UiT The Arctic University of Norway

Faculty of Biosciences, Fisheries and Economics
School of Business and Economics

Market volatility and new evidence from media sentiment

An AI-driven approach

Hooman Abdollahi

A dissertation for the degree of Philosophiae Doctor (PhD) June 2024

Market volatility and new evidence from media sentiment:

An AI-driven approach

By

Hooman Abdollahi

A dissertation for the degree of Philosophiae Doctor

School of Business and Economics

UiT - The Artic University of Norway

Tromsø, 2024

Acknowledgment

The research underlying this thesis has taken place from 2020 to 2024 at UiT-The Arctic University of Norway. I acknowledge the research support from my supervisors, Espen Sirnes and Sturla Lyngnes Fjesme. They have been a source of constant support and a formative influence on my academic career, consistently showing a remarkable commitment to excellence and a high level of professional ethics that have been both inspiring and encouraging. Their guidance and expertise have been invaluable, and I am deeply grateful for their supervision.

I have also benefited from comments, suggestions, and talks on the papers forming this thesis. I would like to express my gratitude to Antonio Gargano, Dan Zhang, Denis Becker, Derek John Clark, Heikki Lehkonen, Jens Kvaerner, Juha-Pekka Junttila, Karen DeVivo, Knut Nygaard, Miguel Karlo Reyes de Jesús, Mikko Moilanen, Øivind Devik Schøyen, Oystein Strom, Stein Østbye, and the seminar participants at the University of Tromsø, Norwegian School of Economics (NHH), Hanken School of Economics, and the University of Jyväskylä.

Thanks are also due to Kamilla Michalsen, Andreas Mikkelsen, and Irena Kustec for having me in their teaching group, which provided an edifying teaching experience. I am also grateful for all the administrative assistance provided by Heidi Hemmingsen, Inger Janne Johansen, Ingjerd Gauslaa Nilsen, Ingrid Dyrkorn Heimland, and Tone Berre. Much appreciation is extended to the Heads of Department, Kåre Skallerud and Mette Talseth Solnørdal, for providing an excellent environment conducive to serious research.

I also thank Amsalu, Babak, Hanna, Marius, Masood, and Samira for their company, as well as my colleagues at HHT.

Finally, I thank my wife, Kaisa, for her continuous support and encouragement. Thanks are also extended to my family and in-laws.

Abstract

This dissertation stands at the intersection of finance and machine learning to empirically delve into the intricate relation between media sentiment and market volatility, covering various aspects of this dynamic. Through three research papers, the thesis presents a detailed investigation into how media—ranging from traditional news to social media platforms—shapes market behaviors and sentiments, influencing market volatility.

In the first paper, we employ a novel approach to quantify media sentiment's impact on market volatility across various financial markets. Using advanced natural language processing techniques, we extract semantic sentiment from news headlines and social media posts. Our findings reveal a time-varying connection between media sentiment and market volatility, indicating how changes in media-driven sentiment can lead to fluctuations in market volatility.

In the second paper, we shift our focus to the predictive power of sentiment in financial markets. We examine the role of media sentiment, derived from both news and social media, in forecasting oil price volatility, given its critical role in the global economy. Employing a hybrid model that integrates sentiment analysis with forecasting models, we reveal the distinct predictive abilities of news versus social media sentiment. The analysis demonstrates that incorporating media sentiment significantly enhances the accuracy of these predictions.

In the third paper, we explore the information transmission pattern from real and fake news to market volatility. Through an advanced analysis of political news stories, we classify news items as either real or fake and then measure their respective impacts on stock market volatility. Our study uncovers the distinct patterns of volatility associated with each type of news, highlighting the more pronounced yet short-lived influence of fake news compared to the sustained but steadier impact of real news.

Keywords: Market and media sentiment, Transmission mechanism, Machine learning in finance

Abstrakt

Denne avhandlingen befinner seg i skjæringspunktet mellom finans og maskinlæring for å empirisk utforske det komplekse forholdet mellom mediesentiment og markedsvolatilitet, og dekker ulike aspekter av denne dynamikken. Gjennom tre forskningsartikler presenterer avhandlingen en detaljert undersøkelse av hvordan medier – alt fra tradisjonelle nyheter til sosiale medieplattformer – former markedsadferd og sentimenter, og til slutt påvirker markedsvolatiliteten.

I den første artikkelen benytter vi en ny tilnærming for å kvantifisere mediesentimentets innvirkning på markedsvolatilitet på tvers av ulike finansmarkeder. Ved å bruke avanserte teknikker for naturlig språkbehandling, trekker vi ut semantisk sentiment fra nyhetsoverskrifter og innlegg på sosiale medier. Våre funn avslører en tidsvarierende forbindelse mellom mediesentiment og markedsvolatilitet, som indikerer hvordan endringer i mediedrevet sentiment kan føre til svingninger i markedsvolatilitet.

I den andre artikkelen skifter vi fokus til den prediktive kraften av sentiment i finansmarkedene. Vi undersøker rollen til mediesentiment, hentet fra både nyheter og sosiale medier, i å forutsi oljeprisvolatilitet, gitt dens kritiske rolle i den globale økonomien. Ved å bruke en hybridmodell som integrerer sentimentanalyse med prognosemodeller, avslører vi de distinkte prediktive evnene til nyhets- versus sosiale mediesentiment. Analysen viser at inkorporering av mediesentiment betydelig forbedrer nøyaktigheten av disse prediksjonene.

I den tredje artikkelen utforsker vi informasjonstransmisjonsmønsteret fra ekte og falske nyheter til markedsvolatilitet. Gjennom en avansert analyse av politiske nyhetshistorier klassifiserer vi nyhetselementer som enten ekte eller falske og måler deretter deres respektive påvirkninger på aksjemarkedsvolatilitet. Vår studie avdekker de distinkte volatilitetsmønstrene assosiert med hver type nyhet, og fremhever den mer uttalte og samtidig kortvarige innflytelsen av falske nyheter sammenlignet med den vedvarende, men jevnere innflytelsen av ekte nyheter.

Nøkkelord: Markeds- og mediasentiment, Transmisjonsmekanisme, Maskinlæring i finans

List of Papers and Contributions

Candidate's name: Hooman Abdollahi

The following articles are included in this thesis:

Paper I: Abdollahi, H., Fjesme, S.L. and Sirnes, E., 2024. Measuring market volatility connectedness to media sentiment. *The North American Journal of Economics and Finance, 71*, p.102091.

Paper II: Abdollahi, H., 2023. Oil price volatility and new evidence from news and Twitter. *Energy Economics, 122*, p.106711.

Paper III: Abdollahi, H., Fjesme, S.L. and Sirnes, E., 2024. Fake news and market volatility: Insights from a large language model (*manuscript*)

Contributions

Contributions/roles	**Paper I**	**Paper II**	**Paper III**
Conceptualization	HA	HA	SF, ES, HA
Data Curation	HA	HA	HA, ES
Methodology	HA, ES	HA	HA
Formal Analysis and Software	HA	HA	HA
Manuscript preparation	HA, SF, ES	HA	HA, ES, SF
Supervision	ES, SF	ES, SF	ES, SF

HA = Hooman Abdollahi ES = Espen Sirnes SF = Sturla Lyngnes Fjesme

Conflict of interest: None.

Table of Contents

1. Introduction

The progress in artificial intelligence (AI) and natural language processing (NLP) has opened new avenues to explore the role of media sentiment in market volatility. This thesis aims to investigate the dynamic nature of the relation between textual media sentiment and market volatility using advanced NLP and machine learning techniques. By uncovering the nuanced influence of media sentiment on market behavior, this research has the potential to benefit investors with more informed decision-making, policymakers with tools to navigate market complexities, and financial analysts with the development of more accurate forecasting models.

Literature theorizes aspects of media's impact on the market. Grossman and Stiglitz (1980) note that perfect information efficiency in markets is unattainable due to the presence of noise traders—investors who make decisions based on erroneous or irrelevant information. The noise trader theory provides a foundation for understanding how media-driven sentiment, often viewed as noise within the context of market fundamentals, can impact financial markets. It emphasizes the role of information asymmetry, where not all market participants have equal access to or interpretation of information, thus introducing volatility unrelated to fundamental values. Shiller (2000) points out that newspapers and, by extension, other forms of media have historically played a key role in the emergence of speculative bubbles. He suggests that the narratives and sentiments disseminated through media channels can amplify investor reactions, creating a conducive environment for the formation and bursting of speculative bubbles. Enhancing this theoretical framework, the limited attention theory provides insights into how investors process information. This theory posits that investors have limited cognitive resources for information processing, causing them to prioritize certain news over others based on salience. This selective attention mechanism, heavily influenced by media, can significantly affect market volatility as investors react to the influx of information with varying degrees of credibility.

Empirical evidence further underscores the significant role of media in the dynamics of markets. Niederhoffer (1971) examines the association between investment-related news and market behavior. He finds a reactionary response by market participants to the news, indicating that positive news often leads to temporary price increases while negative news triggers short-term price falls. Tetlock (2007) explores the relation between media coverage and stock market movements, finding that media pessimism—proxied by the prevalence of negative words in news articles—exerts downward pressure on stock prices. The focus on the relation between media and market dynamics then advances towards the details within the media. Goldman et al. (2022) find that the quality and integrity of journalism can significantly influence market efficiency and stability, highlighting the latent interplay between media content and financial market responses.

Within the finance literature, studying media's impact on the market has been primarily built upon non-econometric tools for sentiment analysis. Central to this methodical approach is textual analysis, which processes media content into numerical indices prepared for financial analysis. Commonly, finance literature employs dictionary analysis, focusing on word count to gauge sentiment. However, this method, 'a passive collection of words' as described by Shllier (2017), falls short of capturing the sentiment truly embedded in the text. Beyond facing technical challenges in language processing, it primarily reveals only the surface level of content, neglecting the deeper psychological and emotional factors that significantly impact its influence. Shiller (2020) posits the idea of adopting more sophisticated approaches, such as semantic sentiment analysis, which reveals the meaning and psychological significance behind words. This advancement allows researchers to go beyond the collection of words and capture how the meaning of the content shapes emotions, perceptions, and ultimately influences market behavior. The advent of large language models (LLMs) presents an opportunity to explore this innovative approach further.

The concept of the availability heuristic, introduced by Tversky and Kahneman (1973), proposes that novel information is weighted more heavily by investors, and this inherently implies a time-varying effect where the impact of media sentiment on market volatility diminishes or strengthens over time. However, evidence detailing the characteristics of this time-varying impact of media on the market remains scarce as prior research has predominantly focused on media's overall impact on the market. The advent of methodological advancements enables the exploration of how media information is transmitted to the market over time. Additionally, Bae et al. (2003) explore the concept of cross-regional contagion, which posits that events in one region can affect the likelihood of similar responses in other markets. This introduces a potential geographical dimension to the media's impact that may illuminate how markets respond to news across national borders. Although Engelberg and Parsons (2011) examine how local news directly influences local stock markets, the cross-market relation is yet to be further examined.

Recent advancements in computational techniques allow scholars to further quantify the informational value media brings to financial markets. Audrino et al. (2020) demonstrate that media sentiment variables can improve stock volatility forecasts marginally. The modest enhancement may be attributed to the limitations of their sentiment analysis model and the basic nature of their forecasting approach. Such findings indicate the potential for revisiting this area, especially in light of the significant progress in NLP and predictive machine learning algorithms. Howard and Ruder (2018) emphasize the substantial improvements in sentiment analysis achieved through language models, while Timmermann (2006) highlights the superior performance of hybrid forecasting models that combine multiple predictive algorithms to more accurately capture a wider range of time series dynamics and reach a higher quality prediction. This evolution underlines the importance of re-evaluating the informational value of both traditional and social media in the financial forecasting realm, indicating a crucial area for future research in understanding their distinct impacts on financial markets.

In the discourse on the media's impact on financial markets, it is also essential to understand the diverse nature of media, which finance literature has often regarded as a single, unified entity. Recent studies, such as Dougal et al. (2012), delve into the complexities of media, focusing on how specific aspects like media slant can directly impact market dynamics. Additionally, the rising concern about the spread of fake news and its not-yet-fully-understood effects on market behavior represents a significant field for further scrutiny. Clarke et al. (2021) highlight the efficacy of machine learning tools in detecting fake news. This detailed view of media emphasizes the necessity for more in-depth analysis to thoroughly grasp the intricate ways in which media influences financial markets.

Building upon these theoretical concepts and empirical evidence, this dissertation aims to shed light on the complex mechanisms by which media sentiment affects market volatility. It not only contributes to the academic discourse on finance, but also provides insights for investors and market participants. Therefore, this dissertation poses the following primary and secondary research questions:

Primary research question:

- How can advancements in machine learning be employed to deepen our understanding of the relation between media and market volatility?

Secondary sub-questions:

- In what ways does the quantification of media sentiment enhance our comprehension of its time-varying impact on market volatility across different financial sectors?
- What role does media sentiment assume in the predictive modeling of market volatility, and how do the different media types (traditional vs. social) contribute to this process?
- With an increasing focus on the spread of fake news, how does fake news affect market volatility compared to real news, and what distinguishes the impacts of these two news types?

2. Data and Methods

2.1 Methodical framework

This thesis employs a suite of advanced computational techniques to analyze the influence of media sentiment on market volatility. The methodical approach is broadly categorized into two phases: data generation and econometric analysis. In the data generation phase, NLP techniques are used to extract sentiment from media sources. The extracted sentiment data is then employed in the econometric analysis phase to assess its influence on market volatility. The details of this methodological framework are further elaborated upon in each individual paper.

2.1.1 Natural language processing

Central to our NLP approach is the application of the Bidirectional Encoder Representations from Transformers (BERT), a pre-trained machine learning model proposed by Devlin et al.

(2018) and developed by Google for textual analysis. The BERT model, known for its depth and complexity, is employed to classify, interpret, and quantify sentiment from textual news and social media content.

For sentiment analysis, BERT is fine-tuned using the Financial PhraseBank developed by Malo et al. (2014)—a collection of sentences from financial news labeled by finance professionals. This fine-tuning process customizes BERT's capabilities to the finance domain, enabling a better understanding of media sentiment. Sentiment scores are derived by calculating the difference between positive and negative prediction probabilities (logits), providing an index of sentiment from -1 (most negative) to 1 (most positive).

For news classification, a process that distinguishes between fake and real news, we primarily employ the dataset by Horne and Adali (2017). This dataset includes labeled examples of real and fake news, serving as the primary source for training the BERT model. This training equips the model to recognize typical stylistic characteristics of fake news, thereby enabling the classification of news articles based on a set of empirically validated journalistic criteria. Classifying news as either fake or real allows for a detailed analysis of how various types of media content impact market behavior.

2.1.2 Econometric framework

To explore the directional connectedness from media sentiment to market volatility, the Diebold and Yilmaz (DY) connectedness framework is employed (Diebold and Yilmaz, 2012). This framework, based on vector autoregression (VAR) and generalized variance decomposition, quantifies the magnitude and direction of sentiment's influence on market volatility. This econometric approach facilitates a granular analysis of how information and sentiment propagate through financial markets over time.

We also employ a diverse set of methods including Generalized Autoregressive Conditional Heteroskedasticity (GARCH) (Bollerslev, 1986), Bidirectional Long Short-Term Memory (BiLSTM) (Hochreiter and Schmidhuber, 1997) models, and Ordinary Least Squares (OLS) regression among others to rigorously analyze the dynamics between media sentiment and market volatility. Each method is detailed extensively in their respective papers, underlining their application and efficacy in uncovering the nuanced relationships within our dataset.

2.2 Data

The dissertation utilizes an extensive dataset encompassing market data and media content related to several major financial markets and commodities, including the US (the Dow Jones Industrial Average (DJIA)), the UK (the Financial Times Stock Exchange (FTSE)), France (the CAC 40), Germany (the DAX 30), Japan (the Nikkei), and Brent crude oil. These asset markets are selected for their global significance and the rich media coverage they receive, allowing for a comprehensive analysis of sentiment's effects.

2.2.1 Market data

Market data, including daily closing, minimum, and maximum prices, are sourced from Refinitiv Eikon and investing.com from 2014 to 2022. Market volatility is computed using the Parkinson (1980) range-based measure, which offers a detailed reflection of market movements and the potential influence of daily information.

2.2.2 Unstructured data

Media content, including news stories and tweets, is collected from investing.com and Twitter, respectively. The choice of investing.com as a news source is justified by its aggregation of content from reputable outlets like Reuters and Bloomberg, ensuring the reliability of the data. Tweets related to the selected markets are gathered using relevant hashtags. The period of study, from 2014 to 2022, encompasses significant global events, providing a rich context for analyzing sentiment's impact on markets.

