regressors z-scored. Clustered SE in parentheses. Stars: BH-adjusted p-values. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 3.8: Rolling Realized Beta Change Regressions (30-min Window)

	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-0.2364	0.0098	-0.0025	0.0058	-0.0933	0.1275*
	(0.1651)	(0.0192)	(0.0134)	(0.0074)	(0.0650)	(0.0758)
Stance	-0.0190	0.0118	0.0195	0.0137	-0.0008	0.1244
	(0.2653)	(0.0198)	(0.0158)	(0.0093)	(0.0891)	(0.0869)
Novelty	0.1096	-0.0096	-0.0064	-0.0086	-0.1290	-0.0539
	(0.0917)	(0.0102)	(0.0073)	(0.0075)	(0.1020)	(0.0344)
Stance × Novelty	-0.0818	0.0064	0.0074	0.0011	-0.1772	-0.0852
	(0.1018)	(0.0118)	(0.0074)	(0.0080)	(0.2106)	(0.0755)
N	71	88	88	88	88	87
R^2	0.009	0.003	0.008	0.045	0.064	0.047
Adj. R^2	-0.036	-0.033	-0.027	0.010	0.031	0.013

Notes: Dependent variable: $\Delta\hat{\beta}$ (change in rolling realized beta relative to ES, 30-min window). Coefficients are in beta units; a coefficient of 0.12 means that a one-standard-deviation increase in the regressor is associated with a 0.12 increase in co-movement with the S&P 500. All regressors z-scored. Clustered SE in parentheses. Stars: BH-adjusted p-values. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table 3.9: Price IRF: Policy Stance on Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	0.23	1.98	3.76	6.82	11.98
	(0.92)	(2.11)	(2.76)	(3.31)	(6.76)
VIX Futures	-14.74	-19.11	-22.03	-17.73	-14.27
	(6.26)	(8.33)	(9.47)	(14.62)	(31.57)
10Y T-Note	-0.61	-1.14	-0.90	-1.47	-3.88
	(0.43)	(0.85)	(0.93)	(0.92)	(2.20)
5Y T-Note	0.89	0.63	0.90	0.79	-0.71
	(0.34)	(0.35)	(0.55)	(0.51)	(1.34)
Dollar Index	1.01	1.43	1.05	-0.45	-4.01
	(0.43)	(0.61)	(0.45)	(0.89)	(2.14)
Crude Oil WTI	-2.10	-0.81	4.75	3.06	9.32
	(4.28)	(6.28)	(8.46)	(10.23)	(9.69)
Gold	-1.47	-3.08	0.51	0.62	1.31
	(1.30)	(1.42)	(2.04)	(2.46)	(3.96)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1 \text{Stance} + \beta_2 \text{Novelty} + \beta_3 \text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Policy Stance reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. Pre-announcement placebo test passed: no significant pre-event coefficients.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 3.10: Price IRF: Novelty on Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	-0.99	-1.42	-0.65	-0.61	-3.91
	(0.56)	(1.70)	(1.34)	(1.82)	(2.25)
VIX Futures	10.24	4.84	-2.55	-6.02	6.22
	(9.19)	(9.31)	(7.62)	(5.36)	(8.33)
10Y T-Note	0.82	1.17	1.52	1.87	3.70
	(0.60)	(0.97)	(0.91)	(1.09)	(2.15)
5Y T-Note	-0.61	-0.16	0.16	-0.00	1.08
	(0.41)	(0.10)	(0.17)	(0.27)	(0.47)
Dollar Index	0.41	0.30	0.61	0.94	0.60
	(1.34)	(1.74)	(1.51)	(1.62)	(2.22)
Crude Oil WTI	-0.52	-0.82	-3.78	-2.03	-1.29
	(3.54)	(4.29)	(7.83)	(6.83)	(6.20)
Gold	-1.30	0.84	0.55	-1.13	2.00
	(1.42)	(2.62)	(2.33)	(1.89)	(2.36)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1\text{Stance} + \beta_2\text{Novelty} + \beta_3\text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Novelty reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. Pre-announcement placebo test passed: no significant pre-event coefficients.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 3.11: Price IRF: Stance × Novelty on Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	2.78	6.91	8.27	9.47	6.35
	(1.36)	(4.14)	(4.39)	(4.61)	(4.49)
VIX Futures	-28.02**	-47.97**	-61.74***	-51.80***	-44.06*
	(8.62)	(12.75)	(15.54)	(10.66)	(14.53)
10Y T-Note	-1.76	-3.53	-3.68	-4.77	-5.70
	(0.85)	(1.68)	(1.74)	(1.88)	(2.84)
5Y T-Note	1.15	-0.03	0.24	-0.19	0.63
	(0.57)	(0.10)	(0.16)	(0.28)	(0.46)
Dollar Index	2.96	4.05	3.16	3.32	2.83
	(1.15)	(1.53)	(1.27)	(1.41)	(2.21)
Crude Oil WTI	7.62	7.77	19.17	16.07	18.35
	(7.45)	(11.46)	(18.20)	(19.84)	(18.66)
Gold	1.48	-5.97*	-2.19	1.94	-0.60
	(1.59)	(2.08)	(1.57)	(2.34)	(2.13)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1\text{Stance} + \beta_2\text{Novelty} + \beta_3\text{Stance}\times\text{Novelty} + \varepsilon$. Coefficient on Stance × Novelty reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. Pre-announcement placebo test passed: no significant pre-event coefficients.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 3.12: Price IRF: Policy Stance on Abnormal Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	–	–	–	–	–
VIX Futures	-15.04	-22.74	-28.86	-28.65	-30.54
	(7.07)	(9.76)	(10.84)	(16.52)	(37.21)
10Y T-Note	-0.85	-3.12	-4.67	-8.29	-15.86
	(1.20)	(2.81)	(3.59)	(4.03)	(8.44)
5Y T-Note	0.66	-1.35	-2.86	-6.03	-12.68
	(0.85)	(2.20)	(3.10)	(3.63)	(7.84)
Dollar Index	0.78	-0.55	-2.72	-7.27	-15.99
	(0.92)	(1.86)	(2.63)	(3.73)	(8.39)
Crude Oil WTI	-2.33	-2.79	0.99	-3.76	-2.65
	(3.71)	(5.05)	(7.17)	(8.73)	(8.44)
Gold	-1.77	-5.75	-5.34	-8.84**	-13.76
	(1.25)	(2.57)	(2.54)	(2.66)	(5.73)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1\text{Stance} + \beta_2\text{Novelty} + \beta_3\text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Policy Stance reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. ES entries are omitted because abnormal returns are defined relative to the ES benchmark and are zero by construction. Pre-announcement placebo test passed: no significant pre-event coefficients. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 3.13: Price IRF: Novelty on Abnormal Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	–	–	–	–	–
VIX Futures	10.21	1.84	-4.80	-10.09	7.18
	(9.77)	(10.47)	(8.32)	(6.08)	(10.24)
10Y T-Note	1.81	2.59	2.17	2.48	7.61
	(0.97)	(2.14)	(1.87)	(2.07)	(3.41)
5Y T-Note	0.38	1.26	0.81	0.61	5.00
	(0.58)	(1.71)	(1.37)	(1.91)	(2.57)
Dollar Index	1.40	1.72	1.26	1.56	4.51
	(1.48)	(2.77)	(2.28)	(2.92)	(3.68)
Crude Oil WTI	0.47	0.60	-3.13	-1.42	2.62
	(3.34)	(3.64)	(7.50)	(6.75)	(5.73)
Gold	-0.30	2.34	1.43	-0.21	6.27
	(1.46)	(3.08)	(2.70)	(2.63)	(2.70)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1\text{Stance} + \beta_2\text{Novelty} + \beta_3\text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Novelty reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. ES entries are omitted because abnormal returns are defined relative to the ES benchmark and are zero by construction. Pre-announcement placebo test passed: no significant pre-event coefficients. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 3.14: Price IRF: Stance × Novelty on Abnormal Cumulative Return

	$h = 5$ min	$h = 15$ min	$h = 30$ min	$h = 60$ min	$h = 120$ min
E-mini S&P 500	–	–	–	–	–
VIX Futures	-31.74**	-58.80***	-72.89***	-65.76***	-53.38*
	(9.50)	(14.98)	(18.63)	(13.00)	(17.64)
10Y T-Note	-4.54	-10.44	-11.95	-14.24	-12.05
	(2.09)	(5.70)	(6.03)	(6.30)	(6.89)
5Y T-Note	-1.63	-6.94	-8.03	-9.66	-5.72
	(0.99)	(4.18)	(4.36)	(4.58)	(4.66)
Dollar Index	0.18	-2.85	-5.11	-6.15	-3.52
	(1.65)	(3.98)	(4.37)	(5.05)	(4.70)
Crude Oil WTI	4.84	0.86	10.90	6.59	12.00
	(6.44)	(8.45)	(14.43)	(16.04)	(15.12)
Gold	-1.30	-12.83	-10.31	-7.35	-6.73
	(1.32)	(5.31)	(4.61)	(3.60)	(4.95)
N	148	148	148	148	148

Notes: Jordà (2005) local projection: $\text{CumRet}_i(0 \to h) = \alpha + \beta_1\text{Stance} + \beta_2\text{Novelty} + \beta_3\text{Stance} \times \text{Novelty} + \varepsilon$. Coefficient on Stance × Novelty reported in basis points ($\times 10^4$). All regressors z-scored. Newey–West HAC standard errors in parentheses. ES entries are omitted because abnormal returns are defined relative to the ES benchmark and are zero by construction. Pre-announcement placebo test passed: no significant pre-event coefficients. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (Benjamini–Hochberg adjusted within each (DV, model) family).

Table 3.15: Robustness: Event-Level Regressions — ΔRV and $\Delta\log$RV (30-Minute Window, Interaction Model)

Contract	Dep. Var.	Regressor	$\hat{\beta}$	t	p_{NW}	p_{Boot}	p_{Clust}	p_{QReg}	p_{Perm}	#Sig
Crude Oil WTI	ΔRV	Stance	-7.78	-2.94	0.004	0.691	0.001	–	0.007	3
Crude Oil WTI	ΔRV	Novelty	8.20	2.26	0.025	0.068	0.009	–	0.000	4
Crude Oil WTI	ΔRV	St. × Nov.	-8.61	-2.39	0.018	0.212	0.014	–	0.002	3
Dollar Index	ΔRV	Stance	-1.65	-1.83	0.069	0.260	0.095	–	0.014	3
Dollar Index	ΔRV	Novelty	1.15	1.06	0.293	0.069	0.366	–	0.014	2
Dollar Index	ΔRV	St. × Nov.	-1.20	-0.86	0.392	0.124	0.469	–	0.047	1
E-mini S&P 500	ΔRV	Stance	-4.94	-3.22	0.002	0.584	0.011	–	0.019	3
E-mini S&P 500	ΔRV	Novelty	6.82	2.00	0.048	0.013	0.028	–	0.002	4
E-mini S&P 500	ΔRV	St. × Nov.	-5.32	-1.69	0.094	0.150	0.167	–	0.024	2
Gold	ΔRV	Stance	-1.05	-0.55	0.583	0.692	0.602	–	0.561	0
Gold	ΔRV	Novelty	3.34	2.48	0.014	0.318	0.056	–	0.021	3
Gold	ΔRV	St. × Nov.	-4.68	-1.49	0.138	0.266	0.117	–	0.025	1
VIX Futures	ΔRV	Stance	-2.80	-0.46	0.644	0.659	0.580	–	0.619	0
VIX Futures	ΔRV	Novelty	3.42	0.57	0.568	0.492	0.486	–	0.382	0
VIX Futures	ΔRV	St. × Nov.	-7.96	-3.12	0.002	0.601	0.058	–	0.067	3
5Y T-Note	ΔRV	Stance	-1.30	-2.04	0.043	0.236	0.235	–	0.050	2
5Y T-Note	ΔRV	Novelty	0.94	1.13	0.260	0.203	0.235	–	0.041	1
5Y T-Note	ΔRV	St. × Nov.	-2.48	-2.22	0.028	0.070	0.149	–	0.004	3
10Y T-Note	ΔRV	Stance	-1.50	-2.10	0.038	0.210	0.210	–	0.025	2
10Y T-Note	ΔRV	Novelty	0.70	0.87	0.388	0.314	0.313	–	0.082	1
10Y T-Note	ΔRV	St. × Nov.	-1.66	-1.99	0.048	0.135	0.250	–	0.013	2
Crude Oil WTI	$\Delta\log$RV	Stance	-2.1849	-2.04	0.043	0.442	0.035	–	0.022	3
Crude Oil WTI	$\Delta\log$RV	Novelty	0.8902	2.15	0.033	0.614	0.086	–	0.135	2
Crude Oil WTI	$\Delta\log$RV	St. × Nov.	-0.7406	-1.18	0.240	0.650	0.203	–	0.326	0
Dollar Index	$\Delta\log$RV	Stance	-0.9082	-1.79	0.075	0.337	0.121	–	0.041	2
Dollar Index	$\Delta\log$RV	Novelty	0.5076	0.95	0.345	0.209	0.478	–	0.064	1
Dollar Index	$\Delta\log$RV	St. × Nov.	-0.3540	-0.46	0.649	0.464	0.673	–	0.247	0
E-mini S&P 500	$\Delta\log$RV	Stance	-2.9143	-2.86	0.005	0.716	0.001	–	0.004	3
E-mini S&P 500	$\Delta\log$RV	Novelty	0.3441	0.57	0.570	0.639	0.609	–	0.577	0
E-mini S&P 500	$\Delta\log$RV	St. × Nov.	-0.6767	-1.02	0.311	0.644	0.249	–	0.389	0
Gold	$\Delta\log$RV	Stance	-1.7794	-1.67	0.098	0.433	0.093	–	0.066	3
Gold	$\Delta\log$RV	Novelty	0.7810	1.97	0.051	0.583	0.120	–	0.168	1
Gold	$\Delta\log$RV	St. × Nov.	-0.6972	-1.09	0.275	0.618	0.236	–	0.334	0
VIX Futures	$\Delta\log$RV	Stance	-0.8191	-0.45	0.655	0.610	0.563	–	0.558	0
VIX Futures	$\Delta\log$RV	Novelty	-0.2304	-0.24	0.810	0.816	0.794	–	0.817	0
VIX Futures	$\Delta\log$RV	St. × Nov.	-1.5107	-2.11	0.037	0.781	0.037	–	0.192	2
5Y T-Note	$\Delta\log$RV	Stance	-2.3930	-2.24	0.026	0.545	0.012	–	0.014	3
5Y T-Note	$\Delta\log$RV	Novelty	-0.6468	-1.21	0.227	0.624	0.164	–	0.293	0
5Y T-Note	$\Delta\log$RV	St. × Nov.	-1.0850	-2.16	0.032	0.827	0.024	–	0.163	2
10Y T-Note	$\Delta\log$RV	Stance	-2.5521	-2.31	0.022	0.504	0.010	–	0.012	3
10Y T-Note	$\Delta\log$RV	Novelty	-0.7730	-1.42	0.157	0.697	0.090	–	0.224	1
10Y T-Note	$\Delta\log$RV	St. × Nov.	-1.0145	-1.92	0.057	0.791	0.044	–	0.205	2

Notes: This table reports multi-method robustness results for the event-level interaction model $\Delta Y_i = \alpha + \beta_1 \text{Stance}_i + \beta_2 \text{Novelty}_i + \beta_3(\text{Stance}_i \times \text{Novelty}_i) + \varepsilon_i$ using 30-minute rolling windows. The upper panel reports ΔRV coefficients in basis points; the lower panel reports $\Delta\log$RV coefficients in native units. All regressors are z-scored. Five inference methods are compared: p_{NW} = Newey–West HAC; p_{Boot} = wild bootstrap (1,999 replications, Rademacher weights); p_{Clust} = clustered by event date (HC1); p_{QReg} = quantile regression at median ($\tau = 0.5$); p_{Perm} = permutation test (4,999 replications). #Sig = number of methods yielding $p < 0.10$. A coefficient is considered robust when #Sig ≥ 3. Stance effects on $\Delta\log$RV are significant for 5 of 7 contracts with #Sig ≥ 3, confirming the volatility channel.