2.3 Ethical considerations

The research adheres to ethical standards in data collection and analysis. Publicly available data sources are used, eliminating the need for individual consent. However, care is taken to ensure the confidentiality and anonymity of any potentially identifiable information within the dataset. The use of machine learning models and sentiment analysis is conducted with an awareness of their limitations and potential biases, and efforts are made to mitigate these through rigorous model validation and testing.

3. Results

3.1 Paper I: Measuring Market Volatility Connectedness to Media Sentiment

The first paper presents new evidence on how sentiment expressed in media, encompassing both news stories and tweets, is transmitted to financial markets over time. Using advanced analytical techniques, this study pioneers in quantitatively analyzing the relation between media sentiment and market behavior.

The empirical framework initiates with semantic sentiment analysis using the BERT model to build exclusive sentiment indices for selected markets. This innovative approach offers a refined measure that captures the emotional and psychological undertones of textual media. We then employ the DY connectedness framework to analyze how market volatility is connected to the evolution of media sentiment across selected major financial markets and commodities. This comprehensive approach, combining sophisticated sentiment analysis with robust econometric models, enables a deeper examination of sentiment's impact on market dynamics.

The findings reveal that media sentiment exerts a significant, albeit time-varying, influence on market volatility. Particularly, the study documents a pronounced effect of sentiment during major events, characterizing the *spiky* nature of this impact. Notably, we also find a long-lasting connection between media sentiment and market volatility, indicating that the influence of media extends beyond immediate attention. Furthermore, cross-market analysis illustrates that sentiments related to one market can influence the volatility of others, highlighting the global interconnectedness of financial markets. These insights emphasize the critical role of media in shaping financial market dynamics.

In conclusion, this paper advances our understanding of the dynamic interplay between media sentiment and financial markets. It contributes to the existing literature by demonstrating the time-varying and spiky pattern of media sentiment's impact and its cross-market transmission effects on volatility. Moreover, this study provides insights for investors and market participants, underscoring the importance of considering media sentiment in financial decision-making and forecasting strategies. This work lays the groundwork for future explorations, particularly emphasizing the need to scrutinize the transmission patterns of sentiment's impact on market volatility. Additionally, the discovery of a long-lasting linkage between media sentiment and market movements suggests that media sentiment holds potential informational value for market forecasting. The findings and methodologies introduced here are further explored and built upon in subsequent papers, aiming to deepen our understanding of these dynamics.

3.2 Paper II: Oil Price Volatility and New Evidence from News and Twitter

This paper assesses the predictive power of media sentiment, derived from both traditional news sources and a social media platform (Twitter). The study is motivated by the critical role oil plays in the global economy, where fluctuations in oil prices can have far-reaching implications. Recognizing media sentiment as a potential signal for market movements, we aim to integrate sentiment analysis into volatility forecasting models, thereby enhancing the accuracy and reliability of these predictions.

We hypothesize that news and Twitter tweets enhance the forecasting power given their long-lasting connection with oil price volatility. Additionally, the Twitter sentiment, given its immediacy and widespread user engagement, might outperform traditional news in forecasting oil price volatility. This assumption is based on the idea that social media can capture real-time public sentiment and market reactions more swiftly than conventional news outlets, potentially offering a more accurate reflection of market dynamics.

To test these hypotheses, the study uses oil-related sentiment data from Paper I, namely a dataset of 14,833 news headlines and 116,463 tweets related to the oil market. We then use sentiment indices to feed a hybrid forecasting model combining GARCH with BiLSTM to forecast oil price volatility.

The analysis yields interesting results: incorporating sentiment indices into the forecasting model considerably enhances its accuracy, affirming the substantial informational value of media sentiment. However, contrary to our initial anticipation, it is the sentiment derived from traditional news sources, not Twitter, that significantly outperforms in forecasting oil price volatility. This revelation highlights the intricacy of media influence on financial markets, suggesting that the depth and comprehensiveness of news content offer more profound market insights than public interpretations on social media platforms.

In conclusion, this paper underscores the enduring relevance and superiority of traditional news sentiment in enhancing the predictive accuracy of oil price volatility models. The superior performance of news sentiment highlights its relevance and reliability as a forecasting tool, suggesting that investors and market analysts should prioritize it in their analytical frameworks. This paper serves as a foundation for future research to further explore and exploit the informational content of media, paving the way for the development of more sophisticated and accurate forecasting models that adeptly incorporate the signals of refined sentiment.

3.3 Paper III: Fake News and Market Volatility: Insights from a Large Language Model

This paper delves into the relation between the dissemination of fake and real news and its subsequent impact on financial market volatility. It employs advanced NLP techniques to distinguish between fake and real news, further analyzing their sentiment and how each contributes to market dynamics. The research is crucial, particularly in a period where the internet has exponentially increased the spread of fake news, posing significant challenges and implications for market participants and the broader financial ecosystem.

We initially hypothesize that fake news has no impact on market volatility when compared to real news. This assumption is grounded in the belief that rational market participants are capable of discerning the credibility of news, thereby minimizing the potential effects of fake news on market behavior. Expanding on this foundation, the second hypothesis suggests that while fake news may not significantly affect market volatility over the long term, it can nonetheless induce short-term market reactions. These transient effects are attributed to the initial shock and uncertainty that fake news might introduce to the market before rational evaluation and discounting take place. Focusing on news tone and based on the established literature, we also hypothesize that the presence of negative and weak modal words in news significantly influences market volatility, highlighting the role of linguistic cues in shaping market responses.

We collect a large dataset of 25,444 political news stories and employ the BERT model for fake news detection and sentiment analysis. This study constructs numerical sentiment indices for both fake and real news. These indices are analyzed within the DY framework to measure the directional connectedness from news sentiment development to market volatility, providing a detailed understanding of how information flows influence the market.

The findings reveal that real news has a sustained and significant effect on market volatility, while fake news impacts are transient and less significant. The analysis highlights that real news contributes to a continuous, wavy pattern of volatility spillovers, contrasting with the spiky, short-lived shocks induced by fake news. This differential impact underscores the importance of content credibility and the distinct ways investors process real and fake news. Additionally, we find that other tonal characteristics in fake news—namely, positive, litigious, and constraining tones—significantly influence market volatility, offering insights into how linguistic features of news content can affect financial markets.

In conclusion, this research contributes to the understanding of how fake and real news differently influence financial markets. It indicates the resilience of markets in filtering out the noise introduced by fake news, while also highlighting the lasting importance of credible, real news in informing investors. These findings lay the groundwork for future research into integrating news sentiment analysis into investment strategies.

4. Discussion

4.1 Interconnectedness of findings

The thesis collectively advances the understanding of how external information sources—namely, traditional news, social media, and the credibility of information (fake vs. real news)—influence market volatility. A crucial aspect of our findings makes a connection between information diffusion theory and information cascades to explain the link between media sentiment and market volatility.

Information diffusion theory describes how information spreads through a population over time, focusing on the mechanisms driving information flow and adoption (Rogers, 1962; Rogers et al., 2014). We use news outlets and social media platforms, which act as key channels for information diffusion (Valente, 1996). Concise news stories by the press and rapid information dissemination on social media can create strong emotional responses, influencing investor sentiment without necessarily providing in-depth analysis. This creates an environment ripe for information cascades.

Information cascades are a specific type of information diffusion phenomenon, where people base decisions on observed sentiment-driven information rather than solely on their own private signals. Similar to the way Bikhchandani et al. (1992, 1998) illustrate in their works on cascade formation, this can lead to cascades where everyone makes the same decision, regardless of their private signals' strength. For instance, positive news and social media posts can create an illusion of widespread confidence, leading to a buying frenzy. Conversely, negative sentiment can trigger a cascade of selling.

Additionally, the use of advanced NLP techniques and large datasets in this thesis probes into the modern-day complexity of information cascades. It is not just about the actions of others anymore; it is also about the sentiments and opinions expressed across media platforms. The rapid dissemination and accessibility of information via social media and news outlets have amplified the speed at which information cascades can occur. This phenomenon aligns with the concept of information herding explored within the field of behavioral finance (Hirshleifer, 2015).

A notable finding comes from our analysis of static connectedness, highlighting how different levels of source authenticity influence the magnitude of the cascades they initiate. For instance, Paper I compares market connectedness to news outlets and social media. Here, the credibility of the information source becomes a crucial factor. We find that most of the time, markets exhibit a higher connectedness to news outlets than social media. We interpret this as markets, in their collective decision-making process, placing more weight on the authenticity of the source—a critical consideration within the framework of information cascades. This finding aligns partly with Yu et al. (2023), who suggest that information credibility can have differing effects on asset returns based on investor types. While we focus on market volatility connectedness, Yu et al. (2023) explore how credibility impacts returns on assets preferred by experts versus those favored by gamblers. However, the role of information credibility in influencing market dynamics is found to be significant in both studies, highlighting its importance for understanding market behavior.

Our findings solidify the role of information diffusion in explaining market volatility. By highlighting the influence of media sentiment via news and social media, we extend the theory to encompass the role of media channels in shaping information flow within financial markets. These channels not only transmit information, but also influence investor sentiment through emotional language and social contagion effects. This reinforces the notion that information adoption is not solely based on rational analysis but is also swayed by the collective sentiment expressed through various media channels.

Papers II and III further support this notion. Paper II highlights the differential impact of news sentiment compared to Twitter sentiment on market volatility forecasting, again suggesting that source credibility plays a role in market decision-making. Similarly, Paper III finds that market volatility exhibits a long-lasting connection to real news, while connectedness to fake news is negligible over the entire sample.

Our findings indicate that media sentiment can create cascades by fostering an illusion of widespread confidence. In Paper II, we delve deeper into how this information can be utilized for practical purposes. In this context, we might expect the confidence bias proposed by Griffin and Tversky (1992) to play a role. This theory suggests a tendency to focus more on the *strength* of evidence (here frequency or reportage of information) than its *weight* (here source credibility). While social media might offer a higher frequency of information, news outlets typically carry more weight due to their credibility. Contrary to this bias, our finding reveals that, in the context of market volatility, especially regarding oil prices, the market collectively places greater value on the weight of evidence. This observation diverges from the expected patterns of confidence bias, indicating that, at the market level, the credibility and refinement of information from regular media overshadow the voluminous but less reliable streams from social media.

It is important to acknowledge that confidence bias is a well-established theory, but it might not fully capture the complexities of investor decision-making in the oil market. One possible explanation for the divergence from the expected bias is the role of professional investors. These actors likely possess the resources and expertise to critically evaluate information from various sources, prioritizing credibility over mere frequency. Additionally, the specific context of the oil market, which relies heavily on established information channels, might play a role.

We delve into another crucial aspect: the media's time-varying influence on market volatility, analyzed in Paper I's dynamic analysis section. Our findings offer valuable insights into how market sensitivity to news evolves over time. The information transmission pattern from media sentiment development to market volatility—how it grows, peaks, and fades—offers an analogy to the epidemic pattern of narrative spread described by Shiller (2020). Initially, a news story might gain traction, leading to a rising connectedness with market volatility. However, as the story unfolds and more information emerges, its impact weakens, mirroring the fading punch of an epidemic narrative. Market sensitivity to the story eventually subsides.

Further analysis of the time-varying connectedness to media sentiment reveals a narrative dimension to some market-moving news events. Shiller (2020) argues that narratives with strong internal consistency and alignment can significantly influence market behavior, especially when they resonate with investor experiences and expectations. For example, consider the narrative surrounding then-US President Donald Trump and his perceived influence on the stock market. When the market went up, Trump would often make comments or tweets. However, when the market plunged, his silence fueled a narrative suggesting he was somehow responsible for the downturn. News stories based on this narrative emerged in February 2018, grabbing attention and further impacting the market. This connection highlights the importance of narrative coherence and alignment in determining the impact of news on market volatility, indicating how coherent stories drive market movements.

Paper III further refines this understanding by examining the differential effects of real and fake news sentiment on volatility. The findings reveal distinct information transmission patterns. Fake news sentiment exhibits a more transient connection with market volatility, spiking briefly upon its dissemination before fading over time. Here, the concept of limited attention theory (Kahneman, 1973; Fiske and Taylor, 1991; Hirshleifer and Teoh, 2003) becomes crucial. Investors, with finite cognitive resources, struggle to process all available information and may initially react strongly to attention-grabbing fake news. However, as the stylistic embellishments wear off and more information emerges, the influence of fake news diminishes. Real news sentiment, on the other hand, demonstrates a more persistent, wavy pattern of connection with market volatility. This reflects the gradual, yet constant, dissemination and processing of credible information by the market. Investors may take more time to integrate real news into their decision-making, leading to a more sustained impact on market volatility.

These findings refine the understanding of information cascades within financial markets by highlighting the interplay between source credibility and the time-varying nature of media influence. Traditionally, information cascades were viewed as a binary phenomenon—triggered or not triggered (Bikhchandani et al., 1992). Our findings suggest a more nuanced view. Credible sources, like reliable news, can trigger long-lasting cascades with a sustained impact on market volatility. Conversely, cascades initiated by less credible framing, like fake news, can be transient, spiking briefly before fading away as the limited attention of investors shifts focus. This highlights the importance of considering both the credibility and the evolving nature of media sentiment when analyzing information cascades in financial markets.

Furthermore, this finding emphasizes the importance of considering the duration of these cascades based on information source credibility. Behavioral finance research suggests that investors exhibit limited memory (Hirshleifer, 2015), meaning they tend to overweight recent information when making decisions. This can lead to biases such as *overreaction to news* or *neglect of base rates*. In the context of information cascades, limited memory implies that investors may be more susceptible to attention-grabbing stories (here fake news) in the short term. However, as the initial excitement fades and new information emerges, these short-lived bursts of sentiment are unlikely to have a lasting impact on market behavior. This aligns with our observation of transient cascades triggered by fake news.

Paper III also reveals how fake news is perceived in financial markets. Fake news employing a specific language style mimics credibility through advanced sentence structures. During periods of heightened market sensitivity to fake news, these articles are characterized by their utilization of litigious, constraining, positive, and negative tones, which our analysis has found to be statistically significant in influencing market volatility. While the positive tone in fake news has a more pronounced impact on market dynamics than the negative tone, the litigious and constraining tones underscore the perceived credibility and potential implications of the information presented. These insights into the linguistic strategies deployed in fake news further refine our understanding of the interplay between news dissemination and market behavior.

4.2 Main contributions

The thesis collectively advances our understanding of the directional relation from textual media to market dynamics and volatility. Specifically:

- We employ advanced NLP techniques like BERT to extract semantic sentiment indices from media content. This innovative approach offers a groundbreaking understanding of the psychological impact of news sentiment on market dynamics (Shiller 2017, 2020). In contrast to prior studies that have primarily focused on the static relationship between media coverage and market volatility, such as Hsu et al. (2021), Fraiberger et al. (2021), and Behera and Rath (2022), our analysis also delves into the dynamic nature of this connection. The analysis of time-varying connectedness provides valuable insights into how market sensitivity to media sentiment evolves over time. We uncover the patterns through which media sentiment is transmitted to market volatility. This furthers our understanding of how news stories, with their inherent narrative structures, shape market behavior.
- Through the innovative use of sentiment analysis in market volatility prediction, we contribute to the literature in comparing the effects of traditional news and social media (Twitter) on oil price volatility forecasting. We show that a news-based sentiment index significantly enhances forecasting accuracy over a Twitter-based sentiment index, highlighting that the market places greater weight on the verified information from traditional news outlets compared to the potentially less-reliable sentiment from social media. This partly contrasts with Audrino et al. (2020), who report that sentiment is of marginal informational value for forecasting. We show that media sentiment can provide substantial signals when used in an efficient forecasting system.
- We make a notable contribution by formally differentiating the effects of fake and real news on market volatility. We empirically uncover the distinct transmission patterns of

each, providing concrete evidence that financial markets can discern and react differently to these types of information. Unlike previous studies that employ a form of fact-checking (e.g., Kogan et al., 2019, 2023; Arcuri et al., 2023), we identify fake news through validated stylistic features established in journalism and NLP literatures (see Horne and Adali, 2017; Damstra et al., 2021; Clarke et al., 2021). Our analysis of fake news reveals a significant influence of positive tone, which defies the established literature (e.g., Tetlock, 2007; Ahern and Sosyura, 2015). By delineating the characteristics that make fake news fleeting in its market impact, we contribute to the broader discourse on market dynamics, investor psychology, and information credibility in the digital age.

4.3 On media and natural language processing in the finance domain

The importance of media in finance arises from the fact that it plays a role in reducing the cognitive load of information processing on consumers, including investors. Tuchman (1978), in his classic book on media theory, explains how well-crafted news provides pre-analyzed and contextually rich information, reducing the cognitive burden on the audience who might otherwise struggle to process raw data, information, or events. By presenting information in a digestible format, news narratives can influence the decision-making processes of agents, which is crucial in the fast-paced environment of financial markets.

The significant role of media in finance became more pronounced as studies began to recognize its impact on investor perceptions and market trends. Shiller (1989) discusses how news media can drive market volatility. His work suggests that media coverage can amplify market trends by shaping investor expectations and behavior. Boudoukh et al. (2004) introduce quantitative techniques for analyzing financial news and its impact on stock prices. Their work demonstrates that specific words and phrases in news articles can be systematically linked to market movements. Tetlock (2007) made a major contribution by applying sentiment analysis to financial news. Using the tonal word lists from the psychological Harvard-IV dictionary, he shows that the media tone, particularly the use of negative words in news, can predict stock market declines. This study highlights the importance of textual analysis in understanding how media sentiment affects market behavior. However, the underlying mechanism to measure the media sentiment remained an inchoate word frequency metric based on word lists of general dictionaries. A major problem was that these dictionaries do not contain financial jargon and do not list the words based on their financial implications, which could reduce sentiment analysis accuracy. Later, Loughran and McDonald (2011) develop a finance-specific dictionary to improve the accuracy of sentiment analysis within the finance context. Their work emphasizes that traditional sentiment analysis tools often misinterpret financial jargon, leading to inaccurate predictions.

Although these efforts significantly improved media sentiment analysis within the finance domain, the word metric methods are inherently fallible due to their lack of capability in capturing nuances in natural language, such as modifiers, negations, out-of-the-box vocabulary, and more. However, advances in computational methods and machine learning have allowed for more sophisticated textual analyses. Later, deep learning techniques, such as convolutional neural networks, were employed to build small local language models, representing a significant progress as this practice elevated the textual analysis from word level to sentence level. Malo et al. (2013) provide the foundation for such analysis by demonstrating its superior performance compared to dictionary-based analysis. However, the high level of technicality and calibration required for the successful implementation of these models obstructed their wide adoption among finance scholars. Interestingly, the advent of LLMs in recent years has made it much easier for financial researchers and institutions to use NLP analysis.

In the 2020s, the integration of AI has further transformed textual analysis of media in finance. This period also saw an increased emphasis on the use of alternative data sources, such as social media and blogs, to capture a broader range of market sentiments. For instance, the constituent papers of this thesis employ an LLM to analyze large volumes of news articles and social media feeds, achieving new insights into studying market behavior. The development of large-scale pre-trained language models, such as BERT and the Generative Pre-trained Transformer (GPT), has transformed NLP. These models are trained on vast amounts of data and can be fine-tuned for specific tasks, achieving unprecedented levels of accuracy and fluency. Their ability to understand and generate human-like text has opened new avenues for research and application in finance. LLMs can analyze vast volumes of financial news, social media posts, and other textual data in real-time, providing insights that were previously unattainable. This capability allows for more accurate sentiment analysis, trend detection, and risk assessment.

Furthermore, NLP has the potential to extend the research to other areas such as financial technology. For instance, smart agents, powered by advanced NLP and AI technologies, represent the next frontier in automated financial services. These agents can engage in more complex interactions with users, providing real-time advice and executing trades on their behalf. The development of these smart agents will benefit significantly from advancements in NLP. Techniques such as sentiment analysis, entity recognition, and contextual understanding will enable these agents to process and interpret financial news, social media, and other data sources in real-time. This capability is crucial for making timely investment decisions, as market conditions can change rapidly.

Future research will also focus on improving the interpretability and transparency of NLP models, ensuring they can be trusted in critical financial decision-making processes. For instance, robo-advisors have emerged as a major innovation in financial technology, using AI algorithms and NLP to provide automated financial advice to investors. These platforms analyze large volumes of financial data and user inputs to offer personalized investment recommendations. They utilize NLP to interpret and analyze user queries and provide relevant and timely advice. For instance, an investor might ask a robo-advisor about the implications of recent market volatility on their portfolio. The NLP system would then process the user's query, analyze current market data, and generate an appropriate response. However, the complexity of financial language and the need for precise interpretation highlight the necessity of improving NLP models' accuracy and contextual understanding.

As NLP models become more integral to financial decision-making, enhancing their interpretability and transparency is crucial. For instance, users must trust that the recommendations provided by robo-advisors are based on sound analysis and free from biases. This transparency is vital in the finance domain, where the stakes of investment decisions are high. Moreover, regulatory bodies such as the Securities and Exchange Commission (SEC) are increasingly scrutinizing the algorithms used by financial institutions to make sure they are fair and transparent (SEC, 2023). Ensuring that NLP models used in robo-advisors meet these regulatory standards will potentially be a focus of future research.

In conclusion, the evolution of textual analysis, from simple word counting to sophisticated deep learning models, has had a profound impact on financial research. As these technologies continue to advance, their application in finance will become increasingly powerful, offering new tools and methodologies for understanding and navigating not only media sentiment but also the complexities of financial markets.

4.4 Implications, limitations, and further research

The implications of this thesis extend beyond the academic realm, offering insights for practitioners and policymakers. Our research sheds light on the profound influence of media sentiment on market dynamics, highlighting the necessity for a critical approach to information source credibility. This need is especially pressing in the digital age, where information can be both informative and misleading. Understanding how fake news exploits language to manipulate sentiment offers policymakers valuable tools for navigating the complexities of market movement. These insights can also inform the development of tools for detecting and mitigating the impact of fake news on markets.

Moreover, the demonstrated superiority of news sentiment analysis in forecasting market volatility presents a compelling case for its incorporation into existing financial models. Integrating sentiment indices derived from trusted news sources could enable financial analysts and investors to achieve enhanced predictive accuracy, leading to more informed and strategic decision-making processes.

However, the necessity to streamline our expansive dataset in Paper I—due to the limited availability of related datasets and the computational demands of processing such large volumes of data—suggests the specificity of our findings to certain markets and news categories. Therefore, future research should extend the investigation to explore how these insights apply across different markets and asset classes, strengthening the robustness and generalizability of our conclusions.

Furthermore, the dynamic interplay between media sentiment and investor behavior invites a more interdisciplinary approach to research. There is a burgeoning need for studies that integrate insights from behavioral finance, data science, and information technology. These collaborative endeavors hold promise to yield more comprehensive models of investor behavior, capturing the multifaceted influences of media sentiment in the digital age.

Looking ahead, future research avenues are rich with potential. These include expanding the scope to emerging social media platforms, investigating the interplay between algorithmic trading and sentiment-driven market movements, and delving deeper into the psychological mechanisms behind investor responses to news. Addressing these themes will contribute to creating more resilient financial ecosystems that can navigate the challenges of the evolving digital media landscape.

5. Conclusion

This thesis has made significant strides in bridging the gap between textual media sentiment and market volatility. By using advanced NLP techniques, we are able to extract refined sentiment from media content, revealing its dynamic influence on market behavior. Our research not only delves into how news narratives shape investor decisions, but also demonstrates the superiority of news sentiment analysis in forecasting compared to social media sentiment. Furthermore, we break new ground by differentiating the effects of real and fake news on markets, highlighting the importance of information credibility in the digital age. The findings of this thesis offer insights for financial practitioners, policymakers, and academics alike, paving the way for more informed decision-making and resilient financial ecosystems.

Reference

Ahern, K.R. and Sosyura, D., 2015. Rumor has it: Sensationalism in financial media. *The Review of Financial Studies*, *28*(7), pp.2050-2093.

Arcuri, M.C., Gandolfi, G. and Russo, I., 2023. Does fake news impact stock returns? Evidence from US and EU stock markets. *Journal of Economics and Business*, *125*, p.106130.

Audrino, F., Sigrist, F. and Ballinari, D., 2020. The impact of sentiment and attention measures on stock market volatility. *International Journal of Forecasting*, *36*(2), pp.334-357.

Bae, K.H., Karolyi, G.A. and Stulz, R.M., 2003. A new approach to measuring financial contagion. *The Review of Financial Studies*, *16*(3), pp.717-763.

Behera, C. and Rath, B.N., 2022. The connectedness between Twitter uncertainty index and stock return volatility in the G7 countries. *Applied Economics Letters*, *29*(20), pp.1876-1879.

Bikhchandani, S., Hirshleifer, D. and Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, *100*(5), pp.992-1026.

Bikhchandani, S., Hirshleifer, D. and Welch, I., 1998. Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of economic perspectives*, *12*(3), pp.151-170.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), pp.307-327.

Boudoukh, J., Feldman, R., Kogan, S. and Richardson, M., 2013. *Which news moves stock prices? A textual analysis* (No. w18725). National Bureau of Economic Research.

Brenner, M., Pasquariello, P. and Subrahmanyam, M., 2009. On the volatility and comovement of US financial markets around macroeconomic news announcements. *Journal of Financial and Quantitative Analysis*, *44*(6), pp.1265-1289.

Clarke, J., Chen, H., Du, D. and Hu, Y.J., 2021. Fake News, Investor Attention, and Market Reaction. *Information Systems Research*.

Damstra, A., Boomgaarden, H.G., Broda, E., Lindgren, E., Strömbäck, J., Tsfati, Y. and Vliegenthart, R., 2021. What does fake look like? A review of the literature on intentional deception in the news and on social media. *Journalism Studies*, *22*(14), pp.1947-1963.

Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Diebold, F.X. and Yilmaz, K., 2012. Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of Forecasting*, *28*(1), pp.57-66.

Dougal, C., Engelberg, J., Garcia, D. and Parsons, C.A., 2012. Journalists and the stock market. *The Review of Financial Studies*, *25*(3), pp.639-679.

Engelberg, J.E. and Parsons, C.A., 2011. The causal impact of media in financial markets. *The Journal of Finance*, *66*(1), pp.67-97.

Fiske, S.T. and Taylor, S.E., 1991. *Social cognition*. McGraw-Hill Book Company.

Fraiberger, S.P., Lee, D., Puy, D. and Ranciere, R., 2021. Media sentiment and international asset prices. *Journal of International Economics*, *133*, p.103526.

Goldman, E., Martel, J. and Schneemeier, J., 2022. A theory of financial media. *Journal of Financial Economics*, *145*(1), pp.239-258.

Grossman, S.J. and Stiglitz, J.E., 1980. On the impossibility of informationally efficient markets. *The American economic review*, *70*(3), pp.393-408.

Hirshleifer, D. and Teoh, S.H., 2003. Limited attention, information disclosure, and financial reporting. *Journal of accounting and economics*, *36*(1-3), pp.337-386.

Hirshleifer, D., 2015. Behavioral finance. *Annual Review of Financial Economics*, *7*, pp.133-159.

Hochreiter, S. and Schmidhuber, J., 1997. Long short-term memory. *Neural computation*, *9*(8), pp.1735-1780.

Horne, B.D. and Adali, S., 2017, May. This just in: Fake news packs a lot in title, uses simpler, repetitive content in text body, more similar to satire than real news. In *Eleventh international AAAI conference on web and social media*.

Howard, J. and Ruder, S., 2018, July. Universal Language Model Fine-tuning for Text Classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 328-339).

Hsu, Y.J., Lu, Y.C. and Yang, J.J., 2021. News sentiment and stock market volatility. *Review of Quantitative Finance and Accounting*, *57*(3), pp.1093-1122.

Kahneman, D., 1973. *Attention and effort* (Vol. 1063, pp. 218-226). Englewood Cliffs, NJ: Prentice-Hall.

Kogan, S., Moskowitz, T.J. and Niessner, M., 2019. Fake news: Evidence from financial markets. *Available at SSRN*, *3237763*.

Kogan, S., Moskowitz, T.J. and Niessner, M., 2023. Social media and financial news manipulation. *Review of Finance*, *27*(4), pp.1229-1268.

Malo, P., Sinha, A., Korhonen, P., Wallenius, J. and Takala, P., 2014. Good debt or bad debt: Detecting semantic orientations in economic texts. *Journal of the Association for Information Science and Technology*, *65*(4), pp.782-796.

Malo, P., Sinha, A., Takala, P., Ahlgren, O. and Lappalainen, I., 2013, December. Learning the roles of directional expressions and domain concepts in financial news analysis. In *2013 IEEE 13th international conference on data mining workshops* (pp. 945-954). IEEE.

Niederhoffer, V., 1971. The analysis of world events and stock prices. *The Journal of Business*, *44*(2), pp.193-219.

Parkinson, M., 1980. The extreme value method for estimating the variance of the rate of return. *Journal of Business*, pp.61-65.

Rogers, E.M., 1962. Diffusion of innovations. *Diffusion of innovations.*

Rogers, E.M., Singhal, A. and Quinlan, M.M., 2014. Diffusion of innovations. In *An integrated approach to communication theory and research* (pp. 432-448). Routledge.

SEC., 2023. *Establishment of an Ethical Artificial Intelligence Framework for Investment Advisors*. U.S. Securities and Exchange Commission.

Shiller, R.J., 1989. *Market volatility*. MIT press.

Shiller, R.J., 2020. *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

Tetlock, P.C., 2007. Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, *62*(3), pp.1139-1168.

Timmermann, A., 2006. Forecast combinations. *Handbook of economic forecasting*, *1*, pp.135-196.

Tuchman, G., 1978. *Making news: A study in the construction of reality*. Free Press.

Tversky, A. and Kahneman, D., 1973. Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, *5*(2), pp.207-232.

Valente, T.W., 1996. Network models of the diffusion of innovations. *Computational & Mathematical Organization Theory*, *2*(2), pp.163-164.

Yu, Z., Wang, M.D., Wei, X. and Lou, J., 2023. News credibility and influence within the financial markets. *Journal of Behavioral Finance*, *24*(2), pp.238-257.

Paper I

Measuring Market Volatility Connectedness to Media Sentiment

North American Journal of Economics and Finance 71 (2024) 102091

Contents lists available at ScienceDirect

North American Journal of Economics and Finance

journal homepage: www.elsevier.com/locate/najef

Measuring market volatility connectedness to media sentiment

Hooman Abdollahi[a,*], Sturla L. Fjesme[b,c], Espen Sirnes[a]

[a] School of Business and Economics, UiT-The Arctic University of Norway, Tromsø, Norway
[b] Oslo Business School, Oslo Metropolitan University, Norway
[c] School of Economics, Innovation and Technology, Kristiania University College, Oslo, Norway

A R T I C L E I N F O

JEL Classification:
G41
G14
E32

Keywords:
Financial market volatility
Time connectedness
Transmission mechanism
News sentiment
Twitter sentiment
Natural language processing

A B S T R A C T

We examine directional connectedness patterns from news and social media to financial market volatility using textual analysis and high-frequency data. We find that media sentiment induces market volatility, but the magnitude of that connectedness is time-varying. In addition, news and social media sentiment pertinent to one market transmits volatility to other markets. Finally, we find that sentiment transmits sharp shocks to markets during major events. At other times, there are smaller spillover effects, indicating that the directional connectedness from sentiment to markets follows a spiky pattern over time. We conclude that news and social media play an important (but not constant) role in transmitting volatility across financial markets. This insight explains earlier divergent findings in the literature.

1. Introduction

A long line of research shows that, in addition to fundamentals and the interconnection of markets, news stories on topics such as political conflicts, the state of public opinion, economic events, or the general business climate can be sources of market volatility (e.g., Brenner et al. (2009), Lucca and Moench (2015), and de Oliveira et al. (2018)). Media can then spread investor sentiment and influence expectations about future market behavior (Brown and Cliff, 2004). In some instances, these mechanisms appear to have significant effects. For instance, there is a prevailing view that a surge in sentiment might have been a contributing factor to the challenges faced by the Silicon Valley Bank in 2023.

This evidence suggests that part of financial market volatility is connected to media sentiment. However, to date, limited attention has been devoted to examining how sentiment behaves across financial markets because quantifying sentiment from news stories and social media posts has been a challenge. In this paper, we use recent advances in natural language processing to extract and quantify market sentiment in higher volumes and with greater precision than has been possible previously. We then develop a new measure of news and social media sentiment, which we use to answer the following questions: (i) How much of the volatility seen in the market can be attributed to the sentiment embedded in news stories and social media? (ii) To what degree does this relation explain volatility transmitted to other markets? (iii) How do sentiment shocks behave across international markets?