Table 3.16: Robustness: Event-Level Regressions — RV Ratio and $\Delta\beta$ (30-Minute Window, Interaction Model)

Contract	Dep. Var.	Regressor	$\hat{\beta}$	t	p_{NW}	p_{Boot}	p_{Clust}	p_{QReg}	p_{Perm}	#Sig
Crude Oil WTI	RV Ratio	Stance	0.0611	0.13	0.895	0.894	0.894	–	0.886	0
Crude Oil WTI	RV Ratio	Novelty	0.4481	1.16	0.249	0.479	0.232	–	0.123	0
Crude Oil WTI	RV Ratio	St. × Nov.	-0.1615	-0.28	0.784	0.777	0.782	–	0.694	0
Dollar Index	RV Ratio	Stance	-1.1473	-1.50	0.158	–	–	–	–	0
Dollar Index	RV Ratio	Novelty	-0.4007	-2.15	0.051	–	–	–	–	1
Dollar Index	RV Ratio	St. × Nov.	0.3796	1.66	0.121	–	–	–	–	0
E-mini S&P 500	RV Ratio	Stance	-0.4896	-1.62	0.110	0.846	0.021	–	0.167	1
E-mini S&P 500	RV Ratio	Novelty	-0.1825	-1.58	0.118	0.851	0.079	–	0.403	1
E-mini S&P 500	RV Ratio	St. × Nov.	0.2511	1.20	0.232	0.856	0.088	–	0.382	1
Gold	RV Ratio	Stance	0.0845	0.16	0.872	0.896	0.881	–	0.876	0
Gold	RV Ratio	Novelty	0.6432	1.41	0.162	0.432	0.159	–	0.093	1
Gold	RV Ratio	St. × Nov.	-0.5527	-1.08	0.284	0.650	0.254	–	0.296	0
VIX Futures	RV Ratio	Stance	0.4669	1.79	0.081	0.655	0.142	–	0.214	1
VIX Futures	RV Ratio	Novelty	-0.3809	-1.33	0.191	0.641	0.289	–	0.290	0
VIX Futures	RV Ratio	St. × Nov.	-0.0118	-0.08	0.940	0.984	0.954	–	0.963	0
5Y T-Note	RV Ratio	Stance	-0.0861	-0.65	0.523	0.559	0.486	–	0.360	0
5Y T-Note	RV Ratio	Novelty	0.0596	0.42	0.675	0.717	0.666	–	0.667	0
5Y T-Note	RV Ratio	St. × Nov.	-0.0415	-0.57	0.574	0.716	0.624	–	0.666	0
10Y T-Note	RV Ratio	Stance	-0.0692	-0.63	0.536	0.642	0.605	–	0.476	0
10Y T-Note	RV Ratio	Novelty	0.0436	0.36	0.717	0.781	0.775	–	0.744	0
10Y T-Note	RV Ratio	St. × Nov.	0.1231	1.54	0.133	0.562	0.182	–	0.234	0
Crude Oil WTI	$\Delta\beta$	Stance	-0.0008	-0.01	0.993	0.998	0.994	–	0.996	0
Crude Oil WTI	$\Delta\beta$	Novelty	-0.1290	-1.26	0.210	0.586	0.213	–	0.149	0
Crude Oil WTI	$\Delta\beta$	St. × Nov.	-0.1772	-0.84	0.403	0.442	0.393	–	0.157	0
Dollar Index	$\Delta\beta$	Stance	0.0137	1.48	0.144	0.456	0.162	–	0.100	1
Dollar Index	$\Delta\beta$	Novelty	-0.0086	-1.15	0.254	0.391	0.310	–	0.118	0
Dollar Index	$\Delta\beta$	St. × Nov.	0.0011	0.13	0.894	0.935	0.889	–	0.813	0
Gold	$\Delta\beta$	Stance	0.1244	1.43	0.156	0.579	0.116	–	0.157	0
Gold	$\Delta\beta$	Novelty	-0.0539	-1.56	0.121	0.721	0.127	–	0.335	0
Gold	$\Delta\beta$	St. × Nov.	-0.0852	-1.13	0.262	0.646	0.177	–	0.270	0
VIX Futures	$\Delta\beta$	Stance	-0.0190	-0.07	0.943	0.944	0.946	–	0.936	0
VIX Futures	$\Delta\beta$	Novelty	0.1096	1.20	0.236	0.704	0.272	–	0.497	0
VIX Futures	$\Delta\beta$	St. × Nov.	-0.0818	-0.80	0.425	0.757	0.495	–	0.642	0
5Y T-Note	$\Delta\beta$	Stance	0.0195	1.23	0.221	0.583	0.393	–	0.409	0
5Y T-Note	$\Delta\beta$	Novelty	-0.0064	-0.88	0.381	0.786	0.481	–	0.639	0
5Y T-Note	$\Delta\beta$	St. × Nov.	0.0074	0.99	0.325	0.820	0.381	–	0.693	0
10Y T-Note	$\Delta\beta$	Stance	0.0118	0.59	0.554	0.772	0.693	–	0.714	0
10Y T-Note	$\Delta\beta$	Novelty	-0.0096	-0.94	0.349	0.736	0.403	–	0.665	0
10Y T-Note	$\Delta\beta$	St. × Nov.	0.0064	0.55	0.585	0.849	0.599	–	0.830	0

Notes: Continuation of Table 3.15. The upper panel reports RV Ratio coefficients (post/pre RV); the lower panel reports $\Delta\beta$ coefficients (change in realized beta relative to ES). All regressors are z-scored. The same five inference methods are used: Newey–West HAC, wild bootstrap, clustered SE, quantile regression, and permutation test. RV Ratio and $\Delta\beta$ results are generally weaker than the ΔRV and Δ logRV results in Table 3.15, suggesting that the volatility channel operates primarily through level changes rather than ratio adjustments or co-movement shifts.

3.9 Figures

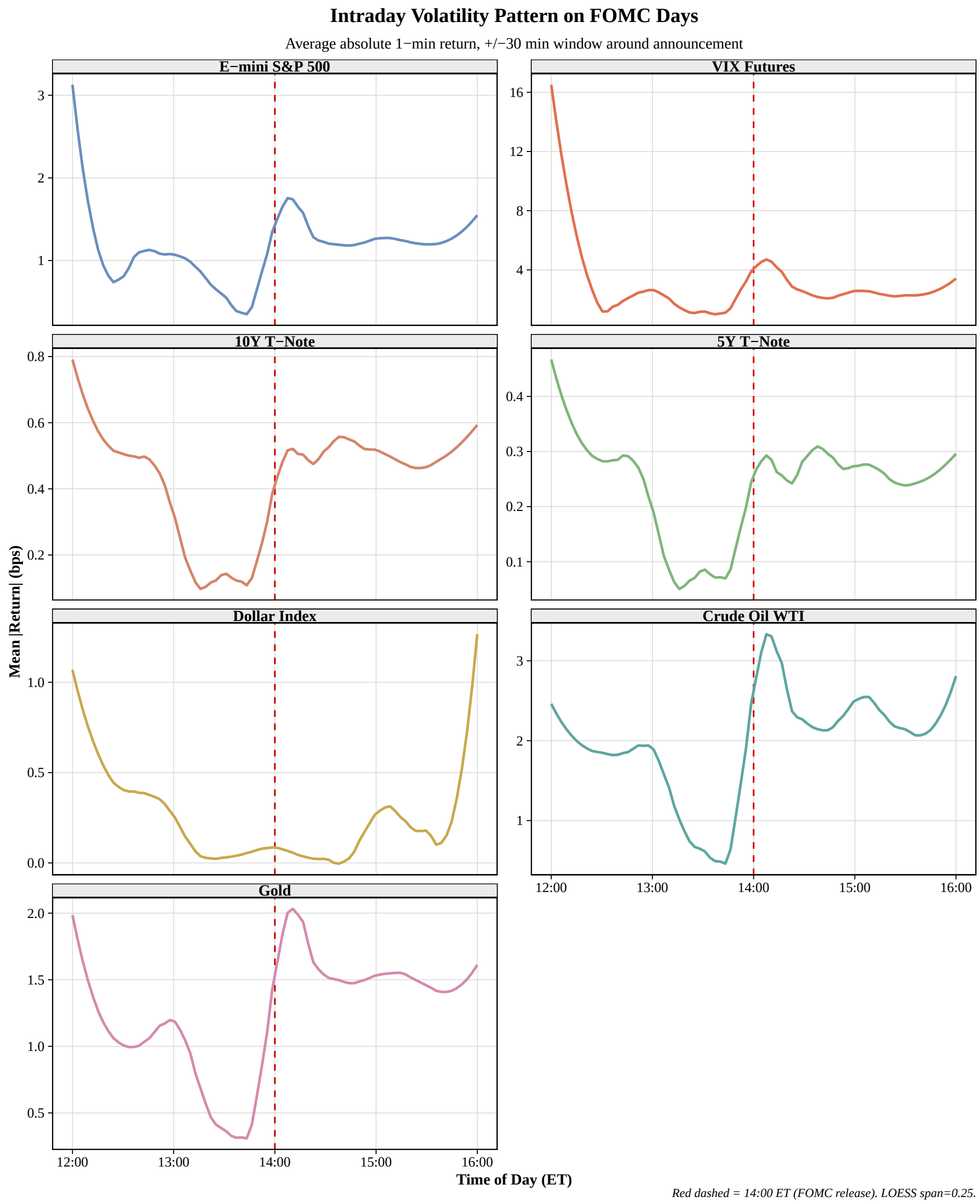


Figure 3.1: Intraday volatility pattern on FOMC days. Red dashed line marks the 14:00 ET announcement. This figure documents the characteristic volatility spike at announcement time, motivating our high-frequency identification strategy. The pre-announcement period shows relatively stable volatility, while the post-announcement spike and gradual decay are consistent with information processing models.

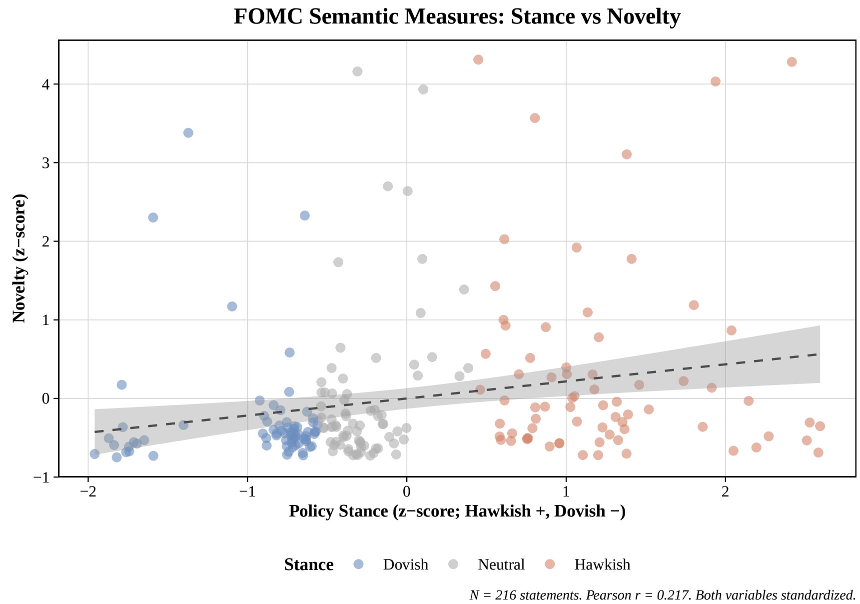


Figure 3.2: FOMC semantic measures: stance vs. novelty (z-scored). Points colored by stance tercile. The low correlation between stance and novelty ($r = 0.12$) supports the conditional independence assumption (Assumption 3.1) and motivates the separate estimation of tone and novelty effects in our decomposition analysis.

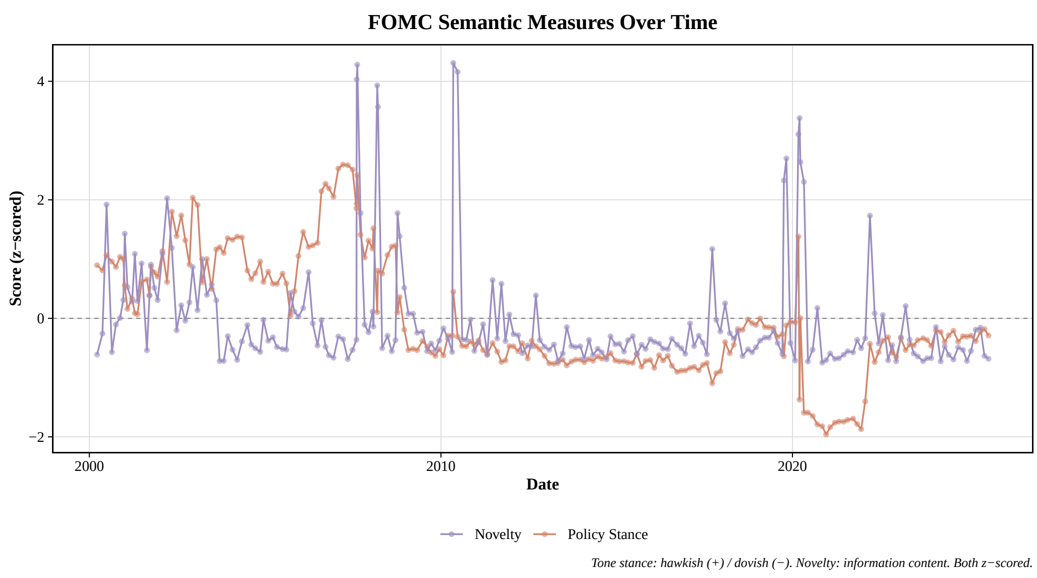
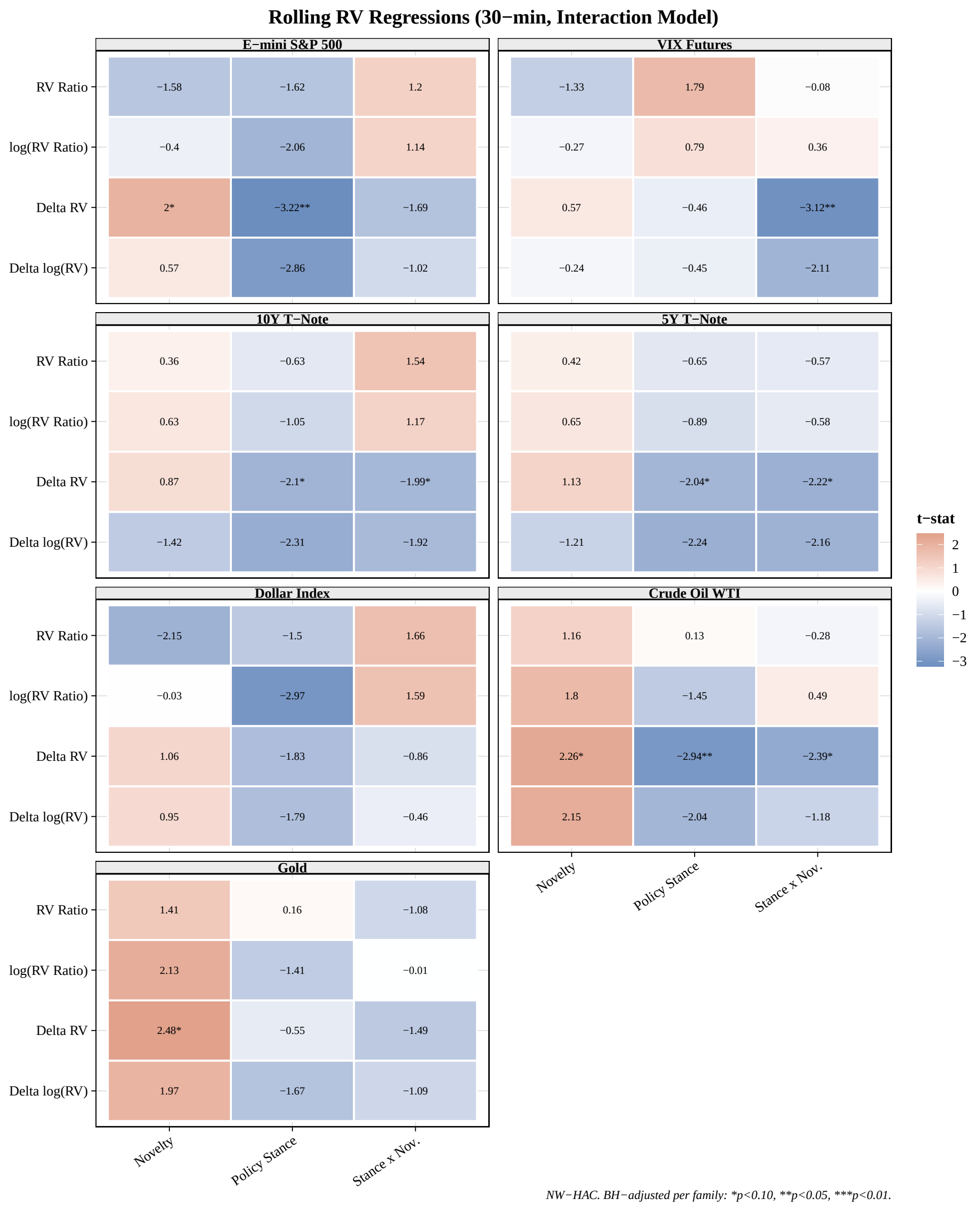