To investigate the connectedness between sentiment and market volatility, we obtain daily values for six financial markets, including Brent crude oil and markets in five industrialized countries: the US (the Dow Jones Industrial Average (DJIA)), the UK (the Financial Times Stock Exchange (FTSE)), France (the CAC40), Germany (the DAX30), and Japan (the Nikkei). We gather news

* Corresponding author at: UiT, Hansine Hansens veg 18, Tromsø N-9019, Norway.
E-mail addresses: hooman.abdollahi@uit.no (H. Abdollahi), sturla.fjesme@oslomet.no (S.L. Fjesme), espen.sirnes@uit.no (E. Sirnes).

https://doi.org/10.1016/j.najef.2024.102091
Received 8 August 2023; Received in revised form 9 November 2023; Accepted 21 January 2024
Available online 29 January 2024

headlines and tweets regarding each market from 4 August 2014, to 22 December 2020.

We extract sentiment from news headlines and Twitter tweets using advanced natural language processing, which enables us to convert textual data into numerical values. We employ the Bidirectional Encoder Representations from Transformers (BERT) model proposed by Devlin et al. (2018), and train it on financial language. We produce a total of 12 sentiment series for the news and tweets of each market with daily values between $+1$ (the most positive) and -1 (the most negative). Using the computed sentiment series, we measure the magnitude of volatility spillovers through the framework proposed by Diebold and Yilmaz (2012) (DY henceforth).[1] This framework is built upon predictive modeling under misspecification and can be used to measure causal connections. This approach enables us to uncover the mechanism behind volatility transmission.

Our main empirical finding is evidence of a degree of interdependence among the selected markets and media sentiment. How information affects markets on a general level is reasonably well understood in the classical finance literature. This study contributes by analyzing in more detail the information market participants use. We find that the effect of sentiment on markets varies over time and is more intense during extreme events.

Although the static results over our sample period do not capture the evolution of time-varying dependencies, they still document that sentiment has a long-lasting but relatively small effect on market volatility. Secondly, we document that the sentiment associated with a specific market induces volatility across other markets, indicating that volatility is transmitted between markets through news and social media. By including social media, a relatively new information medium, in our sample, we incorporate public perceptions into our analyses. We find that social media sentiment has a stronger long-lasting effect than news sentiment for some markets. Finally, using rolling-window analyses to examine the evolution of connectedness, we find that the connection between sentiment and volatility varies significantly over time. We further document how sentiment behaves within markets, showing that sentiment follows a spiky pattern, its effect appearing and disappearing over time.

From these findings, we conclude that news and the social media play a crucial role in transmitting volatility across financial markets. News stories—either fact-based, such as reporting posted on official news websites, or baseless like viral rumors on Twitter—contribute to the transmission of volatility across markets, though the salience of this transmission channel varies over time. Given the durability of massive contagions, we find that markets are highly capable of absorbing new information. In addition, the long-lasting effect we document implies that sentiment provides distinct information valuable for predicting volatility in the selected markets.

Our main contribution to the literature is showing that sentiment contributes to country-level volatility transmission in a time-varying and spiky pattern. Our study advances existing research on the relation between sentiment and financial markets (Behera and Rath, 2022; Costola et al., 2023; Mensi et al., 2023) by measuring dynamic directional connectedness from media sentiment to market volatility. This approach allows us to provide a granular analysis of the extent to which daily market volatility arises from media sentiment. We build exclusive sentiment indices for each market using both news headlines and Twitter feeds. This practice facilitates a broader analysis on both national and international scales, distinguishing our research from prior studies (Behera and Rath, 2022; Feng et al., 2022; Qiao et al., 2022; Basak et al., 2023; Zeitun et al., 2023). We document a tenuous connection between sentiment corresponding to one market and volatility across markets in other countries. Furthermore, our research distinguishes from prior studies by using exclusive sentiment indices for each market, which allows us to be the first to provide a comparative analysis of media sentiment related to different markets and their corresponding market volatility. We document that the long-lasting effect of news sentiment on market volatility is more pronounced than that of Twitter sentiment in all markets except for the UK and German markets. This discovery contributes new insights into the varying influences of different types of sentiment on financial market dynamics. Finally, we extend the literature on media sentiment and market dynamics as our sentiment measure is more sophisticated than those used in previous studies in two ways. Unlike prior studies that focus only on one aspect of news (see Rangel, 2011; Brandt and Gao, 2019; Dong et al., 2022; Feng et al., 2022; Gao et al., 2022; Koch et al., 2022; Basak et al., 2023), we cover the whole spectrum of news, ranging from business and economic to political and terrorism headlines. Furthermore, unlike previous studies that employ Google Trends, word-counting, or lexicon-based approaches to sentiment extraction, we capture semantic sentiment through a BERT model fine-tuned with a financial PhraseBank, which generates sentiment based on the meaning of the text. Our innovative use of BERT contributes to the literature by measuring the semantic sentiment of material posted on Twitter by the public in addition to news published by traditional media outlets, reducing the potential bias driven by news media. These contributions can help policymakers and investors who wish to hedge their risks across international markets understand behavioral drivers of volatility.

The paper proceeds as follows. Section 2 presents an overview of the literature together with hypothesis development. Section 3 describes the methodologies for sentiment extraction and econometric analysis. The data are described in Section 4. Section 5 reports the results, and Section 6 concludes the paper.

2. Related literature and hypothesis development

The literature has theorized the role of sentiment in the market dynamics. Grossman and Stiglitz (1980) suggest that market fluctuations can, in part, result from unexplained variations in public opinion, or *noise trading*. De Long et al. (1990) note that market dynamics can be influenced by nonfundamental factors, such as traders' erroneous stochastic beliefs, which deflect prices from

[1] The DY framework decomposes the total spillover into the spillover from one variable to another variable and overcomes the impact of ordering in the orthogonal decomposition results. Therefore, it fits to our study, whose aim is to capture the directional connectedness in a multivariate network.

fundamental values even when fundamental risk is absent. Klibanoff et al. (1998) propose that news is a source of financial information that affects investors' opinions, triggering some to react more quickly when news is particularly salient.

Building on this foundation, scholars have posited an analytical framework to quantify the sentiment from textual sources and assess its impact on financial data. Using word counts and a pre-existing sentiment dictionary, Tetlock (2007) finds that news pessimism puts downward pressure on prices. Subsequently, finance scholars began to pay more attention to textual analysis, leading Loughran and McDonald (2011) to build an exclusive sentiment dictionary for financial filings. These textual analytic techniques have advanced this line of research by providing a basic understanding of specific parts of textual data. More recently, Shiller (2020) suggests that more finely tuned quantitative methods should be developed to explore how various types of media influence financial markets. More specifically, Shiller points out the necessity of constructing a semantic sentiment proxy.[2]

Shifting to empirical studies, the relation between media and financial markets was first investigated by Niederhoffer (1971), who shows that investment-related news headlines impact price movement. Later, Ederington and Lee (1993) document that news announcements have a strong impact on volatility patterns across financial markets. Griffin et al. (2011) study the relation between volatility and public news arrival and find that in most developed markets the volatility is more on the news announcement day. Da et al. (2015) employ Google Trends metrics for specific economic keywords and document that their metrics predict temporary increases in market volatility. Manela and Moreira (2017) measure uncertainty from front-page articles in *The Wall Street Journal*, which they call news-based implied volatility (NVIX), and find that volatility peaks during extreme events like financial crises and world wars. Su et al. (2019) examine the spillover of three sentiment indices across nine markets and find that NVIX is the most powerful predictor of market volatility. Liu et al. (2019) use a geopolitical risk index to examine oil price volatility and conclude that serious geopolitical risk is important in determining oil price volatility. Also, a new line of research investigates the impact of sentiment embedded in social media on market volatility. Nishimura and Sun (2021) capture the impact of the US president tweets concerning the US-China economic conflict on market volatility and show the expanding impact of tweets on market volatility. Bouri et al. (2022) build a proxy for investor happiness through Twitter feeds and document that their proxy affects volatility more than returns. Aharon et al. (2022) employ a media coverage index tailored to COVID-19 and indicate the significant role of such a sentiment in volatility transmission across the G7 countries. Shen et al. (2022) use a dictionary-based approach to measure news tone and discover that sentiment has an asymmetric impact on future market volatility. Mensi et al. (2023) utilize established uncertainty indices, such as economic policy uncertainty, and find that these indices predominantly transmit volatility to the market during both bearish and tranquil market periods. Katsafados et al. (2023) use a lexicon-based approach to analyze tweets for positive and negative sentiments during a specific phase of the COVID-19 pandemic. Their findings indicate that positive sentiment is associated with lower volatility. Apergis et al. (2023) utilize Google Trends news to examine the influence of COVID-19-related sentiment on market volatility and find significant impacts from the pandemic.

Joseph et al. (2011) and Da et al. (2011) show that increased positive sentiment is related to increased prices and subsequently volatility. Fedyk (2018) finds that front-page financial news is incorporated in financial markets more rapidly than non-front-page news. Based on this literature, we expect that news sentiment can have a consistent and long-lasting effect on market dynamics, leading us to our first hypothesis:

H1. There is a long-lasting connectedness between sentiment expressed through news and social media and market volatility.

Ross (1989) finds that the volatility across a market directly corresponds to the intensity of information flow coming into that market. Williams (1999) links information efficiency to the speed of incorporating new information into prices. Nishimura and Sun (2018) propose that if a market is informationally efficient, information shocks in that market can cause volatility in other markets. Based on these findings, we expect that volatility will spread between different markets through news and social media. We formalize this reasoning in our second hypothesis:

H2. Volatility is transmitted between markets by news and social media.

Birz (2017) examines the effect of macroeconomic news on stock prices and documents that economic news stories influence prices during the announcement week, before reversing over the following week. Shiller (2020) argues that the contagious impact of news stories increases as the public talks about the stories, but the effects slow down when people eventually lose interest. Shiller proposes that the effect of news stories on the market can be analogous to the spread of an epidemic disease, where the rate of contagion is higher than the recovery in the early stages before it slows down. Hence, we anticipate that the behavior follows a spiky pattern. Based on this research, we expect that volatility spreads across markets in a spiky pattern, as stated in our third hypothesis:

H3. The transmitted volatility from media sentiment to the markets follows a spiky pattern.

3. Methodology

3.1. Textual analytics

Bidirectional Encoder Representations from Transformers (BERT) is a pre-trained machine learning technique for textual analysis developed by Google. The model is trained on English Wikipedia and BookCorpus. BERT is a multilayer deep-learning model composed of an encoder in the first layer that takes the input text, such as news headlines or tweets, and a final layer that predicts the probabilities of the given text being positive, neutral, or negative. The final layer (output) of the neural network, called logits, returns the prediction probabilities, which here means how a given news headline/tweet is to be perceived. To calculate sentiment scores, a common practice is to

[2] "There should be more serious efforts at collecting further time series data on narratives, going beyond the passive collection of others' words, towards experiments that reveal meaning and psychological significance" (Shiller 2017, p. 48).

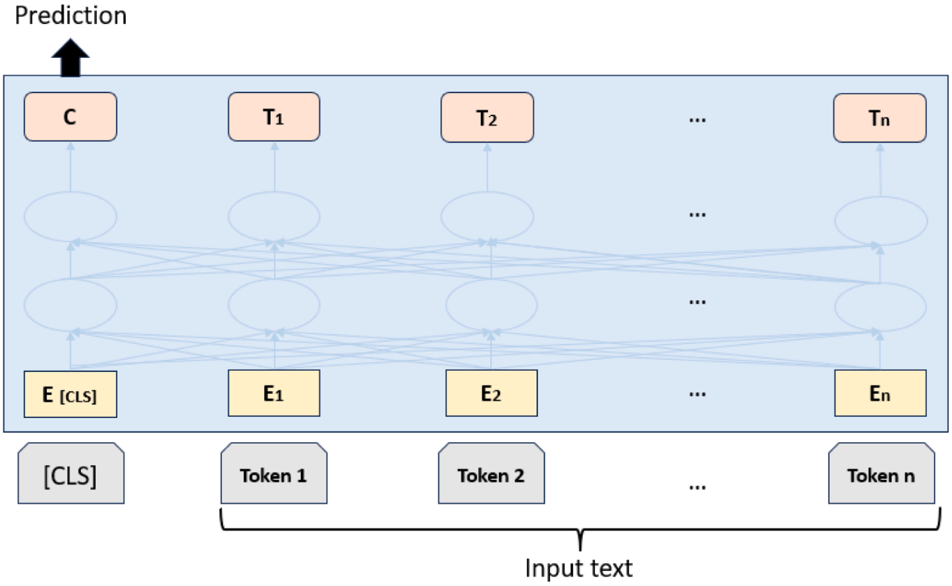


Fig. 1. BERT Simplified Structure (Adapted from Devlin et al. (2018)). This figure shows the structure of BERT for sentiment prediction. Starting with the input text which is divided into tokens of word sequentially. *[CLS]* is a special token added to the beginning of every input sequence, serving as an aggregation representation for classification tasks. Each token is embedded into vectors ($E_{[CLS]}$, E_1, E_2,…,E_n). These embeddings are then processed through several encoder layers (represented by ellipses), resulting in contextualized token representations (T_1, T_2,…,T_n). The prediction, in the context of sentiment analysis, is derived from the final encoder layer's output from the *[CLS]* token, denoted as *C*.

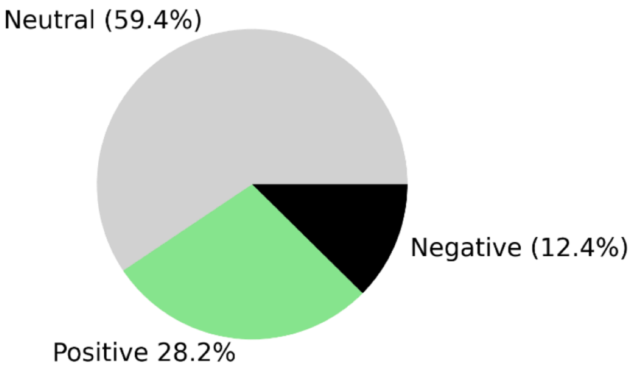


Fig. 2. Classification of Sentiment Labels in Financial PhraseBank. This figure shows the portion of sentiment labels in the Financial PhraseBank, a dataset of 4,845 news sentences annotated by 16 finance domain professionals.

simply calculate the difference between positive and negative logits (Lin and Luo, 2020), as shown in Equation (1). Therefore, the sentiment score always falls between −1 (the most negative) and 1 (the most positive). Fig. 1 depicts the simplified structure of BERT for sentiment analysis task.

$$SentimentScore = logit_{positive} - logit_{negative} \tag{1}$$

We use the BERT base version, which has 12 encoder layers, 12 multihead attention heads, a hidden size of 768, and 110 million parameters. Because our aim is to predict the sentiment label and score of news headlines and tweets in the context of finance, it is essential to fine-tune the model across this domain. Thus, we use the Financial PhraseBank (Malo et al., 2014), which comprises 4,845 selected sentences from financial news labeled as positive, negative, or neutral by 16 professionals within the finance domain. Labels indicate how the annotators perceive that the information embedded in a sentence might influence prices. Fig. 2 shows the proportion of positive, negative, and neutral sentences in the PhraseBank. Training the BERT with the Financial Phrasebank enhances the model's performance in processing text within the finance context as the model learns specialized terms used in the finance literature. We train the model using different split ratios and evaluate its performance using two standard metrics of accuracy and macro F1 average as shown in Equations (2) and (3) (Baeza-Yates and Ribeiro-Neto, 1999).

$$Accuracy = \frac{Numberofcorrectpredictions}{Totalnumberofsample} \tag{2}$$

$$F1scoreineachlable = 2 \times \frac{1}{\frac{1}{percision} + \frac{1}{recall}}$$

$$MacroF1average = \frac{\sum F1scoresineachlable}{Totalnumberoflables} \tag{3}$$

Finally, to obtain the sentiment score for each day, we compute the mean of all sentiment scores for that day. This produces daily proxies for both news and Twitter sentiment and leads to a set of novel and unique data that we incorporate into the econometric analysis.

3.2. The Diebold and Yilmaz (DY) framework

To investigate the connectedness between news headlines, tweets, and market volatility, we follow the framework proposed by Diebold and Yilmaz (2012, 2014).[3] We begin by modeling market volatility and our sentiment series as an n-variable vector autoregression (VAR) model in the first step. Then, we employ the generalized variance decomposition (GVD) introduced by Koop et al. (1996) and Pesaran and Shin (1998) to construct the H-step-ahead forecast and to decompose the forecast error variance for each variable corresponding to shocks coming from other variables at time t. Let $y_t = (y_{1t}, y_{2t}, \cdots, y_{nt})$ be a vector of market volatility and sentiment series:

$$y_t = \sum_{i=1}^{p} \theta_i y_{t-i} + \epsilon_t, \tag{4}$$

where θ_i (for $i = 1, 2, \ldots, p$) are $n \times n$ matrices, and ϵ_t, as the error vector, has zero means and a variance–covariance matrix Σ. The moving-average representation of this $VAR_{(p)}$ process is:

$$y_t = \sum_{i=0}^{\infty} \varphi_i \epsilon_{t-1} t = 1, \cdots, T$$

$$\varphi_i = \theta_1 \varphi_{i-1} + \theta_2 \varphi_{i-2} + \cdots + \theta_m \varphi_{i-m}, \tag{5}$$

where y_t denotes a $(K \times 1)$ vector for the involved series; φ_0 is $n \times n$ identity matrices and equals zero for $i < 0$; θ_i is $n \times n$ autoregression coefficient matrices; and ϵ_t denotes the vector of error terms ($i.i.d.$). Based on the generalized VAR, the H-step-ahead forecast error variance decomposition of the i^{th} variable coming from the j^{th} variable is:

$$\varphi_{ij}^{H} = \frac{\sigma_{ij}^{-1} \sum_{h=0}^{H-1} (e\hat{A}_i \varphi_h \Sigma e_j)^2}{\sum_{h=0}^{H-1} (e\hat{A}_i \varphi_h \Sigma \varphi \hat{A}_h e_j)}, \tag{6}$$

where σ_{ij} is the standard deviation for the error term in the j^{th} equation, and e_i denotes a selection vector. Finally, the directional connectedness, which is the shock transmitted from sentiment j to market i, can be calculated as follows:

$$C_{i \leftarrow j} = \frac{\varphi_{ij}^{H}}{\sum_{ij=1}^{N} \varphi_{ij}^{H}}. \tag{7}$$

The DY framework allows us to estimate both static and dynamic connectedness. Static connectedness provides evidence of spillovers over the entire period, whereas dynamic connectedness provides evidence of the evolution of connectedness across time.

4. Data

We investigate market connectedness based on sentiment and historical data from the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil. Our selection of these markets is derived from multiple considerations. Firstly, these markets benefit from comprehensive coverage across both conventional news sources and social media platforms. Moreover, they stand as some of the most globally pivotal and interconnected financial centers. The dynamics of these markets have profound implications for worldwide financial stability, underscoring their significance in our analysis. Additionally, by opting for markets spread across diverse geographical regions, we try to encompass a wide spectrum of economic climates, policy environments,

[3] See Diebold and Yilmaz (2012, 2014) for further details about the applied framework. Note that we employ the DY framework only to measure the directional impact from sentiment to the market.

Table 1
The Number of News and Tweets. This table shows the number of relevant news headlines (panel A) and tweets (panel B) over the period from 04/08/2014 to 22/12/2020 for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets.

Market	2014	2015	2016	2017	2018	2019	2020	Total
Panel A- News								
DJIA	348	1,047	1,443	1,795	1,550	1,286	1,898	9,367
FTSE	381	716	662	907	715	556	694	4,631
CAC40	369	662	554	646	552	249	531	3,563
DAX	385	675	764	868	727	556	656	4,631
Nikkei	145	437	816	784	923	736	826	4,667
Brent	510	2,685	2,392	2,252	2,043	1,763	3,188	14,833
Total news 41,692								
Panel B- Tweets								
DJIA	4,878	12,327	14,034	24,311	20,595	36,614	53,906	166,665
FTSE	15,729	42,276	43,069	22,462	21,841	23,510	24,069	192,956
CAC40	1,483	4,052	3,177	2,208	2,719	3,126	5,275	22,040
DAX	7,173	15,506	16,172	10,425	7,017	10,837	15,707	82,837
Nikkei	711	7,897	6,762	4,274	3,779	2,597	2,759	28,779
Brent	2,329	14,127	20,376	17,778	20,257	13,702	27,894	116,463
Total tweets 609,740								

and prevailing sentiments. Lastly, recognizing the integral role oil plays in the global economy, the inclusion of Brent –a preeminent benchmark– offers a deeper understanding of the wider commodities market and its intricate relationship with media sentiment.

For the selected markets, we scrape Twitter tweets and news headlines from *investing.com*.[4] The news section of *investing.com* contains wide-ranging news from other sources like Reuters, Bloomberg, Business Insider, and Seeking Alpha, among others, as well as exclusive news stories written for the website. Drawing from esteemed outlets like Reuters and Bloomberg ensures the credibility and reliability of our news sample. Also, the market-centric approach of *investing.com* provides comprehensive coverage of relevant news specific to each market, offering real-time information. Moreover, with its global presence and publicly available news section, this website caters to a wide audience. These features justify our use of this website as our news source. Table 1 shows that our textual data between 4 August 2014, and 22 December 2020, contains a total of 41,692 news headlines and 609,740 tweets on specific hashtags germane to each market (see the list of hashtags in Appendix A1).[5]

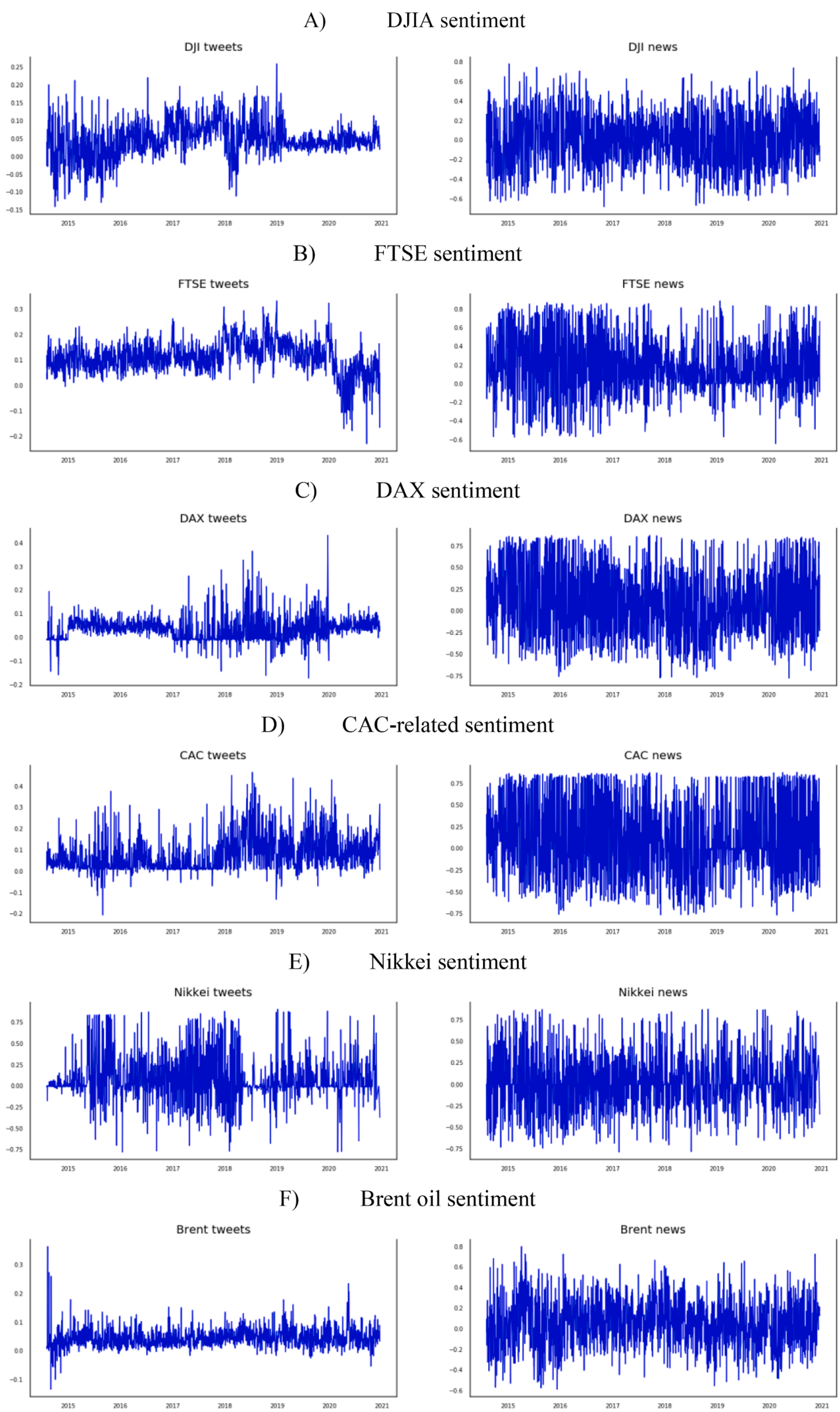
We select this timeframe based on several key considerations. Primarily, this interval coincides with when our data sources become robust and reliable, granting us access to a comprehensive compilation of news and tweets. Moreover, there are pivotal subperiods such as the Brexit implications for the UK, pivotal elections in the US, the Eurozone challenges faced by France and Germany, and Japan's evolving economic and foreign policies within this span. Additionally, this period bears witness to fluctuations in the Brent commodity market due to geopolitical tensions and global supply–demand dynamics. There are also instances of stability and growth, like the landmark Paris Climate Agreement or economic and political deals. Examining these nuanced phases facilitates a holistic understanding of how various news stories influence market volatilities. Furthermore, our dataset represents a continuous stream of public news along with tweets, ensuring a wide spectrum of information for an expansive investor demographic.

We train BERT using the Financial PhraseBank. For this purpose, we divide the PhraseBank into three subsets of training, validation, and test. The aim is to find the best generalization of the model. The training and validation data are used to fit the parameters to the model and tuning hyperparameters. The test data is then used to acquire an evaluation of the final model. In doing so, we apply different relative ratios of 70:30, 80:20, and 90:10 to find the best generalization of the model. Using Equation (2), the obtained accuracy metrics are 0.77, 0.83, and 0.81 for the 70:30, 80:20, and 90:10 split ratios, respectively. The macro F1 averages are obtained using Equation (3), which are 0.77, 0.82, 0.81 for the 70:30, 80:20, and 90:10 split ratios, respectively. These metrics are on a scale of 1 to 0, with 1 being the best value. We obtain the most quality outcomes using the relative ratio of 80:20. This means we use 80 % of the sentences for training and validation and 20 % for the test set. Therefore, we proceed with the BERT model trained on this ratio because it produces more accurate and quality data compared to the model trained on the other ratios. Fig. 3 shows the sentiment series obtained for each index for our sample period. Appendix A2 presents an example of the sentiment calculation process.

To explore the reliability of our sentiment series, we match them to similar established indicators. While we are unable to identify a widely recognized index directly analogous to our sentiment series, we utilize the most comparable benchmarks. For our news sentiment series, we reference the Geopolitical Risk Index introduced by Caldara and Iacoviello (2022), primarily focused on geopolitical, war, and terrorist events. For our Twitter sentiment series, we draw upon the Twitter Economic Uncertainty (Baker et al.,

[4] Investing.com is an online financial markets platform that provides real-time news updates on more than 250 international exchanges; see https://www.investing.com/about-us/.

[5] We use news headlines because those can be easily retrieved, and they convey the meaning of the full article (Li et al., 2019). Furthermore, using news headlines boosts the accuracy of the textual analysis as headlines generally include fewer repetitive or irrelevant words.



(caption on next page)

Fig. 3. Sentiment Series Extracted from News and Tweets for Each Market. News and Twitter sentiment series over the period from 04/08/2014 to 22/12/2020 for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil. Panel A shows the DJIA sentiment series extracted from 166,665 tweets and 9,367 news headlines. Panel B illustrates the calculated FTSE sentiment series using 192,956 tweets and 4,631 news headlines. Panel C portrays the calculated sentiment series for the DAX using 82,837 tweets and 4631 news headlines. Panel D shows the CAC sentiment series extracted from 22,040 tweets and 3,553 news headlines. Panel E shows the calculated sentiment series for the Nikkei using 28,779 tweets and 4,667 news headlines. Panel F shows the calculated sentiment series for Brent oil using 116,463 tweets and 14,833 news headlines.

2021), which quantifies the usage of economic uncertainty-related words in tweets. It is important to note that our sentiment series encompass a comprehensive range of topics. Therefore, they should partly predict other sentiment indices, given the overlaps in content. Panel OLS regressions show that there is a negative and significant relation between our indices and the benchmarks.[6] These inverse relations stem from our sentiment series capturing both positive and negative aspects of news and tweets, while benchmark indices solely capture negative aspect. This means when our sentiment series shift towards a positive spectrum, the severity of the benchmark indices subsides. These results justify our use of the sentiment measures.

We obtain daily closing, minimum, and maximum prices for the selected markets from Refinitiv Eikon for 4 August 2014, to 22 December 2020. Because volatility is unobserved, we compute it using the following daily range-based measure proposed by Parkinson (1980) in Equation (8):

$$\sigma_t^2 = 0.361\left[ln(H_t) - ln(L_t)\right]^2 \tag{8}$$

where H_t and L_t are the maximum and minimum price, respectively, on day t. This measure uses a factor of 0.361 to make the estimation comparable to the more common square measures. Brandt and Diebold (2006) find that the efficiency of the high-low range-based measures is between realized measures using three-hour and six-hour estimations. Fig. 4 shows the time-series plot and volatility for all indices over the sample period. Table 2 presents descriptive statistics for the series.

For the FTSE– and Brent–Twitter sentiment, the standard deviation values are smaller than the average, while the opposite is true for all other series. Thus, Twitter sentiment regarding the FTSE and Brent is less spread out. Standard deviations confirm that Brent (0.002397) and the FTSE (0.000238) have the highest volatility level of the six markets studied, while the CAC–news sentiment (0.435580) and the Nikkei–Twitter sentiment (0.270947) fluctuate the most over time amongst the sentiment series of news and Twitter, respectively. All news series except for the Nikkei are fairly symmetrical, as the skewness is between −0.5 and 0.5. The sentiment series for the Nikkei–Twitter, Nikkei–news, and FTSE–Twitter are moderately skewed, with values between 0.5 and 1 (or −0.5 and −1). The remaining series are highly skewed.

We check the stationarity of the times series using the augmented Dickey-Fuller (ADF) test. The null hypothesis suggests the existence of a unit root in the series. The last column of Table 2 presents the results of the ADF test, confirming the stationarity of all series at the 5 % significance level. In other words, some features of the series like mean and variance do not vary over time.