Figure 3.3: Evolution of policy stance and novelty scores over time. Novelty peaks during the 2008 financial crisis and 2020 pandemic correspond to major policy regime changes. Stance shifts from dovish (2008–2015) to hawkish (2017–2019, 2022–2025) track well-known monetary policy cycles.



Ticker	Outcome	Novelty	Policy Stance	Stance x Nov.
E−mini S&P 500	RV Ratio	−1.58	−1.62	1.2
E−mini S&P 500	log(RV Ratio)	−0.4	−2.06	1.14
E−mini S&P 500	Delta RV	2*	−3.22**	−1.69
E−mini S&P 500	Delta log(RV)	0.57	−2.86	−1.02
VIX Futures	RV Ratio	−1.33	1.79	−0.08
VIX Futures	log(RV Ratio)	−0.27	0.79	0.36
VIX Futures	Delta RV	0.57	−0.46	−3.12**
VIX Futures	Delta log(RV)	−0.24	−0.45	−2.11
10Y T−Note	RV Ratio	0.36	−0.63	1.54
10Y T−Note	log(RV Ratio)	0.63	−1.05	1.17
10Y T−Note	Delta RV	0.87	−2.1*	−1.99*
10Y T−Note	Delta log(RV)	−1.42	−2.31	−1.92
5Y T−Note	RV Ratio	0.42	−0.65	−0.57
5Y T−Note	log(RV Ratio)	0.65	−0.89	−0.58
5Y T−Note	Delta RV	1.13	−2.04*	−2.22*
5Y T−Note	Delta log(RV)	−1.21	−2.24	−2.16
Dollar Index	RV Ratio	−2.15	−1.5	1.66
Dollar Index	log(RV Ratio)	−0.03	−2.97	1.59
Dollar Index	Delta RV	1.06	−1.83	−0.86
Dollar Index	Delta log(RV)	0.95	−1.79	−0.46
Crude Oil WTI	RV Ratio	1.16	0.13	−0.28
Crude Oil WTI	log(RV Ratio)	1.8	−1.45	0.49
Crude Oil WTI	Delta RV	2.26*	−2.94**	−2.39*
Crude Oil WTI	Delta log(RV)	2.15	−2.04	−1.18
Gold	RV Ratio	1.41	0.16	−1.08
Gold	log(RV Ratio)	2.13	−1.41	−0.01
Gold	Delta RV	2.48*	−0.55	−1.49
Gold	Delta log(RV)	1.97	−1.67	−1.09

Figure 3.4: Heatmap of event-level ΔRV regression coefficients by ticker. Darker shading indicates stronger effects. Stance effects are negative and significant across most contracts, while novelty effects are positive. The stance×novelty interaction is concentrated in VIX and Treasury securities, suggesting that the joint impact of tone and information content operates primarily through the uncertainty channel.

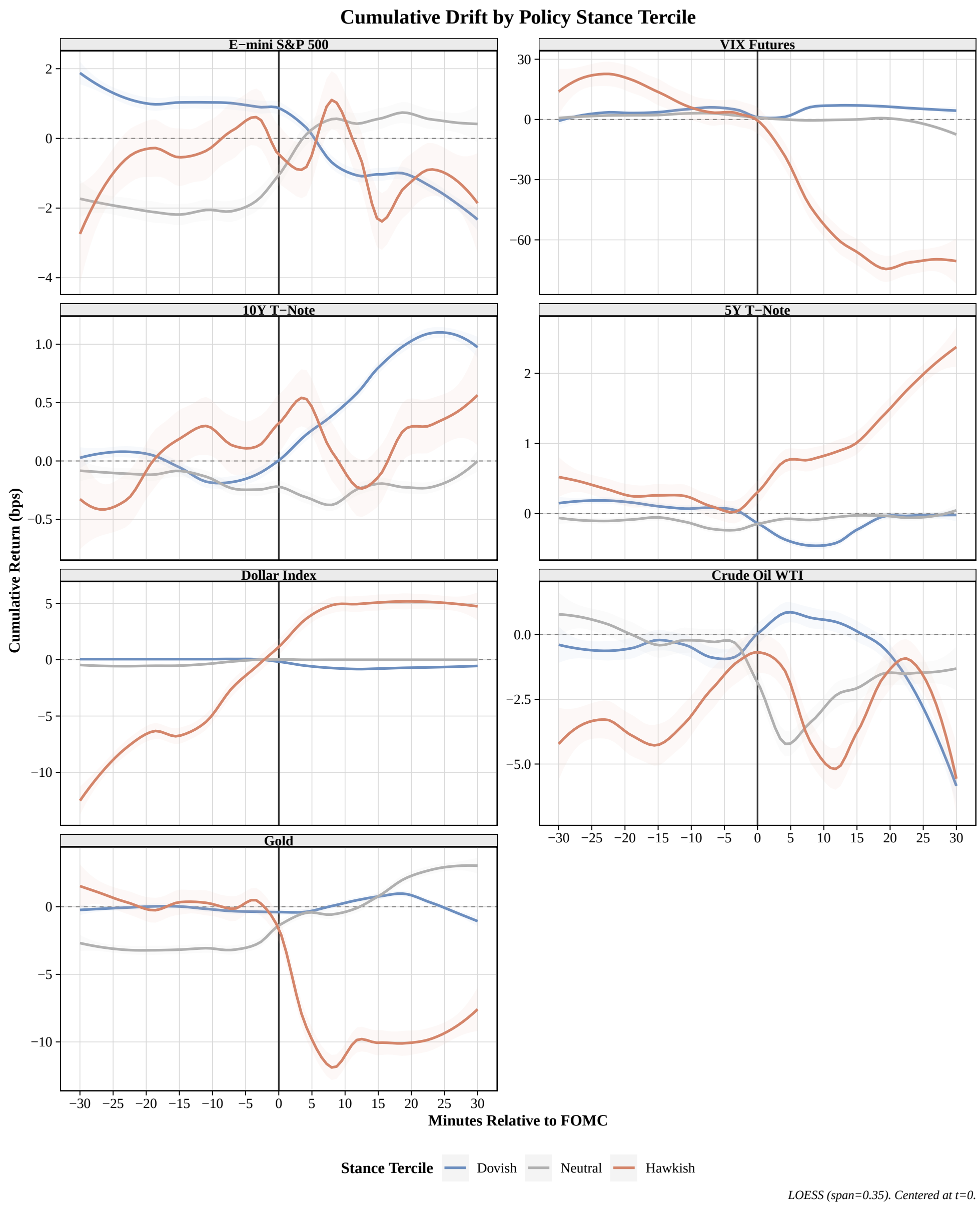


Figure 3.5: Cumulative return drift by stance tercile. Dovish announcements generate positive equity drift that strengthens over 45 minutes, while hawkish announcements produce symmetric negative drift. The monotonic separation between terciles and gradual strengthening over time support H1c (tone effects increasing in magnitude) and the fundamental repricing interpretation.

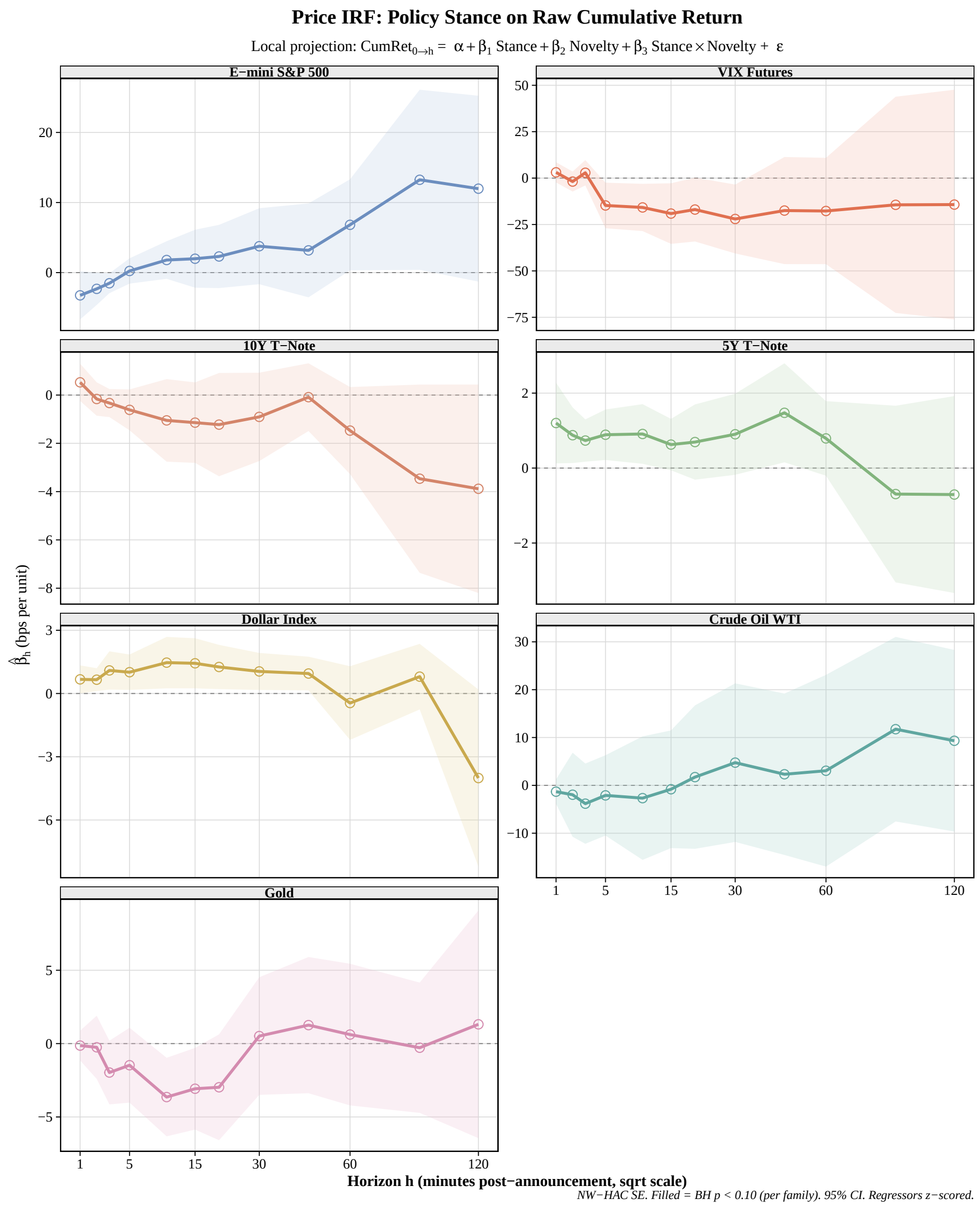


Figure 3.6: Impulse response: cumulative return to policy stance (post-announcement, Newey–West SE). The equity (ES) response builds gradually over the 120-minute window, consistent with slow fundamental repricing by heterogeneous investors (H1c). The VIX response is negative and peaks within 30 minutes, consistent with directionally clear statements resolving uncertainty quickly.