5. Empirical results

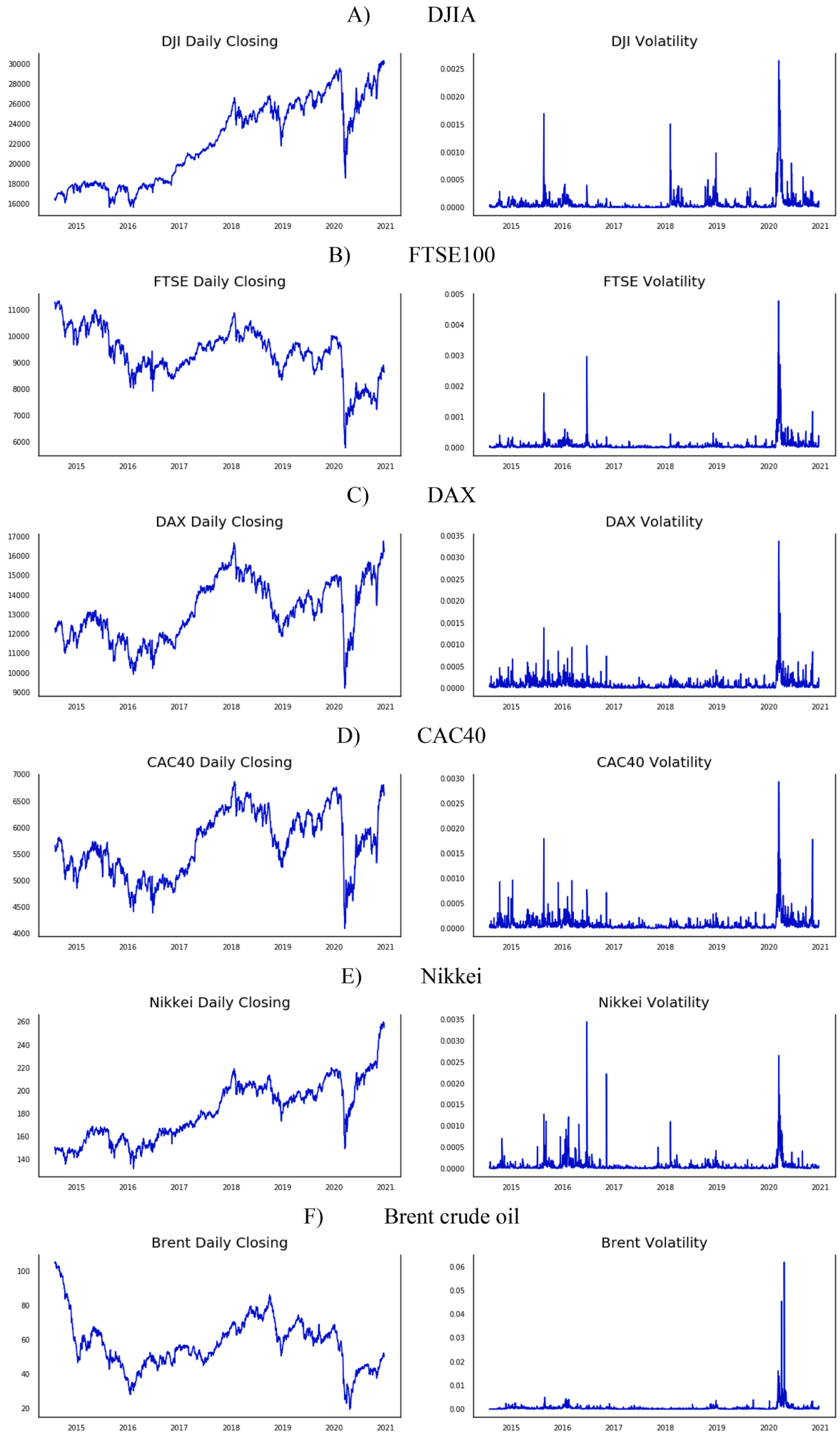
5.1. Market volatility and sentiment

Hypothesis 1 posits a long-lasting connectedness between market volatility and sentiment expressed through news headlines and Twitter tweets. We estimate the existing static connectedness between the sentiment proxies and each market from August 2014 to December 2020 using Equations (5) and (7) with a forecasting time horizon of 15 days. Table 3 shows the magnitude of connectedness between each market and its associated sentiment. The results in Column 1 indicate a statistically significant connectedness between each market and its corresponding news sentiment. Among the markets studied, news sentiment holds the strongest connectedness with DJIA volatility, accounting for 5.23 % of the DJIA's volatility over the entire sample period. We interpret this finding as evidence that the volatility in the US market is influenced by shocks in news sentiment more than is the case in other markets. The German market (the DAX) is the least sensitive to the news, with 0.67 % of its volatility arising from news sentiment. This result suggests that DAX volatility is the least affected by shocks to its news sentiment.

Column 2 of Table 3 shows statistically significant volatility spillovers coming from Twitter sentiment to the relevant markets. The FTSE, with a magnitude of 4.4 %, holds the largest connectedness, while the French market is the least impacted, with a negligible connectedness of 0.46 % over the entire sample period. The DJIA, the CAC, the Nikkei, and Brent receive more shocks from news headlines than from Twitter, while the opposite is true for the FTSE and the DAX.

To further examine the validity of the findings, we assess overall connectedness using Fisher's test. Table 3 shows that the Fisher's test p-value equals zero. This test confirms a statistically significant association between market volatility and news sentiment. Consistent with Hypothesis 1, the static results for the entire sample indicate that sentiment shocks have a long-lasting effect on the volatility of the markets in question, though the estimation accounts for only a minuscule share of the total fluctuations. We conclude that part of the volatility for each market arises from the shocks coming from that market's sentiment series, indicating a long-lasting

[6] A one-unit change in our news sentiment series corresponds to a −0.71-unit change in the Geopolitical Risk Index, a significant relationship at the 5% level. Similarly, a one-unit change in our Twitter sentiment series corresponds to a −4.31-unit change in the Twitter Uncertainty Index, significant at the 10% threshold.



(caption on next page)

Fig. 4. Historical Time Series and Volatility. Daily closing prices and range-based volatility over the period from 04/08/2014 to 22/12/2020 for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany the (DAX index), Japan (the Nikkei index), and Brent crude oil. Panel A-F shows the historical trends in closing price and volatility for the DJIA, FTSE100, DAX, CAC, Nikkei, and Brent crude oil, respectively.

Table 2

Descriptive Statistics. This table shows descriptive statistics for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets, together with their corresponding news and Twitter sentiment series over the period from 04/08/2014 to 22/12/2020 through panel A-C. Values in the first column presents the daily mean for each series followed by the minimum and maximum values in the next two columns. SD stands for standard deviations. Numbers in parenthesis in the last column show the p-value at the 5% significance level.

	Mean	Minimum	Maximum	SD	Skewness	ADF
Panel A- Volatility						
DJIA	0.000073	0.000001	0.002648	0.000190	7.721130	−6.6726 (0)
FTSE	0.000088	0.000002	0.004780	0.000238	10.946080	−5.583 (0)
CAC	0.0000905	0.0000007	0.002942	0.000187	8.7727	−7.1212 (0)
DAX	0.0000973	0.0000007	0.003377	0.000188	8.831979	−6.5157 (0)
Nikkei	0.000055	0	0.003441	0.000167	10.557571	−7.6249 (0)
Brent	0.000651	0.000019	0.061929	0.002397	17.873160	−4.8465 (0.001)
Panel B- News						
DJIA News	0.006113	−0.681537	0.780305	0.278095	0.043738	−37.242 (0)
FTSE News	0.188174	−0.643356	0.889871	0.301199	0.210155	−29.0329 (0)
CAC News	0.116632	−0.764512	0.875995	0.435580	0.1295	−40.2754 (0)
DAX News	0.094747	−0.775066	0.872322	0.392180	0.144225	−9.4413 (0)
Nikkei News	0.017079	−0.786002	0.873290	0.301482	0.586975	−31.5618 (0)
Brent News	0.060612	−0.588032	0.799959	0.229696	0.017622	−7.674 (0)
Panel C- Twitter tweets						
DJIA Twitter	0.046290	−0.140883	0.259133	0.049346	−0.057065	−4.3457 (0.0004)
FTSE Twitter	0.109994	−0.228921	0.333787	0.063758	−0.503394	−2.7089 (0)
CAC Twitter	0.078504	−0.206283	0.465760	0.080821	1.2148	−4.0091 (0.00014)
DAX Twitter	0.036704	−0.172351	0.432159	0.050949	1,209642	−5.0889 (0)
Nikkei Twitter	0.091734	−0.778571	0.904280	0.270947	0.586975	−25.2268 (0)
Brent Twitter	0.044512	−0.134497	0.364388	0.033485	1.349825	−17.9132 (0)

Table 3

The DY Static Estimations for Each Market and its Corresponding Sentiments. This table presents the estimation of static connectedness for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets with their related news and Twitter sentiment series over the entire period from 04/08/2014 to 22/12/2020. The ij^{th} entry represents the directional connectedness, namely the percentage of the forecast error variance of market i due to shocks from sentiment j. Numbers in parentheses are standard errors based on bootstrapping. ** and * denote significance at 5% and 10% levels, respectively. The last row presents the p-value of the Fisher's test against the null hypothesis of no connectedness.

Index	News sentiment	Twitter sentiment
DJIA	5.23 (0.407)**	2.27 (0.05401)**
FTSE	1.06 (0.0363)*	4.4 (0.1014)**
CAC	2.93 (0.0389)**	0.46 (0.0396)**
DAX	0.67 (0.0283)*	0.92 (0.0243)**
Nikkei	2.02 (0.0239)**	1.04 (0.0224)*
Brent	2.16 (0.035)**	1.22 (0.0329)**
Fisher's p-value = 0		

connectedness between media sentiment and market volatility.

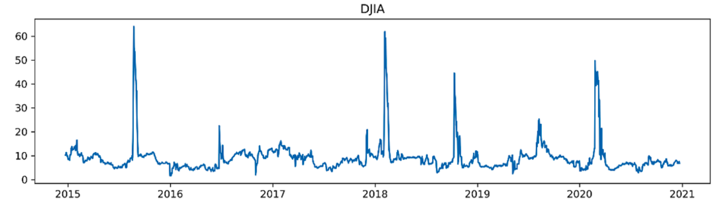
5.2. Market volatility and international sentiment

Hypothesis 2 predicts that sentiment associated with one market can transmit volatility to other markets. That is, volatility can be transmitted between markets by news and social media. We again use the DY static framework, this time including all sentiment series in the analysis, to investigate the extent of connectedness between volatility in each market and sentiment of other markets. We note that in

Table 4
The DY Static Estimation for all Markets and Sentiment series. This table presents the estimation of static connectedness for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets with own and other news and Twitter sentiment series over the entire period from 04/08/2014 to 22/12/2020. The ij^{th} entry represents the directional connectedness, namely the percentage of the forecast error variance of market i due to shocks from sentiment j. Numbers in parentheses are standard errors based on bootstrapping. ** and * denote significance at 5% and 10% levels, respectively. The last row presents the p-value of the Fisher's test against the null hypothesis of no connectedness.

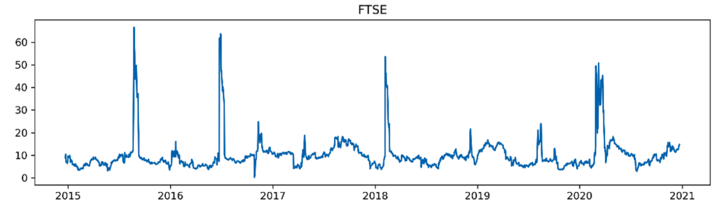
	DJIA News	DJIA Twitter	FTSE News	FTSE Twitter	CAC News	CAC Twitter	DAX News	DAX Twitter	Nikkei News	Nikkei Twitter	Brent News	Brent Twitter
DJIA	3.37** (0.011)	1.58** (0.007)	2.55** (0.008)	0.69** (0.013)	0.48** (0.001)	0.68** (0.006)	1.04** (0.007)	0.16**(0. 005)	1.51** (0.007)	1.33**(0.009)	2.37** (0.009)	0.91** (0.005)
FTSE	0.57** (0.006)	0.58** (0.007)	0.68** (0.007)	2.04** (0.019)	0.6**(0.097)	0.75** (0.006)	0.55** (0.006)	0.44*(0.005)	0.59** (0.006)	0.8**(0.008)	0.78** (0.007)	0.6**(0.006)
CAC	1.66** (0.198)	1.84** (0.417)	0.35** (0.008)	2**(0.017)	1.85** (0.008)	0.29** (0.007)	0.56** (0.006)	0.26** (0.006)	0.90** (0.007)	0.86** (0.008)	2.27** (0.07)	1.34** (0.006)
DAX	2.55** (0.005)	1.10** (0.015)	0.30** (0.008)	1.72** (0.019)	1.65** (0.006)	0.3*(0.007)	0.63** (0.006)	0.51**(0.005)	0.71** (0.006)	1.25**(0.009)	2.5**(0.007)	0.31** (0.005)
Nikkei	1**(0.005)	0.47** (0.005)	0.16*(0.006)	1**(0.05)	0.64** (0.006)	0.56** (0.004)	0.24*(0.005)	0.6*(0.009)	1.08** (0.006)	0.76**(0.006)	0.84**(0.05)	1.03** (0.009)
Brent	0.76** (0.007)	0.15** (0.008)	0.69**(0.01)	2.05** (0.034)	0.46** (0.007)	0.24*(0.008)	0.3**(0.007)	0.22*(0.009)	1.25** (0.009)	0.49**(0.008)	1.26** (0.008)	1.08** (0.008)
Fisher's p-value = 0												

A) Transmitted Volatility to DJIA



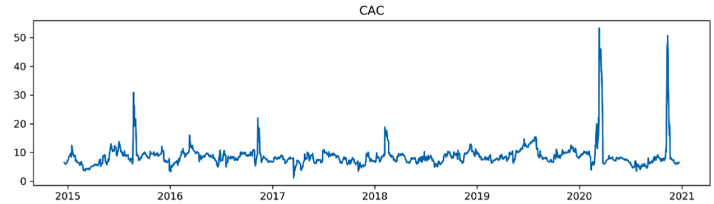
Maximum = 64.05 on24.08.2015, minimum = 1.63 on31.12.2015

B) Transmitted Volatility to FTSE



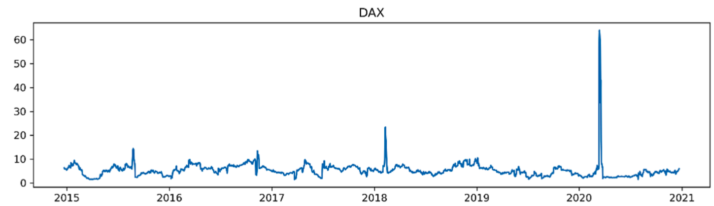
Maximum = 66.6 on 24.08.2016, minimum = 0.23 on 27.10.2016

C) Transmitted Volatility to CAC40



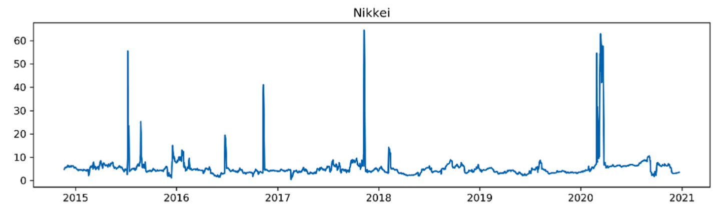
Maximum = 53.3 on 12-03.2020, minimum = 1.21 on 17.03.2017

D) Transmitted Volatility to DAX



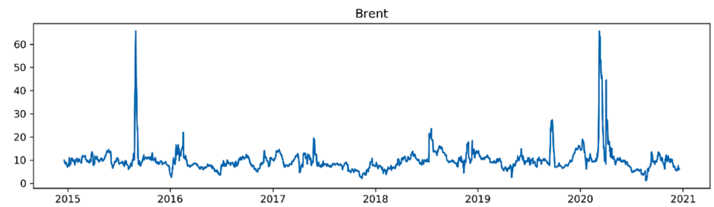
Maximum = 64 on 12.03.2020, minimum = 1.47 on 22.03.2017

E) Transmitted Volatility to Nikkei



Maximum = 64.45 on 09.11.2017, minimum = 0.49 on 18.02.2017

F) Transmitted Volatility to Brent



Maximum = 65.73 on 31.08.2015, minimum = 1.17 on 21.08.2020

(caption on next page)

Fig. 5. Spillovers Transmitted from Sentiment to the Markets over a 100-day Rolling Window. This figure presents total connectedness from both news and Twitter sentiments to the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets through panels A-F over the period from 04/08/2014 to 22/12/2020 through 15-day ahead forecast horizon (H = 15) and 100-day rolling window estimations.

Table 5
F-test for dynamic connectedness. This table shows the result of the F-test on the changes in dynamic volatility spillovers when the connectedness is above or below 20% at the significance level of 5%.

F-statistics	11.2082
P-value	0

recent decades news stories are not limited by national boundaries. Reports show that 57 % of people around the world follow international news, and 48 % follow U.S. news specifically.[7] Hence, global or international news is likely to strongly influence national news. Because our sample of news headlines and tweets contains both national and international news relevant to our six focal markets, a part of the connectedness we capture may stem from global events such as the COVID-19 pandemic. Although our sample of textual data incorporates global events in the context of a particular country, national news is likely still linked to international news, particularly when we source the news from globally prestigious news outlets. To negate such an effect on our estimation, we also add a global financial index to the VAR estimation we use to calculate connectedness.

The DY framework first generates forecast errors through a VAR process and then applies generalized variance decomposition for the final estimation of connectedness. Using a financial global index allows us to cover the forecast errors stemming from global factors such that the new estimation is orthogonal to the global factors. This process may decrease the magnitude of connectedness, but the final estimation is more robust, allowing us to more effectively test our second hypothesis. For this purpose, we use the MSCI world index, which is a market index of 1,546 companies all over the world.

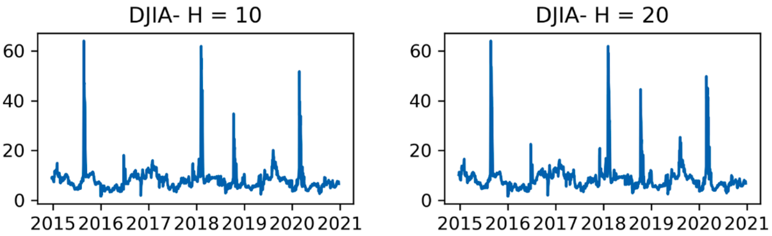
The results in Table 4 show a statistically significant connectedness between markets and various sentiment series. The connectedness between each market and its own sentiment series is a little smaller than the connectedness values presented in Table 3 due to the inclusion of the global factor and international sentiment series. Because our aim is to examine the connectedness across markets and sentiment, we only report the connection between each market and international sentiment series. The DJIA holds the largest connectedness with news about the FTSE and Brent, with magnitudes of 2.55 % and 2.37 %, respectively. We interpret this result as evidence that volatility in the US market is more connected with shocks in news sentiment associated with the FTSE and Brent than with other markets. The connection between CAC news sentiment and the DJIA is the weakest (0.48 %) of the relationships we test.

As for Twitter sentiment, the Nikkei and Brent send the biggest shocks to the DJIA; 1.33 % and 0.91 % of DJIA variability can be attributed to them, respectively. We also see that the US market is influenced by the Twitter sentiment of other markets. The second row of Table 4 shows a small connectedness between the FTSE and all sentiment series, with each international sentiment series accounting for less than 1 % of the FTSE volatility. The French market is most connected with news sentiment regarding Brent (2.27 %) and the least connected with Twitter sentiment regarding the DAX (0.26 %). For the German market, news sentiment regarding the DJIA and Brent transmit the largest shocks to the DAX, with magnitudes of 2.55 % and 2.5 %, respectively. However, CAC Twitter sentiment is the least connected with the DAX, with a magnitude of just 0.3 %. Twitter sentiment regarding the German market displays only negligible connectedness with the Nikkei (0.67 %). In addition, the Nikkei is most connected with the DJIA in terms of news sentiment (1 %). In regard to Twitter sentiment, however, Brent and the FTSE send the greatest shocks to the Nikkei, with values of 1.03 % and 1 %, respectively. Lastly, Brent has the strongest connectedness with FTSE Twitter sentiment (2.05 %) and the weakest connectedness with DJIA Twitter sentiment (0.15 %).

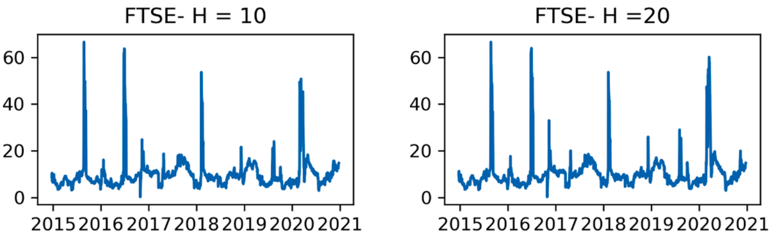
We examine the validity of these findings using Fisher's test, which checks the significance of connectedness between volatility and international sentiment. From the last row in Table 4, we see that the p-value is zero. Therefore, we reject the no-connectedness assumption at any conventional levels and confirm that there is a statistically significant association between market volatility and

[7] https://www.pewresearch.org/global/2018/01/11/publics-around-the-world-follow-national-and-local-news-more-closely-than-international/
.

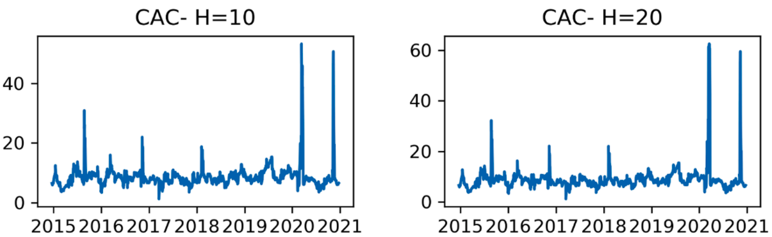
A) Transmitted volatility to DJIA



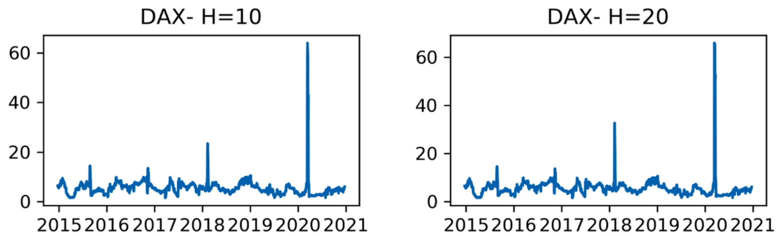
B) Transmitted volatility to FTSE100



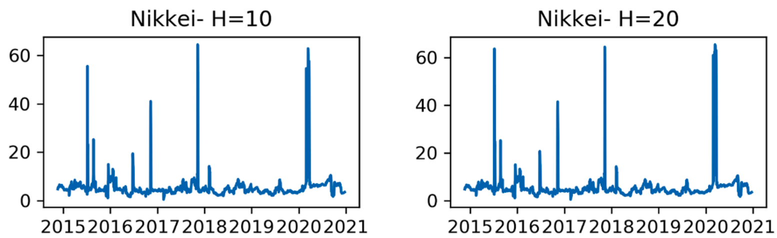
C) Transmitted volatility to CAC40



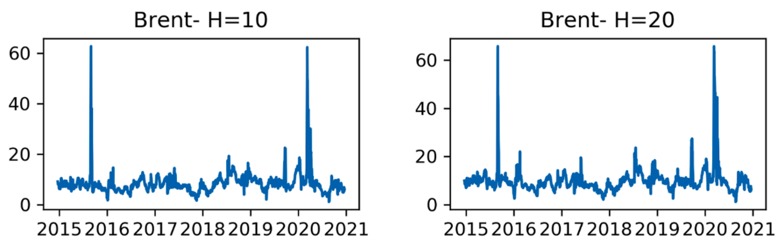
D) Transmitted volatility to DAX



E) Transmitted volatility to Nikkei



F) Transmitted volatility to Brent



(caption on next page)

Fig. 6. Sensitivity Analysis for Dynamic Connectedness. This figure presents the robustness check of the transmitted volatility from sentiment to the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets through panels A-F over the period from 04/08/2014 to 22/12/2020 through a 100-day rolling window with different forecast horizons.

international sentiment. These results are consistent with Hypothesis 2, indicating that volatility is transmitted between markets through news stories and social media.[8]

5.3. Dynamic connectedness

Hypothesis 3 predicts that the volatility arising from media sentiment and transmitted to the markets follows a spiky pattern. To test this hypothesis, we must examine volatility connectedness over a rolling window.[9] A pitfall of a VAR-based framework is that the same VAR parameters are unlikely to last over the entire sample (Lovcha and Perez-Laborda, 2020). Thus, the drawback of using the static model is that the coefficients cannot vary over time, and therefore it is unable to capture the evolution of time-varying dependencies. As one of the objectives of this study is to determine the trend of volatility transmission from sentiment to the markets over time and given the variability in real-time news and social media sentiment, we need to examine connectedness using a rolling window. Drawing from the literature, we choose 100-day subsamples (Wang et al., 2021) and re-estimate Equations (5) and (7) to find the dynamics.

Fig. 5 portrays the volatility connectedness between markets and relevant sentiment series across the rolling subsamples. Connectedness varies significantly over time, ranging from almost zero to more than 60 %. The variation between static and dynamic results is related to the fact that the VAR calculated over the entire sample smooths the results when the relation between variables is not constant (see Lovcha and Perez-Laborda, 2020). For instance, assume we divide a sample period into two subsamples of similar length. Shocks to the first variable negatively affect the second variable in one subsample, while the effect of the same shock is positive with a similar magnitude in the other. By estimating the VAR in each subsample, we can discern the magnitude and direction of these variables, and we obtain the degree of connectedness separately for each subsample based on their VAR estimations. Nevertheless, the VAR used for the entire sample tends to fit the two subsamples because we take the average of both positive and negative effects transmitting from the first to the second variable. This effect explains why the connectedness in the entire sample tends to be lower than the connectedness in the subsamples.

Fig. 5 shows that the volatility connectedness between each market and sentiment series varies across time. That is, the large degree of volatility in one series does not always spill over into other series. Seemingly, during some periods, sentiment or information circulation is of less importance. On average, spillovers from sentiment to the markets are around 10 % or lower in all markets. However, we identify some noteworthy contagions, highlighting the need for further investigation. Below, we discuss some of the major events.

On 9 July 2015, news and tweets regarding the Greek debt crisis and an unstable Chinese stock market sent shocks through the financial markets. The Japanese market was influenced the most, accounting for 63 % of the volatility in the Nikkei. Moreover, on 24 August 2015, news about uncertainty surrounding whether the Federal Reserve would increase interest rates attracted exceptional attention. This uncertainty coincided with the diffusion of “Black Monday” news in the Chinese market.[10] Later that day, *The New York Times* used the term “upheaval” to depict the condition of the markets. This eventually transmitted a dramatic shock to global stock markets, with the DJIA and FTSE receiving the biggest shocks, followed by the Nikkei the next day (25 % on 25 August 2015). Intensive connectedness arising from this setting lasted until 9 September 2015. The oil market also displayed increased connectedness from 27 August to 7 September 2015, when oil prices had their biggest three-day rally since January 2009. This rally stoked excitement among Twitter users. The move coincided with the release of the OPEC bulletin announcing that OPEC was willing to talk to other producers on a level playing field to protect their own interests. During this same period, Saudi (as the leader of OPEC) and Russian officials met about oil prices. These events eventually elevated the connectedness to 65 %.

From 24 June to 11 July 2016, the Brexit vote and Britain’s decision to leave the European Union produced a strong connectedness between sentiment and markets. During this period, the media was rife with speculative articles, opinions, and predictions, leading to an atmosphere of heightened uncertainty. The magnitude of the connectedness peaked at 63 % for the FTSE index. Such a dramatic response in the connectedness underscores not only the weight of major geopolitical events on financial markets but also the rapid and forceful transmission of sentiments in shaping market behaviors and outcomes. On 9 November 2016, the announcement of Donald Trump’s victory in the US presidential election sent a massive shock to the UK, French, and Japanese markets. The connectedness remained higher than usual for a few days thereafter, partly due to varying interpretations and uncertainty about the meaning of Trump’s “America first” motto. In 2017, the implementation of oil production cuts and the announcement of a reduction in the global supply of nearly 1.5 million barrels per day pushed the connectedness magnitude to 20 % in the oil market.

In early February 2018, an announcement from the Federal Reserve about an inflation rate increase significantly impacted the DJIA on February 5–6. Concurrently, as the stock market faced a sharp decline, then-US President Donald Trump chose not to address the market’s performance in his speech, despite his usual penchant for doing so. This deviation drew notable media attention, further intensifying market sentiments. The shock intensity, however, was tempered for other markets. Also in February 2018, the Bank of England noted that

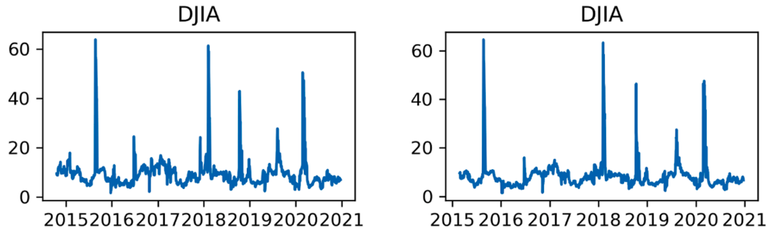
[8] We also check the connectedness by slicing the sample into low-, medium-, and high-volatility regimes. The results, presented in Appendix A3, do not change materially.

[9] The static connectedness between markets and their corresponding sentiment series assumes that spillover coefficients in the VAR analysis are time-invariant. This assumption fails to capture cyclical episodes or shifts in regimes across series. To overcome this shortcoming, we investigate the dynamic connectedness using a rolling window.

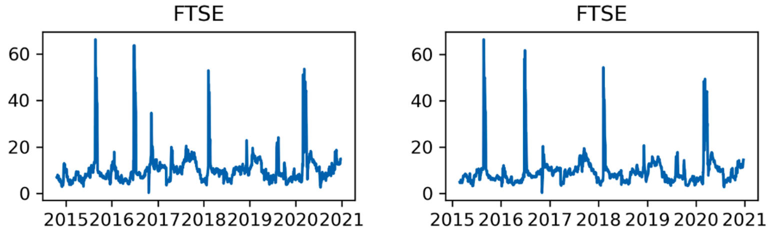
[10] The Chinese state media coined this term following an 8% drop in one day. This was the biggest drop in the Chinese market since 2007.

50 days **150 days**

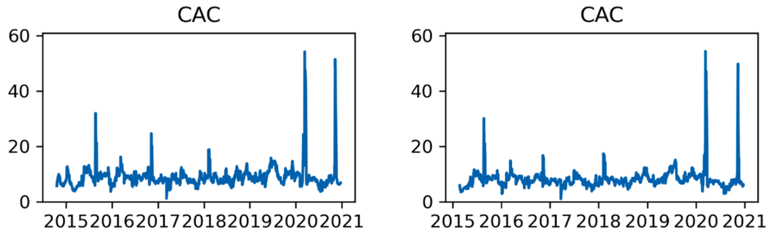
A) Transmitted volatility to DJIA



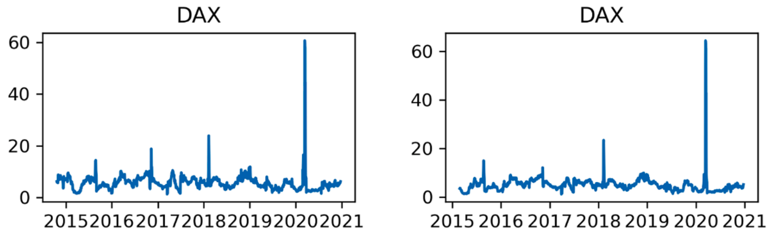
B) Transmitted volatility to FTSE100



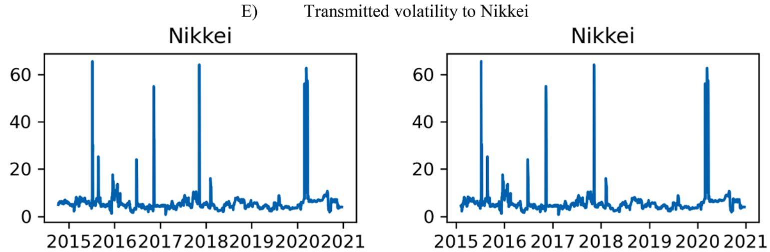
C) Transmitted volatility to CAC40



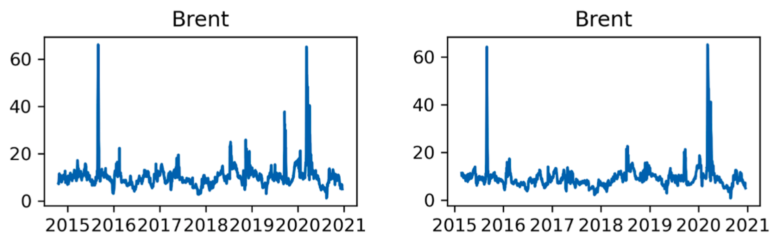
D) Transmitted volatility to DAX



E) Transmitted volatility to Nikkei



F) Transmitted volatility to Brent



(caption on next page)

Fig. 7. Sensitivity Analysis for Dynamic Connectedness. This figure presents the robustness check of the transmitted volatility from sentiment to the US (DJIA index), the UK (FTSE100 index), France (CAC40 index), Germany (DAX index), Japan (Nikkei index), and Brent crude oil markets through panels A-F over the period from 04/08/2014 to 22/12/2020 through rolling windows of 50 and 150 days with a 15-day forecast horizon.

UK interest rates were likely to increase earlier and faster than expected. During the first half of October, tariff clashes between the US and China, together with rising interest rates and concerns about possible overvalued US technology stocks, sent massive shocks to the DJIA. In December, the UK prime minister survived a no-confidence vote, lowering concerns in the media of turmoil in England's politics. During this time, the connectedness increased to around 25 %. In 2019, the US president announced in a series of tweets that a new tariff plan would be imposed on Chinese goods, causing a harsh reaction from China in August. Meanwhile, the conflict between Saudi and Russia over compliance with production cuts added more uncertainty to oil markets. The final dramatic shock during our sample period, the World Health Organization's (WHO) declaration characterizing COVID-19 as a pandemic, took place in March 2020. As media sentiment heightened, investors swiftly recalibrated their expectations. This sentiment-driven shift had a pronounced impact on all markets, inducing spikes in volatility largely attributed to intensified sentiment and the resulting investor apprehension. In Japan, media concerns about disrupted global trade were particularly salient. The Brent market, on the other hand, responded to sentiments of a precipitous decline in global oil demand in light of economic slowdowns. Our analyses indicate that at this point, over 50 % of the observed market volatilities were directly tied to changes in media sentiment.