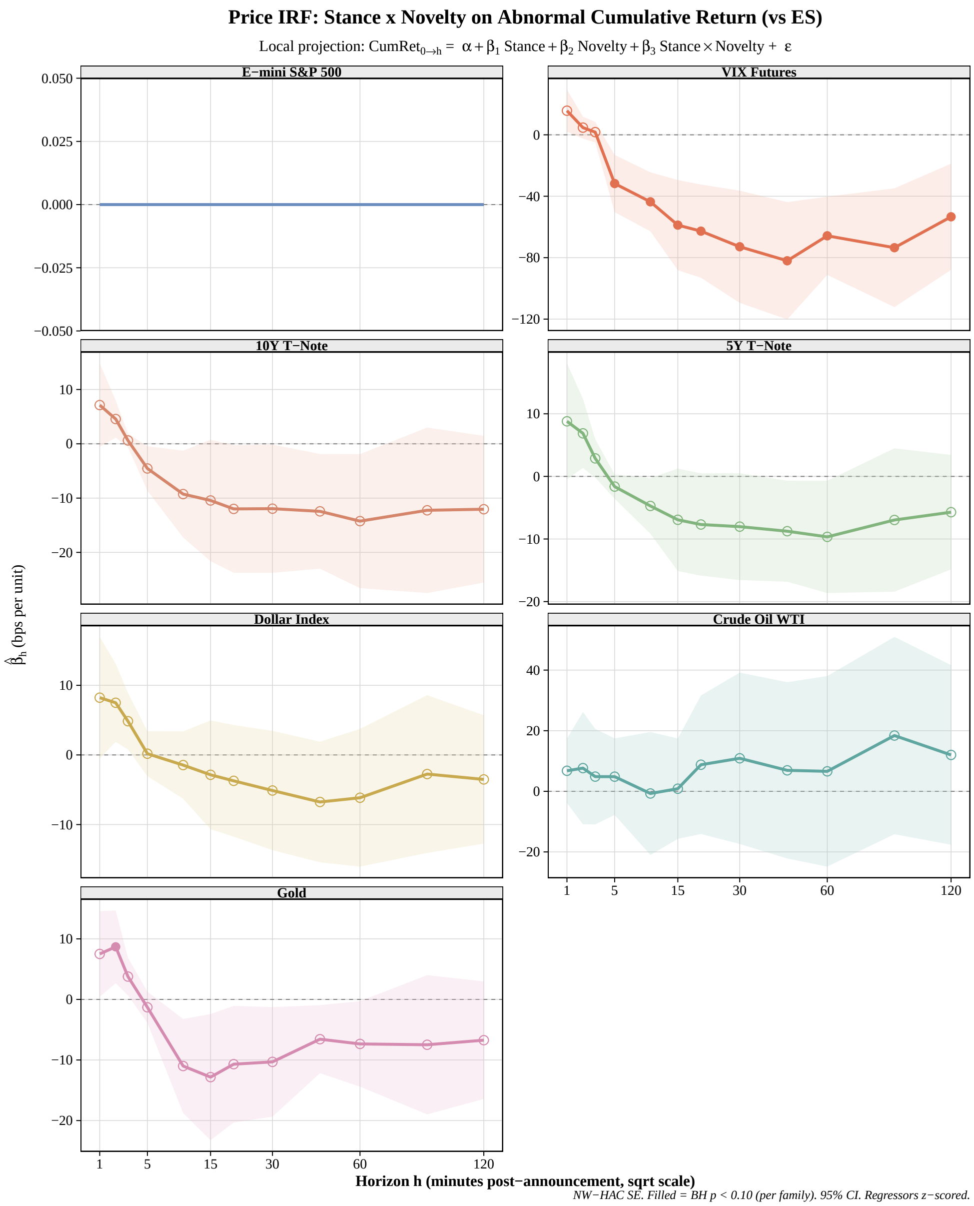


Figure 3.7: Impulse response: abnormal cumulative return to stance × novelty interaction. VIX shows persistent, significant effects from 5 to 120 minutes ($t = -5.06$ at peak), representing the most robust finding in our analysis. This confirms that the joint impact of tone and novelty on uncertainty is both economically large and statistically robust across the full post-announcement window.

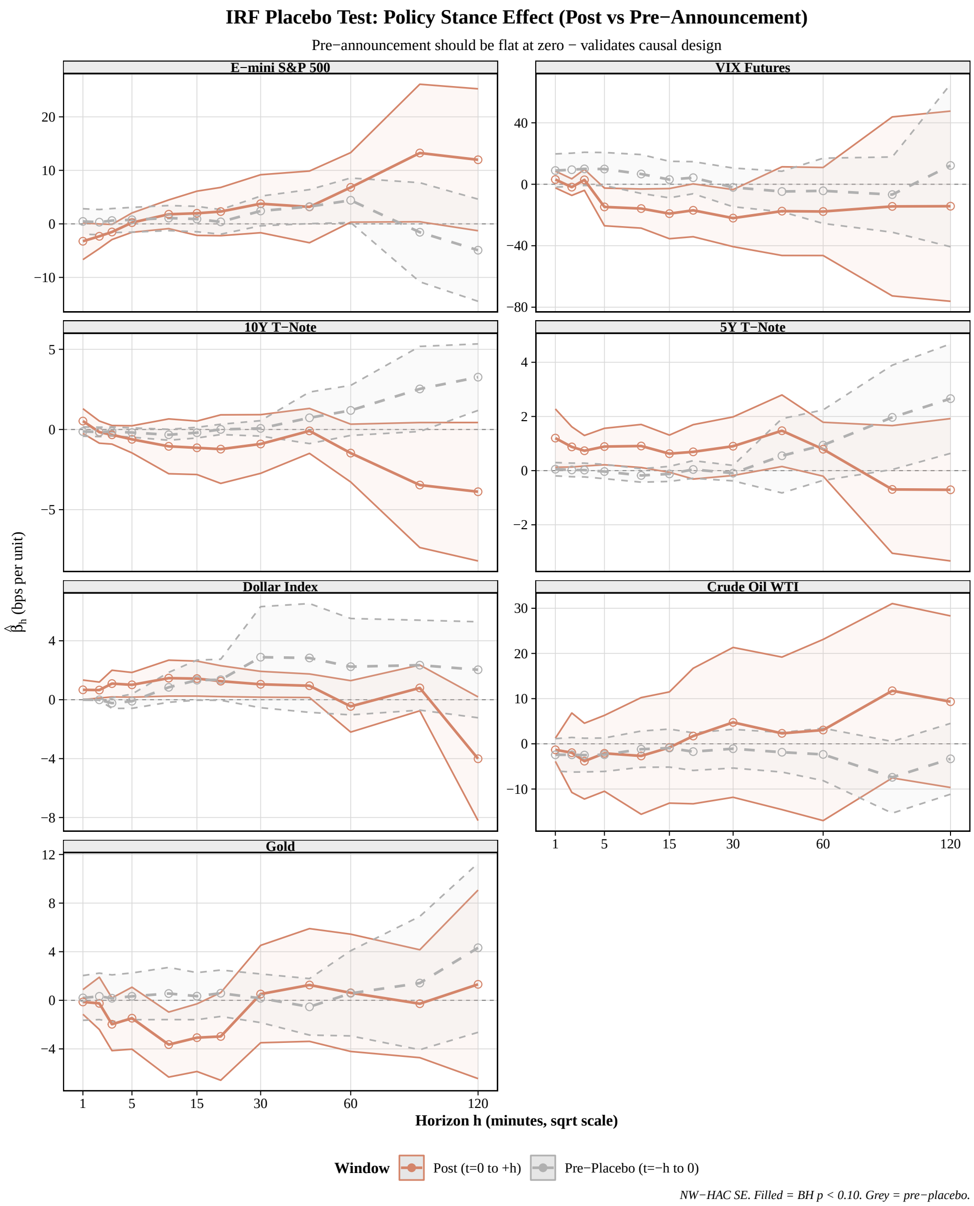


Figure 3.8: Pre-announcement placebo: IRF for stance. All coefficients are near zero and statistically insignificant across all tickers and horizons, validating our event study design. The absence of pre-announcement effects rules out information leakage and confirms that our identification strategy successfully isolates the causal impact of FOMC statement content.

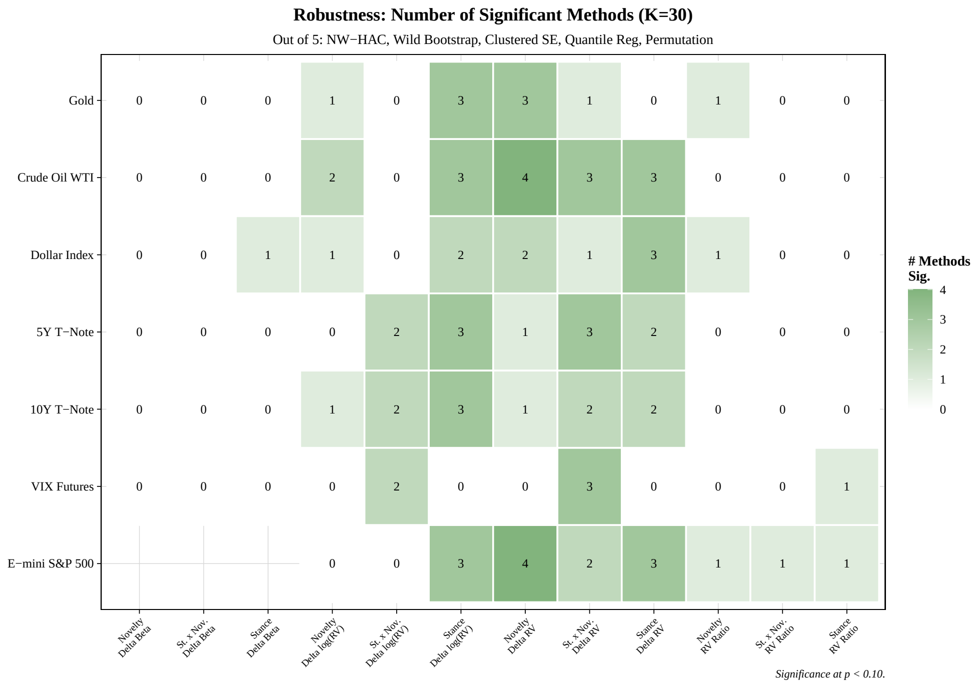


Figure 3.9: Multi-method robustness heatmap: number of methods (out of 5) yielding $p < 0.10$ for each ticker–variable combination. Darker cells indicate more robust results. Stance effects on ΔRV and $\Delta\log$RV are the most robust, with 3–4 methods confirming significance for ES, CL, and Treasury contracts. The interaction term is most robust for VIX, consistent with the uncertainty channel interpretation.

Conclusion

This dissertation set out to understand how information is transmitted to prices and volatility in commodities and financial futures markets, using the identification power of high-frequency data. Three essays examined three distinct transmission mechanisms: the interaction between macroeconomic news and speculative positioning in energy futures; the arbitrage-driven propagation of volatility between commodity ETFs and their underlying baskets; and the semantic channels through which central bank communication moves asset prices. This concluding chapter summarizes the findings, draws out what they imply jointly, acknowledges limitations, and sketches avenues for future research.

Summary of contributions

The first essay showed that speculative trading in energy and commodity futures markets acts as a stabilizing force around macroeconomic announcements. Using 5-minute data over 2007–2024 and a time-varying measure of speculative intensity built from disaggregated CFTC positions, it found that higher speculative activity dampens the reaction of returns and conditional volatility to standardized macroeconomic surprises and narrows bid-ask spreads, with the effects concentrated among money managers rather than swap dealers, and stronger for procyclical energy commodities than for safe-haven gold. The essay thus provides sharply identified evidence against the view that speculation amplifies news-driven volatility, and speaks directly to the design of position limits.

The second essay built the first long-sample, minute-level dataset of indicative net asset values for commodity ETFs and used it to measure volatility transmission between funds and their underlying assets. It found that the direction of transmission reflects the arbitrage technology — one-way, from basket to fund, for physically backed precious metals; two-way and asymmetric for futures-based energy funds — that transmission operates predominantly through jumps rather than the continuous component of volatility, and that 1-minute data reveal transmission up to twice as large as 30-minute estimates. The iNAV emerges as a measurement instrument of independent value for the microstructure of index-linked products.

The third essay decomposed FOMC statements into policy tone and informational novelty using an ensemble of transformer language models with data-driven reference selection, and traced their effects through 1-minute futures data across 148 announcements. Tone predicts directional returns, building over two hours; novelty predicts volatility, resolving within roughly twenty minutes and with a clarifying rather than noise-inducing effect; and their interaction on VIX futures is the single most robust result. The essay shows that the informational content of central bank communication is multidimensional, and that its dimensions travel through different economic channels at different speeds.

Across the three essays, a consistent methodological lesson emerges: the minute-level resolution is not a refinement but a prerequisite. The damping effect of speculation, the jump channel of ETF volatility transmission, and the divergent dynamics of tone and novelty are all features of the first minutes and hours after an information event; at the daily frequency they are attenuated, conflated, or invisible.

Limitations

Several limitations qualify these results and delimit their scope. First, the measures of trader positioning in the first essay derive from weekly CFTC reports, so that speculative intensity is observed at a coarser frequency than the market reactions it conditions; the identification rests on the persistence of positioning rather than on its intraday variation. Second, the iNAV series of the second essay are constructed from disseminated and reconstructed data whose quality, while validated against official NAVs, cannot be audited tick by tick over the whole sample; and the analysis covers four large single-commodity funds, leaving open how the results extend to broad-basket or leveraged products. Third, the semantic measures of the third essay are estimated from a finite corpus of Federal Reserve communications; although reference selection is data-driven and robustness was assessed with multiple inference methods, language-model-based measures inevitably embed modelling choices, and the analysis concerns statements rather than the full communication apparatus (minutes, press conferences, speeches). Finally, all three essays are about the United States markets; the external validity of the findings for other trading venues and regulatory environments remains to be established.

Avenues for future research

The results open several paths. On the policy side, the trader-level heterogeneity documented in the first essay — stabilizing money managers, amplifying swap dealers — invites a finer analysis of position-limit design and of the role of speculative capital in financing the energy transition, where volatility dampening lowers the real-option value of delaying investment. On the market-structure side, the jump-based transmission channel identified in the second essay suggests extending the iNAV apparatus to broad-basket, leveraged, and fixed-income ETFs, and studying how arbitrage frictions — creation/redemption costs, settlement technology, market-maker inventory — shape the propagation of discontinuous risk. On the communication side, the tone/novelty decomposition of the third essay can be carried to press conferences and minutes, to other central banks, and to the cross-section of individual assets, where the divergent horizons of the return and volatility channels may help separate expectations formation from uncertainty resolution. More broadly, the combination of language-model measurement with high-frequency identification — used here for monetary policy — applies to any recurring, text-borne information event: earnings calls, regulatory releases, geopolitical announcements.

High-frequency data have turned questions that were once matters of narrative — does speculation destabilize? do ETFs transmit shocks? do words move markets? — into questions of measurement. The three essays of this dissertation are a step in that direction for commodities and financial futures markets.

Appendix A

Mathematical Proofs (Chapter 3)

This appendix provides formal mathematical derivations for the key theoretical relationships presented in Section 3.5.

A.1 Derivation of Pure Stance Definitions

Theorem A.1 (Pure Stance Characterization)**.** *The pure dovish and hawkish stance measures satisfy the relationships given in Definition 3.7.*

Proof. **Case 1: Pure Dovish Statement.** Assume the Federal Reserve releases a statement that exactly matches the dovish counterfactual: $F_t = F_t^D$.

The tone measure becomes:

$$\begin{aligned}\text{Tone}_t &= \frac{\text{sim}(F_t, F_t^H) - \text{sim}(F_t, F_t^D)}{1 - \text{sim}(F_t^D, F_t^H)} \\ &= \frac{\text{sim}(F_t^D, F_t^H) - \text{sim}(F_t^D, F_t^D)}{1 - \text{sim}(F_t^D, F_t^H)} \\ &= \frac{\text{sim}(F_t^D, F_t^H) - 1}{1 - \text{sim}(F_t^D, F_t^H)} \\ &= -1 \end{aligned} \tag{A.1}$$

The novelty measure is:

$$\text{Novelty}_t = 1 - \text{sim}(F_t, F_{t-1}) = 1 - \text{sim}(F_t^D, F_{t-1}) \tag{A.2}$$

Therefore, the stance measure is:

$$\begin{aligned}
\text{Stance}_t &= \text{Novelty}_t \times \text{Tone}_t \\
&= \left(1 - \text{sim}(F_t^D, F_{t-1})\right) \times (-1) \\
&= -\left(1 - \text{sim}(F_t^D, F_{t-1})\right) \\
&= \text{Stance}_t^{dove}
\end{aligned} \tag{A.3}$$

Case 2: Pure Hawkish Statement. Assume $F_t = F_t^H$:

The tone measure becomes:

$$\begin{aligned}
\text{Tone}_t &= \frac{\text{sim}(F_t^H, F_t^H) - \text{sim}(F_t^H, F_t^D)}{1 - \text{sim}(F_t^D, F_t^H)} \\
&= \frac{1 - \text{sim}(F_t^H, F_t^D)}{1 - \text{sim}(F_t^D, F_t^H)} \\
&= 1
\end{aligned} \tag{A.4}$$

Therefore:

$$\text{Stance}_t = \left(1 - \text{sim}(F_t^H, F_{t-1})\right) \times 1 = \text{Stance}_t^{hawk} \tag{A.5}$$

This completes the proof. □ □

A.2 Derivation of Dovish Weight Parameter

Theorem A.2 (Dovish Weight Parameter Formula)**.** *The weight parameter w_t in the weighted stance representation has the form given in Definition 3.8.*

Proof. From the weighted stance equation:

$$\text{Stance}_t = w_t \cdot \text{Stance}_t^{dove} + (1 - w_t) \cdot \text{Stance}_t^{hawk} \tag{A.6}$$

Substituting the expressions for pure stances from Theorem 1:

$$\begin{aligned}
\text{Stance}_t &= -w_t \left(1 - \text{sim}(F_t^D, F_{t-1})\right) + (1 - w_t) \left(1 - \text{sim}(F_t^H, F_{t-1})\right) \\
&= 1 - 2w_t + w_t \left(\text{sim}(F_t^D, F_{t-1}) + \text{sim}(F_t^H, F_{t-1})\right) - \text{sim}(F_t^H, F_{t-1})
\end{aligned} \tag{A.7}$$

Collecting the terms in w_t, this reads

$$\text{Stance}_t = 1 - \text{sim}(F_t^H, F_{t-1}) - w_t \left(2 - \text{sim}(F_t^D, F_{t-1}) - \text{sim}(F_t^H, F_{t-1})\right). \tag{A.8}$$

Equating with $\text{Stance}_t = (1 - \text{sim}(F_t, F_{t-1})) \times \text{Tone}_t$ and solving for w_t:

$$w_t = \frac{1 - \text{sim}(F_t^H, F_{t-1}) - (1 - \text{sim}(F_t, F_{t-1})) \times \text{Tone}_t}{2 - \text{sim}(F_t^D, F_{t-1}) - \text{sim}(F_t^H, F_{t-1})} \tag{A.9}$$

This establishes the formula. □ □

A.3 Proof of MPS Decomposition

Theorem A.3 (Policy Stance Surprise Decomposition)**.** *The policy stance surprise admits the decomposition given in Proposition 3.1.*

Proof. From the definitions:

$$\begin{aligned}
\text{MPS}_t &= \text{Stance}_t - \mathbb{E}_{t-\Delta}[\text{Stance}_t] \\
&= \text{Novelty}_t \times \text{Tone}_t - (1 - 2p_{t-\Delta}) \cdot \overline{\text{Novelty}}_{t|t-\Delta} \\
&= \left(\overline{\text{Novelty}}_{t|t-\Delta} + \varepsilon_t\right) \times \text{Tone}_t - (1 - 2p_{t-\Delta}) \cdot \overline{\text{Novelty}}_{t|t-\Delta} \\
&= \overline{\text{Novelty}}_{t|t-\Delta} \times \text{Tone}_t - \overline{\text{Novelty}}_{t|t-\Delta} + 2p_{t-\Delta} \cdot \overline{\text{Novelty}}_{t|t-\Delta} + \varepsilon_t \times \text{Tone}_t \\
&= \overline{\text{Novelty}}_{t|t-\Delta}(\text{Tone}_t + 2p_{t-\Delta} - 1) + \text{Tone}_t \cdot \varepsilon_t
\end{aligned} \tag{A.10}$$

This establishes the decomposition. □ □

A.4 Economic Interpretation

The mathematical results provide several economic insights. First, the pure stance characterization shows that our measures correctly identify extreme policy communications, with dovish statements receiving negative stance values and hawkish statements receiving positive values. Second, the weight parameter derivation reveals how actual policy communications can be understood as weighted averages of extreme alternatives. Third, the MPS decomposition shows that policy surprises have two distinct sources: unexpected tone conditional on expected information content, and unexpected information content weighted by actual tone.

Appendix B

Additional Tables and Figures (Chapter 3)

This appendix collects supplementary tables and figures that support the main results but are not essential for following the core argument.

B.1 Descriptive Figures

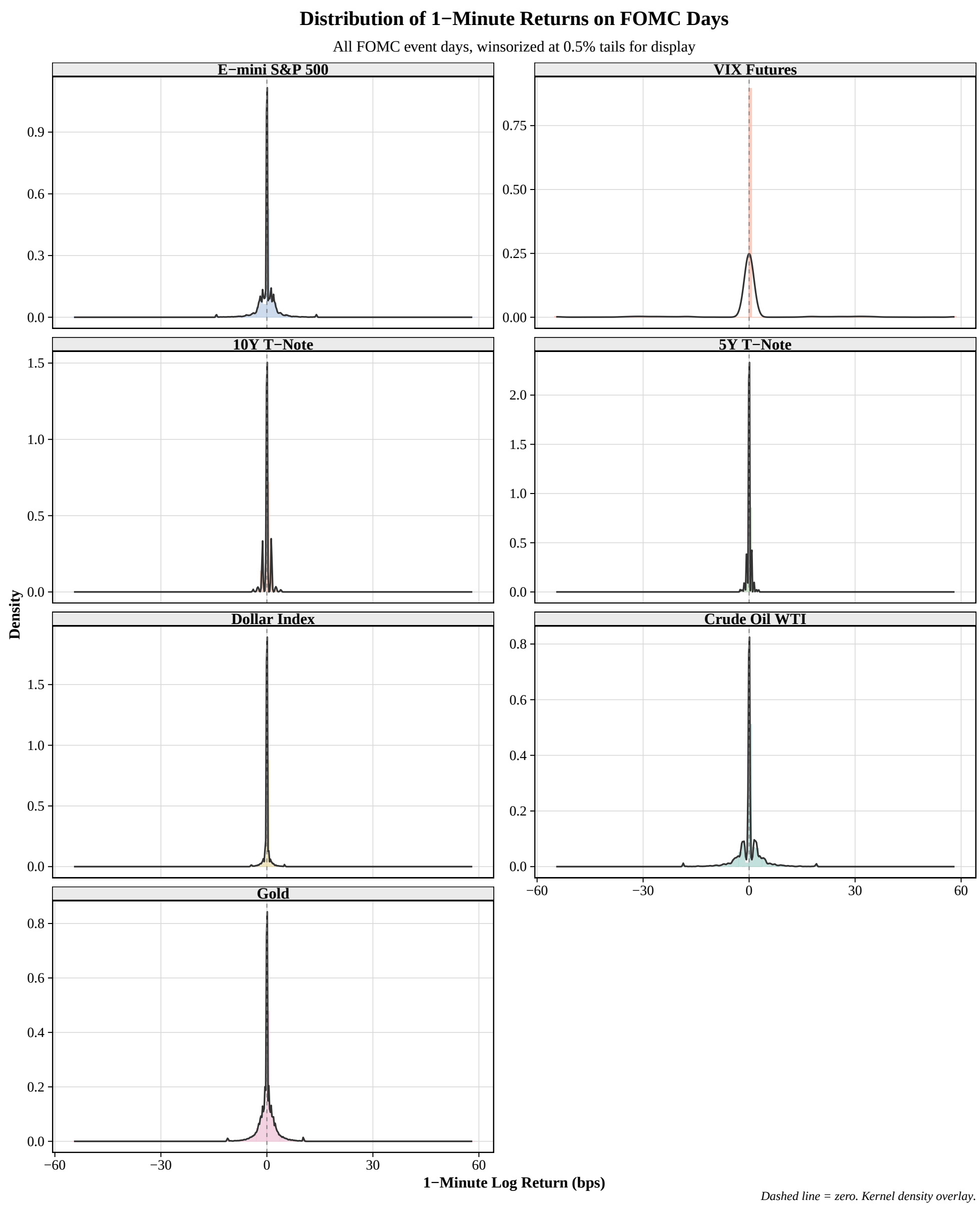


Figure B.1: Distribution of 1-minute log returns on FOMC days. Winsorized at 0.5% tails for display. The heavy tails and leptokurtic shape are consistent with the Jarque–Bera test rejections reported in Table 3.2.

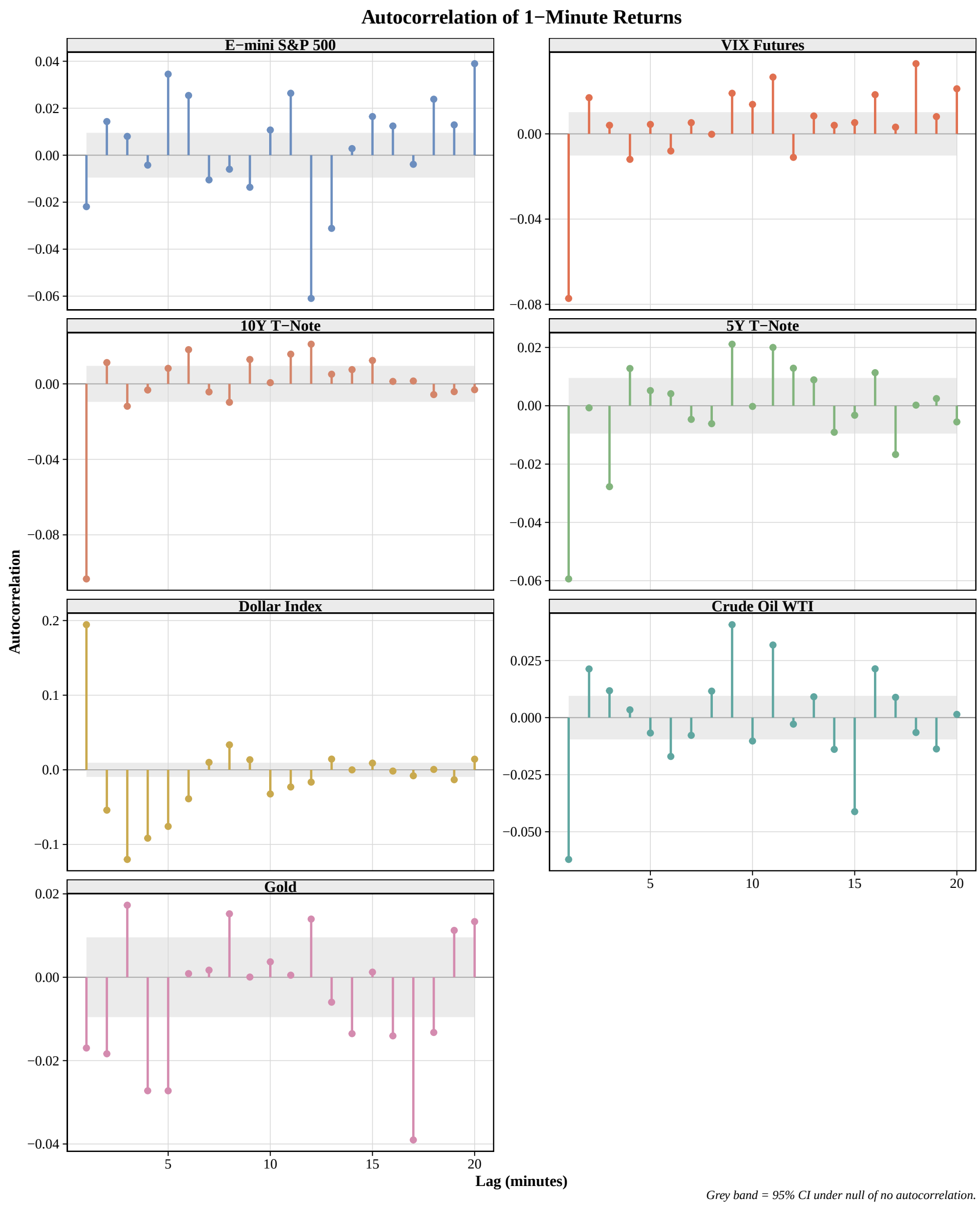


Figure B.2: Autocorrelation of 1-minute returns. Grey band shows 95% confidence interval. Negative first-order autocorrelation reflects bid–ask bounce effects typical of high-frequency data.

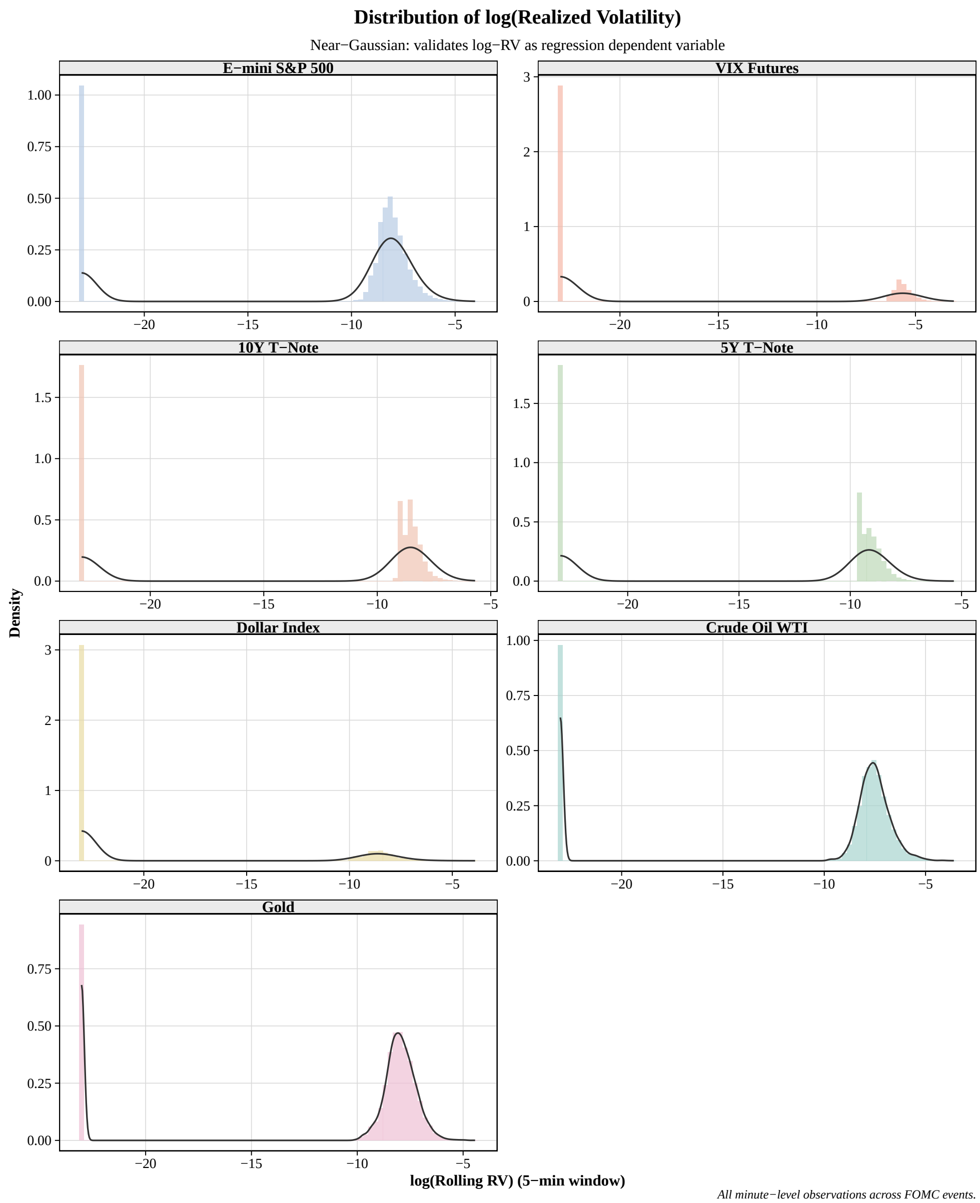


Figure B.3: Distribution of log(RV). Near-Gaussian shape validates the use of log(RV) as the dependent variable in the panel regressions.

B.2 Ensemble Model Diagnostics

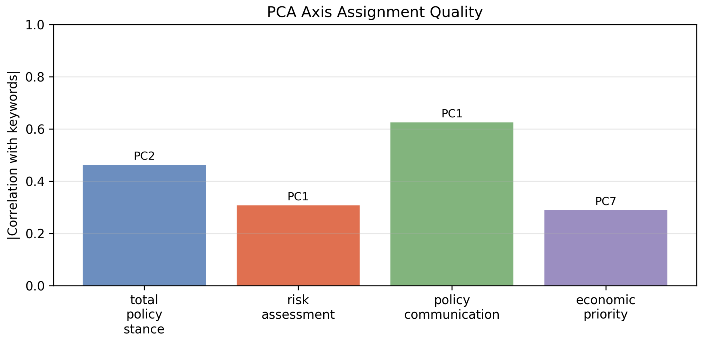
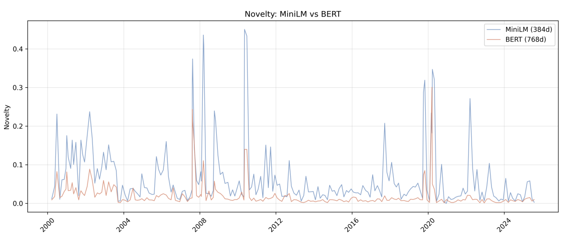
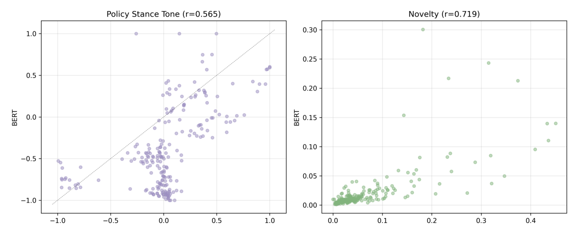


Figure B.4: PCA axis assignment quality. Bar height shows |Corr| between PC projection and keyword differential. All axes exceed the 0.04 minimum separation threshold.

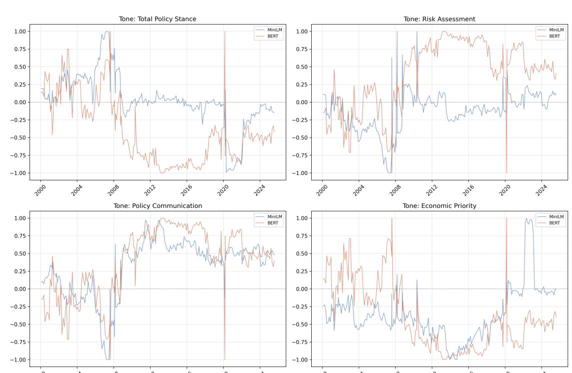


(a) Novelty: MiniLM vs BERT

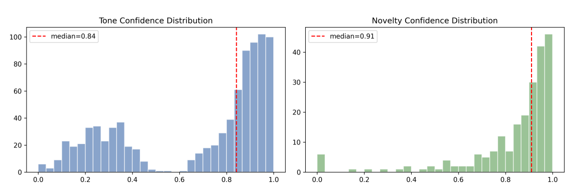


(b) Score scatter plots

Figure B.5: Model comparison: MiniLM (384d) vs. BERT (768d). Novelty correlation $r = 0.72$; policy stance tone $r = 0.56$. The moderate inter-model correlation supports the use of an ensemble approach.



(a) Tone comparison across models



(b) Ensemble confidence distribution

Figure B.6: Tone model comparison and confidence diagnostics. The right panel shows that ensemble confidence is concentrated above 0.7, indicating strong inter-model agreement for most statements.

B.3 Additional Panel Regression Tables

Tables B.1–B.6 report panel minute-level regressions for the remaining dependent variables (realized volatility in levels, realized beta, and returns), separately for stance and novelty.

Table B.1: Panel Regressions: RV on Post × Stance (5-min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	2.79***	7.86***	0.88***	0.54***	0.23**	4.32***	2.92***
	(0.40)	(1.56)	(0.10)	(0.08)	(0.09)	(0.56)	(0.30)
Post × Stance	-1.27	-7.49*	-0.76**	-0.35	0.05	-4.97**	-1.79*
	(0.91)	(3.41)	(0.28)	(0.24)	(0.22)	(1.92)	(0.85)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.008	0.021	0.042	0.012	0.000	0.036	0.021
Adj. R^2	0.008	0.021	0.042	0.012	0.000	0.036	0.021

Notes: Dependent variable: RV (bps, 5-min). Post×Stance $= \mathbf{1}[t > 0] \times$ Stance (z-scored). Post×Novelty $= \mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ± 30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.2: Panel Regressions: RV on Post × Novelty (5-min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	3.26***	10.79***	1.09***	0.64***	0.25***	5.85***	3.49***
	(0.33)	(1.54)	(0.09)	(0.07)	(0.08)	(0.62)	(0.27)
Post × Novelty	3.09	8.32	0.14	0.25	0.63	4.94	1.93
	(1.37)	(5.78)	(0.35)	(0.30)	(0.36)	(2.40)	(1.01)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.097	0.040	0.003	0.013	0.043	0.071	0.050
Adj. R^2	0.097	0.040	0.003	0.012	0.042	0.071	0.050

Notes: Dependent variable: RV (bps, 5-min). Post×Stance $= \mathbf{1}[t > 0] \times$ Stance (z-scored). Post×Novelty $= \mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ±30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.3: Panel Regressions: Beta on Post × Stance (5-min)

	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-0.5789***	-0.0206	-0.0175**	-0.0023	0.0992	0.0405
	(0.1988)	(0.0153)	(0.0087)	(0.0016)	(0.0644)	(0.0342)
Post × Stance	0.1780	-0.0179	-0.0165	-0.0023	-0.1643	-0.0415
	(0.2433)	(0.0178)	(0.0098)	(0.0017)	(0.0782)	(0.0404)
N	4,798	5,724	5,724	5,724	5,724	5,663
R^2	0.000	0.001	0.002	0.001	0.003	0.001
Adj. R^2	0.000	0.001	0.002	0.001	0.003	0.000

Notes: Dependent variable: Realized Beta (5-min). Post×Stance $= \mathbf{1}[t > 0] \times$ Stance (z-scored). Post×Novelty $= \mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ±30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.4: Panel Regressions: Beta on Post × Novelty (5-min)

	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-0.6861***	-0.0144	-0.0114	-0.0017	0.1629***	0.0538**
	(0.1540)	(0.0120)	(0.0070)	(0.0011)	(0.0582)	(0.0268)
Post × Novelty	-0.1361	-0.0116	-0.0064	-0.0037	-0.0295	-0.0445
	(0.1765)	(0.0080)	(0.0048)	(0.0030)	(0.0706)	(0.0204)
N	4,798	5,724	5,724	5,724	5,724	5,663
R^2	0.001	0.001	0.001	0.005	0.000	0.002
Adj. R^2	0.000	0.000	0.000	0.005	0.000	0.001

Notes: Dependent variable: Realized Beta (5-min). Post×Stance = $\mathbf{1}[t > 0] \times$ Stance (z-scored). Post×Novelty = $\mathbf{1}[t > 0] \times$ Novelty (z-scored). $K = 5$ min rolling window, ± 30 min event window. Clustered SE by event date. Stars: BH-adjusted.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.5: Panel Regressions: Returns on Post × Stance (5-min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	0.01	-0.18	0.01	0.01	0.03**	-0.02	0.04
	(0.04)	(0.16)	(0.01)	(0.01)	(0.01)	(0.05)	(0.03)
Post × Stance	0.11	-0.41	-0.02	0.02	0.05	0.15	0.02
	(0.09)	(0.52)	(0.04)	(0.01)	(0.02)	(0.30)	(0.05)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.000	0.000	0.000	0.000	0.001	0.000	0.000
Adj. R^2	0.000	0.000	-0.000	0.000	0.000	0.000	-0.000

Notes: Dependent variable: 1-Min Return (bps, 5-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Coefficients $\times 10,000$. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.6: Panel Regressions: Returns on Post × Novelty (5-min)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	-0.02	0.00	0.01	0.00	0.01	-0.07	0.04
	(0.03)	(0.19)	(0.01)	(0.01)	(0.01)	(0.10)	(0.03)
Post × Novelty	-0.04	0.62	0.06	0.01	0.01	-0.18	0.03
	(0.10)	(0.79)	(0.05)	(0.01)	(0.05)	(0.42)	(0.06)
N	9,028	7,930	9,028	9,028	9,028	9,028	8,967
R^2	0.000	0.001	0.002	0.000	0.000	0.000	0.000
Adj. R^2	-0.000	0.001	0.002	-0.000	-0.000	0.000	-0.000

Notes: Dependent variable: 1-Min Return (bps, 5-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Coefficients $\times 10,000$. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

B.4 Additional Event-Level Regression Tables

Tables B.7 and B.8 report the event-level RV-ratio specifications, which corroborate the pre/post difference results in the main text.

Table B.7: Rolling RV Ratio Regressions (30-min Window, Interaction Model)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	1.6276***	1.6260***	0.9991***	1.0135***	1.8017**	2.9633***	3.7617***
	(0.1997)	(0.2639)	(0.0987)	(0.0537)	(0.6352)	(0.3545)	(0.4085)
Stance	-0.4896	0.4669	-0.0692	-0.0861	-1.1473	0.0611	0.0845
	(0.3029)	(0.2605)	(0.1108)	(0.1334)	(0.7665)	(0.4605)	(0.5241)
Novelty	-0.1825	-0.3809	0.0436	0.0596	-0.4007	0.4481	0.6432
	(0.1157)	(0.2858)	(0.1196)	(0.1413)	(0.1864)	(0.3861)	(0.4561)
Stance × Novelty	0.2511	-0.0118	0.1231	-0.0415	0.3796	-0.1615	-0.5527
	(0.2085)	(0.1563)	(0.0801)	(0.0731)	(0.2289)	(0.5868)	(0.5128)
N	88	41	40	40	17	89	88
R^2	0.050	0.050	0.129	0.025	0.145	0.035	0.044
Adj. R^2	0.016	-0.027	0.057	-0.056	-0.053	0.001	0.010

Notes: Dependent variable: RV Ratio (Rolling 30-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Semantic measures from MiniLM–BERT ensemble.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table B.8: Rolling log(RV Ratio) Regressions (30-min Window, Interaction Model)

	E-mini S&P 500	VIX Futures	10Y T-Note	5Y T-Note	Dollar Index	Crude Oil WTI	Gold
Intercept	0.2577***	-0.6826	-0.1044	-0.0595	-2.2662***	0.2341	0.4248
	(0.0810)	(0.6770)	(0.1403)	(0.0703)	(0.7194)	(0.3526)	(0.3908)
Stance	-0.2575	0.4071	-0.1965	-0.1674	-4.7532	-0.7852	-0.8420
	(0.1249)	(0.5125)	(0.1874)	(0.1886)	(1.5983)	(0.5421)	(0.5977)
Novelty	-0.0173	-0.1680	0.0871	0.1068	-0.0110	0.3561	0.4141
	(0.0437)	(0.6219)	(0.1386)	(0.1650)	(0.3329)	(0.1975)	(0.1945)
Stance × Novelty	0.0890	0.1548	0.1046	-0.0441	1.2119	0.1608	-0.0041
	(0.0779)	(0.4344)	(0.0892)	(0.0755)	(0.7615)	(0.3288)	(0.3215)
N	88	41	40	40	17	89	88
R^2	0.069	0.004	0.183	0.078	0.366	0.086	0.083
Adj. R^2	0.036	-0.077	0.114	0.002	0.220	0.054	0.050

Notes: Dependent variable: log(RV Ratio) (Rolling 30-min). Clustered SE in parentheses. Stars: BH-adjusted p-values. Semantic measures from MiniLM–BERT ensemble.
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

B.5 Additional IRF and Drift Figures

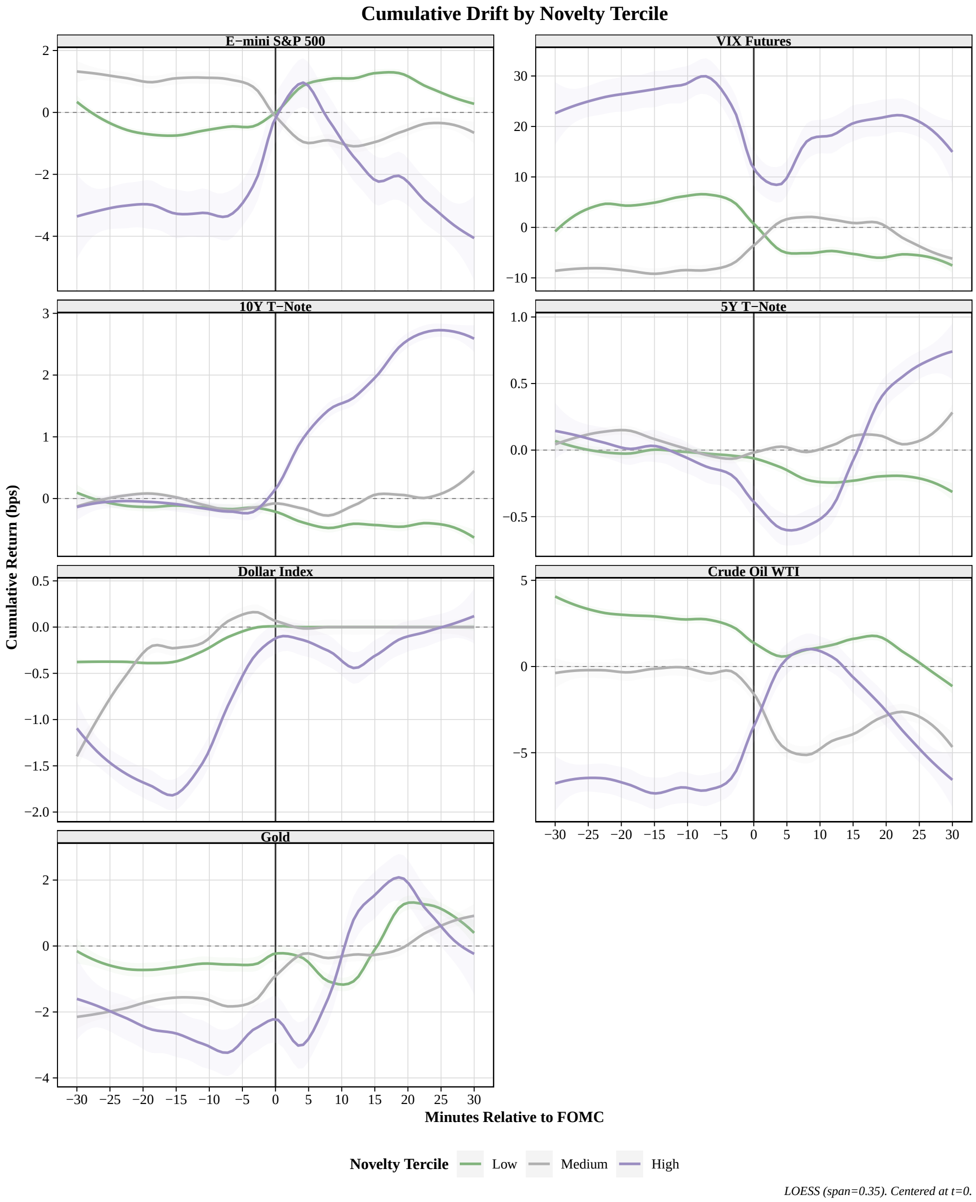


Figure B.7: Cumulative return drift by novelty tercile. Unlike stance-based drift (Figure 3.5), novelty terciles show less directional separation, consistent with novelty affecting volatility rather than returns.

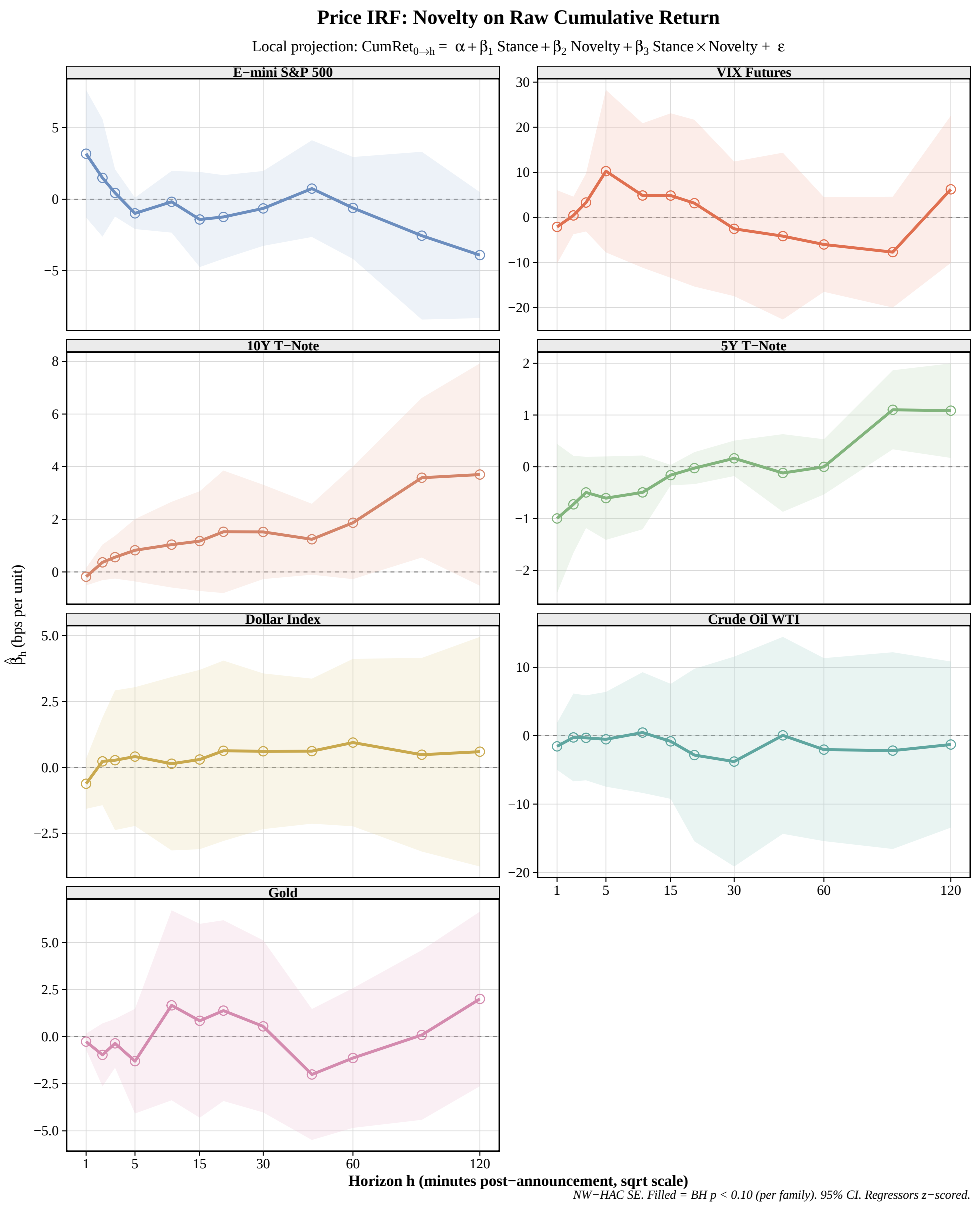


Figure B.8: Impulse response: cumulative return to novelty. Effects are generally smaller and less significant than stance effects (Figure 3.6), supporting the hypothesis that novelty operates through volatility rather than returns.

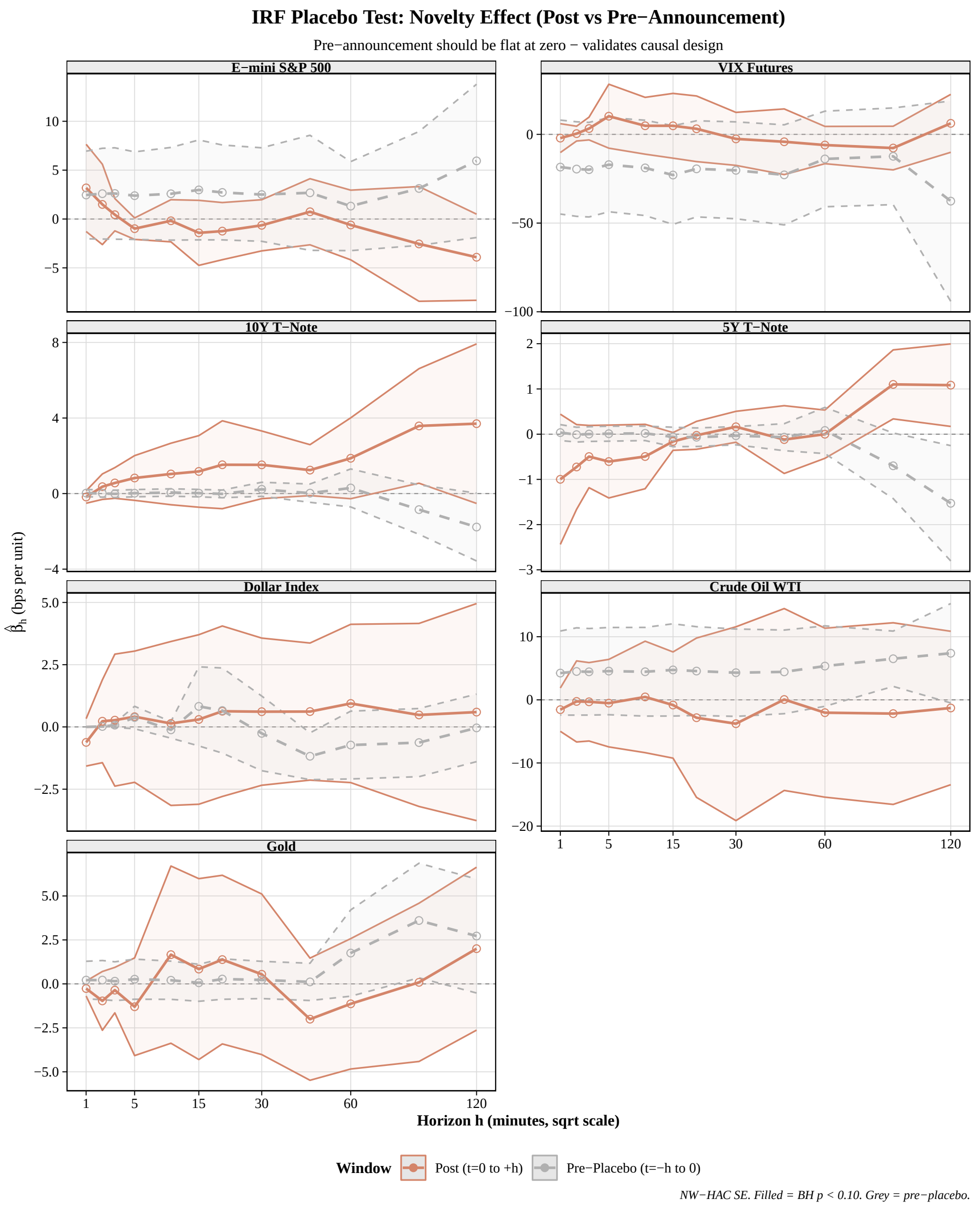


Figure B.9: Pre-announcement placebo: IRF for novelty. All coefficients are near zero, corroborating the placebo results for stance (Figure 3.8).

B.6 Sub-Period Stability

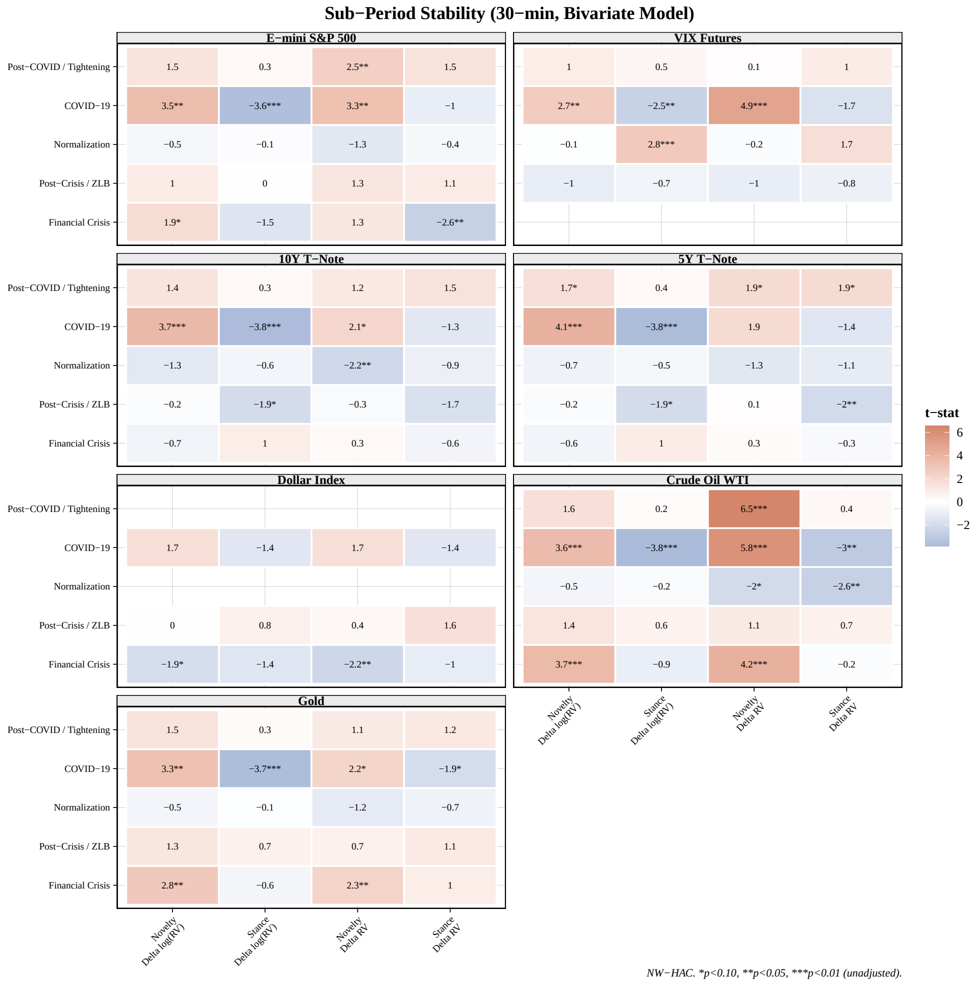


Figure B.10: Sub-period stability: regression coefficients across six Fed policy regimes. Coefficient signs are generally consistent across sub-periods, though magnitudes vary with the degree of policy uncertainty in each regime.

Bibliography

Ackert, L. F. and Y. S. Tian (2000). Arbitrage and valuation in the market for Standard & Poor's depositary receipts. *Financial Management 29*(3), 71–87.

Alquist, R. and O. Gervais (2013). The role of financial speculation in driving the price of crude oil. *The Energy Journal 34*(3), 35–54.

Andersen, T. G. and T. Bollerslev (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance 4*(2–3), 115–158.

Andersen, T. G. and T. Bollerslev (1998). Deutsche mark–dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance 53*(1), 219–265.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. *The Review of Economics and Statistics 89*(4), 701–720.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics 61*(1), 43–76.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2003). Micro effects of macro announcements: Real-time price discovery in foreign exchange. *American Economic Review 93*(1), 38–62.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2007). Real-time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics 73*(2), 251–277.

Apel, M. and M. B. Grimaldi (2012). The information content of central bank minutes. Working Paper Series 261, Sveriges Riksbank.

Araci, D. (2019). Finbert: Financial sentiment analysis with pre-trained language models. *arXiv preprint arXiv:1908.10063*.

Aulerich, N. M., S. H. Irwin, and P. Garcia (2012). Bubbles. In *Food Prices, and Speculation: Evidence from the CFTC's Daily Large Trader Data Files, Paper Prepared for Presentation at the NBER Conference on "Economics of Food Price Volatility" in Seattle, WA*.

Bai, J. and P. Perron (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics 18*(1), 1–22.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics 131*(4), 1593–1636.

Balduzzi, P., E. J. Elton, and T. C. Green (2001). Economic news and bond prices: Evidence from the us treasury market. *Journal of Financial and Quantitative Analysis 36*(4), 523–543.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realized kernels in practice: Trades and quotes. *The Econometrics Journal 12*(3), C1–C32.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B 64*(2), 253–280.

Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics 2*(1), 1–37.

Basak, S. and A. Pavlova (2016). A model of financialization of commodities. *Journal of Finance 71*(4), 1511–1556.

Baumeister, C. and L. Kilian (2014). Do oil price increases cause higher food prices? *Economic Policy 29*(80), 691–747.

Baur, D. G. and B. M. Lucey (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review 45*(2), 217–229.

Ben-David, I., F. Franzoni, and R. Moussawi (2018). Do ETFs increase volatility? *The Journal of Finance 73*(6), 2471–2535.

Benjamini, Y. and Y. Hochberg (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological) 57*(1), 289–300.

Bernanke, B. S. and K. N. Kuttner (2005). What explains the stock market's reaction to federal reserve policy? *Journal of Finance 60*(3), 1221–1257.

Bligh, M. C. and G. D. Hess (2008). The decline of symbolic politics: Federal reserve communications. *Journal of Economic Psychology 29*(4), 535–559.

Blinder, A. S., M. Ehrmann, M. Fratzscher, J. De Haan, and D.-J. Jansen (2008). Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature 46*(4), 910–945.

Boons, M., F. A. de Roon, and M. Szymanowska (2014). The price of commodity risk in stock and futures markets. In *AFA 2012 Chicago Meetings Paper*. Available at SSRN: https://ssrn.com/abstract=1785728 or http://dx.doi.org/10.2139/ssrn.1785728.

Brownlees, C. T. and G. M. Gallo (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis 51*(4), 2232–2245.

Brunetti, C., B. Büyükşahin, and J. H. Harris (2016). Speculators, prices, and market volatility. *Journal of Financial and Quantitative Analysis 51*(5), 1545–1574.

Brunetti, C. and D. Reiffen (2014). Commodity index trading and hedging costs. *Journal of Financial Markets 21*, 153–180.

Brusa, F., P. Savor, and M. Wilson (2015). Asset allocation and fomc announcements. *Review of Financial Studies 28*(5), 1398–1446.

Brusa, F., P. Savor, and M. Wilson (2019). Fomc announcements and market returns: Evidence from the options market. *Journal of Finance 75*(1), 399–441.

Bryant, H. L., D. A. Bessler, and M. S. Haigh (2006). Causality in futures markets. *Journal of Futures Markets 26*(11), 1039–1057.

Büyükşahin, B. and J. H. Harris (2011). Do speculators drive crude oil futures prices? *The Energy Journal 32*(2), 167–202.

Buyuksahin, B. and M. A. Robe (2014). Speculation, commodities and cross-market linkages. *Journal of International Money and Finance 42*, 38–70.

Büyükşahin, B. and M. A. Robe (2014). Speculators, commodities and cross-market linkages. *Journal of International Money and Finance 42*, 38–70.

Campbell, J. R., C. L. Evans, J. D. Fisher, and A. Justiniano (2012). Macroeconomic effects of federal reserve forward guidance. *Brookings Papers on Economic Activity 2012*(1), 1–80.

Cao, F., C.-W. Su, D. Sun, M. Qin, and M. Umar (2024). U.s. monetary policy: The pushing hands of crude oil price? *Energy Economics 134*, 107555.

Carriero, A., G. Kapetanios, and M. Marcellino (2009). Forecasting exchange rates with a large Bayesian VAR. *International Journal of Forecasting 25*(2), 400–417.

Chang, E. C., J. M. Pinegar, and B. Schachter (1997). Interday variations in volume, variance and participation of large speculators. *Journal of Banking & Finance 21*(6), 797–810.

Cheng, I.-H., A. Kirilenko, and W. Xiong (2015). Convective risk flows in commodity futures markets. *Review of Finance 19*(5), 1733–1781.

Cheng, I.-H. and W. Xiong (2014). Financialization of commodity markets. *Annual Review of Financial Economics 6*(1), 419–441.

Chordia, T., R. Roll, and A. Subrahmanyam (2008). Liquidity and market efficiency. *Journal of Financial Economics 87*(2), 249–268.

Corsi, F. (2009). A simple approximate long-memory model of realized volatility. *Journal of Financial Econometrics 7*(2), 174–196.

Da, Z. and S. Shive (2018). Exchange traded funds and asset return correlations. *European Financial Management 24*(1), 136–168.

Da, Z., K. Tang, Y. Tao, and L. Yang (2024). Financialization and commodity markets serial dependence. *Management Science 70*(4), 2122–2143.

Daigler, R. T. and M. K. Wiley (1999). The impact of trader type on the futures volatility-volume relation. *Journal of Finance 54*(6), 2297–2316.

Dannhauser, C. D. (2017). The impact of innovation: Evidence from corporate bond exchange-traded funds. *Journal of Financial Economics 125*(3), 537–560.

Devlin, J., M.-W. Chang, K. Lee, and K. Toutanova (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 4171–4186.

Domanski, D. and A. Heath (2007). Financial investors and commodity markets. *BIS Quarterly Review 3*(1), 53–67.

Ehrmann, M., M. Fratzscher, and R. Rigobon (2011). Global crises and equity market contagion. *Journal of Finance 66*(6), 2597–2649.

Eklund, J. and S. Kim (2024). Fomc statement sentiment and inflation expectations. *Journal of Monetary Economics 141*, 45–62.

Erb, C. B. and C. R. Harvey (2013). The golden dilemma. *Financial Analysts Journal 69*(4), 10–42.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance 25*(2), 383–417.

Fattouh, B., L. Kilian, and L. Mahadeva (2013). The role of speculation in oil markets: What have we learned so far? *The Energy Journal 34*(3), 7–33.

Fishe, R. P. and A. Smith (2012). Identifying informed traders in futures markets. *Journal of Financial Markets 15*(3), 329–359.

Fleming, M. J. and E. M. Remolona (1997). What moves the bond market? *Federal Reserve Bank of New York Economic Policy Review 3*(4).

Fleming, M. J. and E. M. Remolona (1999). Price formation and liquidity in the us treasury market: The response to public information. *Journal of Finance 54*(5), 1901–1915.

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

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The midas touch: Mixed data sampling regression models.

Glosten, L., S. Nallareddy, and Y. Zou (2021). ETF activity and informational efficiency of underlying securities. *Management Science 67*(1), 22–47.

Goldstein, I. and L. Yang (2022). Commodity financialization and information transmission. *The Journal of Finance 77*(5), 2613–2667.

Gorodnichenko, Y., T. Pham, and O. Talavera (2023). The voice of monetary policy. *American Economic Review 113*(2), 548–584.

Gorton, G. and K. G. Rouwenhorst (2006). Facts and fantasies about commodity futures. *Financial Analysts Journal 62*(2), 47–68.

Gromb, D. and D. Vayanos (2010). Limits of arbitrage: The state of the theory. *Annual Review of Financial Economics 2*, 251–275.

Grossman, S. J. and J. E. Stiglitz (1980). On the impossibility of informationally efficient markets. *American Economic Review 70*(3), 393–408.

Gürkaynak, R. S., B. P. Sack, and E. T. Swanson (2005). Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. *International Journal of Central Banking 1*(1), 55–93.

Gürtler, M. and O. Gürtler (2010). The effect of fomc statement language on financial markets. *Journal of Financial Research 33*(4), 369–391.

Haigh, M. S., J. Hranaiova, and J. A. Overdahl (2007). Hedge funds, volatility, and liquidity provision in energy futures markets. *Journal of Alternative Investments 9*(4), 10–38.

Hamilton, J. D. and J. C. Wu (2014). Risk premia in crude oil futures prices. *Journal of International Money and Finance 42*, 9–37.

Hansen, P. R. and A. Lunde (2005). A realized variance for the whole day based on intermittent high-frequency data. *Journal of Financial Econometrics 3*(4), 525–554.

Hansen, S., M. McMahon, and A. Prat (2018). Transparency and deliberation within the fomc: A computational linguistics approach. *Quarterly Journal of Economics 133*(2), 801–870.

Hasbrouck, J. (2003). Intraday price formation in US equity index markets. *The Journal of Finance 58*(6), 2375–2400.

Hayo, B. and M. Neuenkirch (2010). Do federal reserve communications help predict federal funds target rate decisions? *Journal of Macroeconomics 32*(4), 1014–1024.

Hedegaard, E. (2011). How margins are set and affect asset prices. *Job Market Paper*.

Hendershott, T. and R. Riordan (2013). Algorithmic trading and the market for liquidity. *Journal of Financial and Quantitative Analysis 48*(4), 1001–1024.

Henderson, B. J., N. D. Pearson, and L. Wang (2015). New evidence on the financialization of commodity markets. *Review of Financial Studies 28*(5), 1285–1311.

Hollstein, F., M. Prokopczuk, and C. Würsig (2020). Volatility term structures in commodity markets. *Journal of Futures Markets 40*(4), 527–555.

Hu, G. X., J. Pan, and J. Wang (2013). Noise as information for illiquidity. *Journal of Finance 68*(6), 2341–2382.

Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics 3*(4), 456–499.

Irwin, S. H. and B. W. Brorsen (1987). A note on the factors affecting technical trading system returns. *Journal of Futures Markets 7*(5), 591–595.

Irwin, S. H. and B. Holt (2004). The effect of large hedge fund and CTA trading on futures market volatility. In *Commodity Trading Advisors: Risk, Performance Analysis and Selection.* New York, NY: John Wiley and Sons, Inc.

Irwin, S. H. and D. R. Sanders (2011). Index funds, financialization, and commodity futures markets. *Applied Economic Perspectives and Policy 33*(1), 1–31.

Irwin, S. H. and D. R. Sanders (2012). Testing the masters hypothesis in commodity futures markets. *Energy Economics 34*(1), 256–269.

Irwin, S. H. and S. Yoshimaru (1999). Managed futures, positive feedback trading, and futures price volatility. *Journal of Futures Markets 19*(7), 759–776.

Israeli, D., C. M. Lee, and S. A. Sridharan (2017). Is there a dark side to exchange traded funds? An information perspective. *Review of Accounting Studies 22*(3), 1048–1083.

Jarociński, M. and P. Karádi (2020). Deconstructing monetary policy surprises—the role of information shocks. *American Economic Journal: Macroeconomics 12*(2), 1–43.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review 95*(1), 161–182.

Kang, B., C. S. Nikitopoulos, and M. Prokopczuk (2020). Economic determinants of oil futures volatility: A term structure perspective. *Energy Economics 88*, 104743.

Kang, W., K. G. Rouwenhorst, and K. Tang (2020). A tale of two premiums: The role of hedgers and speculators in commodity futures markets. *Journal of Finance 75*(1), 377–417.

Kang, W., K. Tang, and N. Wang (2023). Financialization of commodity markets ten years later. *Journal of Commodity Markets 30*, 100313.

Karali, B. and O. A. Ramirez (2014). Macro determinants of volatility and volatility spillover in energy markets. *Energy Economics 46*, 413–421.

Kilian, L. and D. P. Murphy (2014). The role of inventories and speculative trading in the global market for crude oil. *Journal of Applied Econometrics 29*(3), 454–478.

Kilian, L. and C. Vega (2011). Do energy prices respond to us macroeconomic news? a test of the hypothesis of predetermined energy prices. *Review of Economics and Statistics 93*(2), 660–671.

Koop, G. (2013). Forecasting with medium and large Bayesian VARs. *Journal of Applied Econometrics 28*(2), 177–203.

Kothari, S. P. and J. B. Warner (2007). Econometrics of event studies. pp. 3–36.

Kurov, A., A. Sancetta, G. Strasser, and M. H. Wolfe (2019). Price drift before us macroeconomic news: Private information about public announcements? *Journal of Financial and Quantitative Analysis 54*(1), 449–479.

Kuttner, K. N. (2001). Monetary policy surprises and interest rates: Evidence from the fed funds futures market. *Journal of Monetary Economics 47*(3), 523–544.

Känzig, D. R. (2021, April). The macroeconomic effects of oil supply news: Evidence from OPEC announcements. *American Economic Review 111*(4), 1092–1125.

Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics 4*(1), 25–38.

Liu, L. Y., A. J. Patton, and K. Sheppard (2015). Does anything beat 5-minute RV? A comparison of realized measures across multiple asset classes. *Journal of Econometrics 187*(1), 293–311.

Loughran, T. and B. McDonald (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *Journal of Finance 66*(1), 35–65.

Lucca, D. O. and E. Moench (2015). The pre-fomc announcement drift. *Journal of Finance 70*(1), 329–371.

Madhavan, A. (2012). Exchange-traded funds, market structure, and the flash crash. *Financial Analysts Journal 68*(4), 20–35.

Manela, A. and A. Moreira (2017). News implied volatility and disaster concerns. *Journal of Financial Economics 123*(1), 137–162.

Masters, M. W. (2009). Testimony before the commodity futures trading commission. *Testimony to the Commodity Futures Trading Commission.*

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica 29*(3), 315–335.

Müller, U. A., M. M. Dacorogna, R. D. Davé, R. B. Olsen, O. V. Pictet, and J. E. von Weizsäcker (1997). Volatilities of different time resolutions—analyzing the dynamics of market components. *Journal of Empirical Finance 4*(2-3), 213–239.

Nakamura, E. and J. Steinsson (2018). High-frequency identification of monetary non-neutrality: The information effect. *Quarterly Journal of Economics 133*(3), 1283–1330.

Newey, W. K. and K. D. West (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies 61*(4), 631–653.

O'Hara, M. and X. A. Zhou (2021). Anatomy of a liquidity crisis: Corporate bonds in the COVID-19 crisis. *Journal of Financial Economics 142*(1), 46–68.

Pan, K. and Y. Zeng (2016). ETF arbitrage under liquidity mismatch. *Journal of Financial Economics 120*(3), 617–635.

Patton, A. J. and M. Verardo (2012). Why do markets disagree? evidence from variation in opinion. *Review of Financial Studies 25*(12), 3734–3773.

Petäjistö, A. (2017). Inefficiencies in the pricing of exchange-traded funds. *Financial Analysts Journal 73*(1), 24–54.

Pontiff, J. (1996). Costly arbitrage: Evidence from closed-end funds. *The Quarterly Journal of Economics 111*(4), 1135–1151.

Ready, M. J. and R. C. Ready (2022). Order flows and financial investor impacts in commodity futures markets. *The Review of Financial Studies 35*(10), 4712–4755.

Richie, N., R. T. Daigler, and K. C. Gleason (2008). The limits to stock index arbitrage: Examining S&P 500 futures and SPDRs. *Journal of Futures Markets 28*(12), 1182–1205.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread. *The Journal of Finance 39*(4), 1127–1139.

Rosa, C. (2013). The high-frequency response of exchange rates and interest rates to macroeconomic announcements. *Journal of Banking & Finance 37*(6), 2162–2174.

Savor, P. and M. Wilson (2014). Asset pricing: A tale of two days. *Journal of Financial Economics 113*(2), 171–201.

Schmeling, M. and C. Wagner (2019). Does central bank tone move asset prices? *Review of Finance 23*(5), 933–972.

Scholtus, M., D. Van Dijk, and B. Frijns (2014). Speed, algorithmic trading, and market quality around macroeconomic news announcements. *Journal of Banking & Finance 38*, 89–105.

Shapiro, A. H., M. Sudhof, and D. J. Wilson (2022). Taking the fed at its word: A new approach to estimating central bank objectives using text analysis. *Review of Economics and Statistics 104*(4), 768–784.

Shleifer, A. and L. H. Summers (1990). The noise trader approach to finance. *Journal of Economic perspectives 4*(2), 19–33.

Sims, C. A. and T. Zha (1999). Error bands for impulse responses. *Econometrica 67*(5), 1113–1155.

Singleton, K. (2014). Investor flows and the 2008 boom/bust in oil prices. *Management Science 60*(2), 300–318.

Staer, A. (2017). Asset management via ETFs. *The Review of Financial Studies 30*(9), 3225–3264.

Swanson, E. T. and J. C. Williams (2014). Measuring the effect of the zero lower bound on medium- and longer-term interest rates. *American Economic Review 104*(10), 3154–3185.

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

Todorov, K. (2024). When passive funds affect prices: Evidence from volatility and commodity ETFs. *Review of Finance 28*(3), 831–863.

Veldkamp, L. L. (2011). *Information Choice in Macroeconomics and Finance*. Princeton University Press.

Wang, K., N. Reimers, and I. Gurevych (2021). Tsdae: Using transformer-based sequential denoising auto-encoder for unsupervised sentence embedding learning. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pp. 671–688.

Wongswan, J. (2009). The response of global equity indexes to u.s. monetary policy announcements. *Journal of International Money and Finance 28*(2), 344–365.

Working, H. (1960). Speculation on hedging markets. *Food Research Institute Studies 1*(2), 185–220.

Zhu, Z., L. Sun, J. Tu, and Q. Ji (2022). Oil price shocks and stock market anomalies. *Financial Management 51*(2), 573–612.