News stories can also have a calming effect on the market. Such an impact is generally associated with high media coverage. For instance, the January 2020 statement by the WHO clarifying that the "coronavirus is not yet a global emergency" calmed market volatility. Another notable example arises from the oil market, which seems capable of smoothing the effect of anticipated events well beforehand. For instance, OPEC meetings are scheduled well in advance, and officials usually express their views in the press or on social media beforehand. Thus, the outcomes of the meetings can almost always be correctly predicted and gradually absorbed by the market.

As a formal test of our third hypothesis, we define a spiky pattern event as a period with connectedness above 20 %. We then use an F-test to compare volatility during such events to volatility during periods with no spikes. In Table 5, the F-statistic is 11.2082 and the p-value is zero. These results are consistent with Hypothesis 3, as sentiment extracted from news headlines and Twitter feeds has a noticeable impact on market volatility. Given the durability of the contagions, we conclude that the information flow seems speedy and spiky throughout major news events. During nonspiky times, spillover from sentiment to the markets still occurs but is weaker. We interpret these results as evidence that markets are able to process new information rapidly.

We also examine the robustness of the estimated results to check the validity of the findings using rolling-window subsamples. Similar to Diebold and Yilmaz (2009) and Lovcha and Perez-Laborda (2020), we examine the sensitivity of volatility spillovers to the forecast horizon and the length of rolling window. Fig. 6 depicts volatility transmitted to each market for 10- and 20-day forecast horizons. The sensitivity test shows no significant variation among the time-varying results. That is, the results for volatility connectedness are insensitive to different values for the H-step-ahead forecast horizon. Fig. 7 portrays volatility transmitted to each market over rolling windows of 50 and 150 days. We see no significant variation among the time-varying results. That is, the results for volatility connectedness are also insensitive to different values for rolling windows. In the robustness check, the 100-day rolling window serves as our benchmark. Moreover, we test our findings using different measures of volatility, as detailed in Appendix A4, and extend our robustness checks with the application of the time-varying parameter vector autoregressions (TVP-VAR) model, as presented in Appendix A5. Across these diverse measures and frameworks, our results consistently hold. Therefore, we conclude that our main results are robust to different forecast horizons, rolling windows, measures of volatility, and model configurations.

6. Conclusion

In this paper, we investigate the connectedness between media and market volatility across time. Although the finance literature suggests that media affects market volatility, empirical evidence of the time-varying evolution of such a relation is ambiguous. Additionally, much of the research has centered on economic news from official announcements and news outlets rather than covering the whole spectrum of news. Social media has, for instance, been an increasingly important arena for financial news and opinions in recent years. In addition, new textual analytics methods allow a level of precision in analyzing news that was simply not available only a few years ago.

We extract news sentiment by scraping the *investing.com* website, which covers a wide variety of news posted by various prestigious news outlets. The public may have distinct interpretations of the news. Those interpretations and ideas can be found on social media such as Twitter. We gather a total of 651,432 news stories and tweets for markets based in the US, UK, France, Germany, and Japan as well as Brent crude oil using a financially fine-tuned BERT model to extract daily sentiment for each market from August 2014 to December 2020. We then use the daily sentiment series together with historical data to examine the connectedness between news and market volatility through the DY framework. We find that news and social media sentiment have a long-lasting impact on international market dynamics. Our results also indicate that news germane to one market can transmit volatility to other markets. Lastly, we find that the connectedness between news sentiment and financial markets follows a spiky pattern.

These findings have several empirical implications. Investors, risk managers, hedgers, and other decision makers should consider the important role of news in their decision-making. Not only does media sentiment occasionally transmit massive shocks to markets, but we find evidence of a relatively constant spillover between sentiment and markets. These insights can be particularly interesting for regulatory authorities and market institutions, which may use sentiment analysis to monitor market volatility and distress.

The findings also offer theoretical insights. There is a substantial literature in classical finance on how information affects markets, but much less on how it is actually transmitted. Our work shows that whether it is fundamental information or noise, news and social media are important for transmitting information to financial markets, and the importance varies considerably over time.

Because news and social media sentiment play a role in market dynamics, future models of investment strategies, index tracking, or return/volatility forecasting should incorporate this variable, along with other sources of market volatility. Given the long-lasting effect of sentiment on market volatility we document, future research could investigate how media sentiment impacts market predictability. Future research could also examine the role of non-English news or weight news headlines based on the popularity/viewership of the media outlet and tweets based on their retweet number. Another avenue for future work would be identifying cross-sectional patterns of sentiment-driven volatility or studying the frequency connectedness.

CRediT authorship contribution statement

Hooman Abdollahi: Conceptualization, Methodology, Software, Data Curation, Formal analysis, Writing - Original Draft, Writing - Review & Editing. **Sturla L. Fjesme:** Methodology, Supervision, Formal analysis, Writing - Original Draft, Writing - Review & Editing. **Espen Sirnes:** Methodology, Supervision, Formal analysis, Writing - Original Draft, Writing - Review & Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

We would like to thank Denis Becker, Derek John Clark, Karen DeVivo, Jens Kvaerner, Mikko Moilanen, Knut Nygaard, Oystein Strom, and Dan Zhang for helpful comments.

Appendices

A1

Table A1
Hashtags Searched for Tweet Curation. This table presents different hashtags used to collect pertinent tweets for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets with own and other news and Twitter sentiments over the entire period from 04/08/2014 to 22/12/2020.

Index	Hashtags
Brent	#brentoil, #oilprice, #wtioilprice, #oilmarket
DJIA	#DowJonesIndustrial, #DowJonesIndustrialAverage, #DJI, #DJIA, #DowJonesIndustrialIndex, #DowJones, #DowJonesIndex
FTSE	#FTSE, #FTSE100index, #Footsie #FTSE100, #LondonStockExchange, #englishstockmarket
CAC40	#EuronextParis, #francestockmarket, #frenchstockexchange, #BoursedeParis, #frenchstocks, #cac40, #parisstockexchange, #Frenchstockmarket, #frenchstockexchange
DAX	#DAX, #GermanyDAX30, #DeutscherAktienindex, #Germanstockindex, #GermanyStocks, #DeutscheBörse, #GermanStockExchange, #FrankfurtStockExchange, #FrankfurtStockMarket, #DAXPerformanceIndex
Nikkei	#nikkei225, #日経平均株価, #NikkeiStockAverage, #TokyoStockExchange, #TokyoStockMarket, #Nikkeiindex, #Japanstocks, #JapanStockExchange, #JapanStockMarket, #JapaneseStockMarket, #JapaneseStockExchange, #NikkeiIndexes

A2

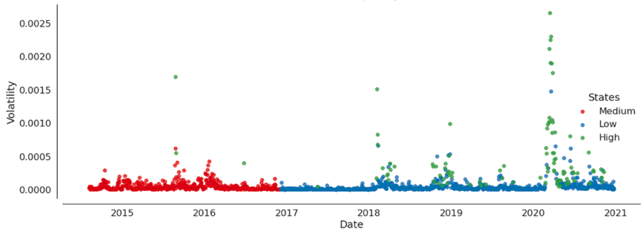
Table A2
Textual Analysis Using Fine-Tuned BERT. Table A2 provides an example of the textual analysis performed in this study. A given text is the input of algorithm, while the sentiment score is the final output. The Logit columns show the estimated probability as to classification for each sentence. The class (positive, negative, or neutral) with the greatest value determines the label of the given sentence. However, the sentiment score is calculated by subtracting negative from positive probability. The more difference in positive and negative logit values, the more positive or negative the sentence is.

Input text	Logit			Label	Score
	Positive	Negative	Neutral		
Oil pushes up on US inventory drop, supply restrictions.	0.7833	0.1227	0.0939	Positive	0.6606
Asian stock markets closed down on January 3rd and European ones opened lower.	0.0837	0.8450	0.0711	Negative	−0.7612
No dodging the oil bullet as emerging economies risk demand hit.	0.1614	0.1697	0.6687	Neutral	−0.0082

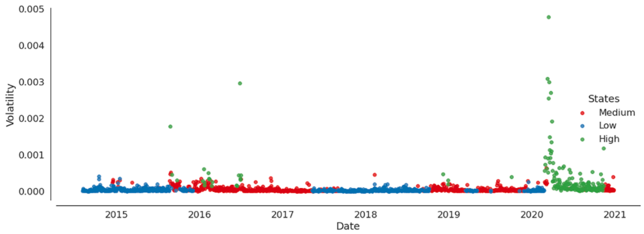
A3

We also examine the connectedness by splitting data samples. The aim is to check the connectedness within low/medium/high volatility states over time. To obtain a valid approximation as to dividing the time series into different volatility regimes, we use hidden Markov model which is an unsupervised machine-learning method for regime detection in stochastic time series. From a statistical vantage point, the hidden Markov model provides a more realistic depiction of the dynamics embedded in financial time series than linear models with constant variance. The core assumption of the model is the existence of hidden states that are not directly observable but have an impact on the observable values (observations). The model determines the existing hidden states from the observations. The known observations, in this case, are the returns that are indirectly affected by the hidden regimes of the market. Fitting the hidden Markov model to return data results in volatility regimes detection. From a quantitative finance perspective, the various regimes lead to adjustments of returns through changes in their means, variances or volatilities, covariances, and serial correlations. A first-order hidden Markov model has two assumptions: (i) The probability of a future state depends only on the current state (Markov assumption); (ii) The probability of an output observation depends only on the state that generated that observation and not on any other states or observations (output independence). We use a standard application of hidden Markov model here; further explanations as to model configuration can be seen in Bhar and Hamori (2004).

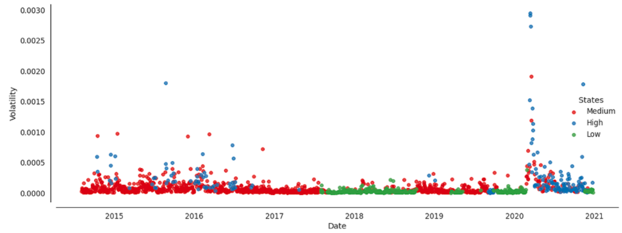
G) DJIA Volatility Regimes



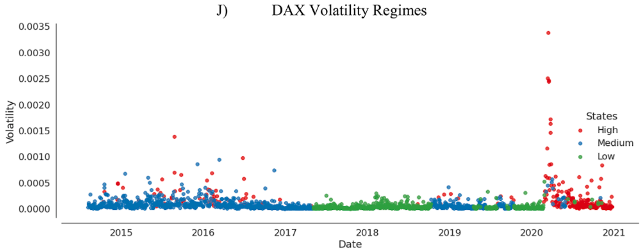
H) FTSE Volatility Regimes



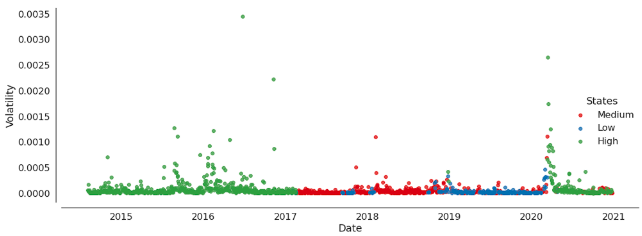
I) CAC40 Volatility Regimes



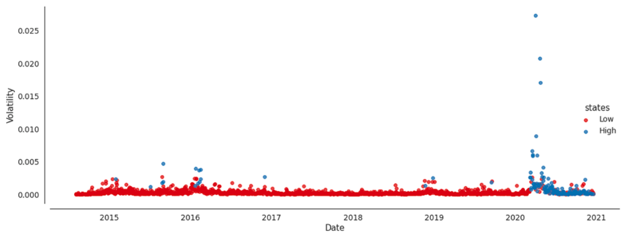
J) DAX Volatility Regimes



K) Nikkei Volatility Regimes



L) Brent Volatility Regimes



(caption on next page)

Fig. A3. Visualization of the Obtained Regimes Detected by a Hidden Markov Model. This figure presents low, medium, and high volatility regimes for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets through panels A-F over the period from 04/08/2014 to 22/12/2020 using hidden Markov model for regime detection.

We re-estimate the DY connectedness to obtain the estimations during various volatility regimes for each market. Table A3 presents the duration, mean and variance, and volatility transmitted to each market by corresponding sentiments over the detected regimes for each index. The calculated means and variances are obtained using the return data for each regime. The lowest variance for returns indicates the low-volatile regime, the second-lowest variance shows medium volatility, and the highest variance for returns implies the high-volatile regime for each market. Columns 'news' and 'Twitter' present the contributions of sentiments to the market volatility within each regime. Results show that connectedness is higher over high-volatile regimes in all series which can be indicative of the contagion effect (sharp shocks) coming from news and Twitter to the market. Also, news sentiment sends more volatility to the markets in high-volatile regimes than Twitter sentiment. There is also a lasting connectedness over the periods with lower volatility which documents volatility spillovers from sentiment to the markets. Except for two low-volatile regimes in the CAC40 and Brent, twitter sentiment contributes to market volatility more than or almost equal with the news contributions. During medium sentiment the degree of connectedness varies between Twitter and news sentiment. For the DJIA and the FTSE, Twitter sends more volatility than news, while the opposite is true for the CAC40 and the Nikkei. Although there are periods in which the estimated connectedness is negligible, all the coefficients are statistically significant except for one occasion for the DAX Twitter sentiment (medium volatile regime), which signifies that sentiment cannot affect the market at some points.

Table A3

The DY Static Estimations for Various Regimes. This table presents the duration, mean and variance, detected volatility regime based on fitting the hidden Markov model to the corresponding return data, and directional volatility connectedness to each market by related news and Twitter sentiments within the detected regimes for each index. Numbers in parentheses are standard errors based on bootstrapping. ** and * denote significance at 5% and 10% levels, respectively. Panel A presents the estimations for the first two periods, while Panel B shows the same estimations for the last periods.

Panel A.

Index	Period I	[Mean, Variance]	Volatility regime	News	Twitter	Period II	[Mean, Variance]	Volatility regime	News	Twitter
DJIA	04.08.2014-02.12.2016	[0.000353, 0.00007]	Medium	0.47** (0.039)	8.1** (0.146)	05.12.2016-21.02.2020	[0.00135, 0.000042]	Low	7.22** (0.085)	7.53** (0.182)
FTSE	4.08.2014-4.11.2015	[0.000196, 0.000061]	Low	3.04** (0.132)	3.73** (0.115)	5.11.2015-11.05.2017	[0.000174, 0.000101]	Medium	0.47* (0.043)	1.21** (0.078)
CAC	4.08.2014-24.04.2017	(-0.000023, 0.000114)	Medium	2.75** (0.043)	0.28** (0.05)	25.04.2017-05.03.2020	(0.00149, 0.00003)	Low	5.8** (0.093)	1.37** (0.065)
DAX	04.08.2014-02.05.2017	(0.000593, 0.000089)	Medium	1.16** (0.027)	0.05 (0.032)	03.05.2017 –26.02.2020	(0.000609, 0.000078)	Low	0.57** (0.043)	0.54** (0.038)
Nikkei	04.08.2014-15.02.2017	(0.000225, 0.000261)	High	1.81** (0.033)	0.47** (0.039)	16.02.2017-18.12.2018	(-0.000183, 0.000153)	Medium	0.56* (0.025)	0.38* (0.031)
Brent	04.08.2014-05.03.2020	(-0.00061, 0.000406)	Low	1.86** (0.052)	0.10** (0.013)	06.03.2020-22.12.2020	(0.006413, 0.008663)	High	3.19** (0.106)	2.19** (0.096)

Panel B.

Index	Period I	[Mean, Variance]	Volatility regime	News	Twitter	Period II	[Mean, Variance]	Volatility regime	News	Twitter
DJIA	22.02.2020-22.12.2020	[-0.00575, 0.001069]	High	12.3** (0.157)	1.68** (0.163)	—	—	—	—	—
FTSE	12.05.2017-21.02.2020	[0.000196, 0.000061]	Low	0.93** (0.049)	2.47** (0.113)	24.02.2020-22.12.2020	[-0.001942, 0.000698]	High	3.69** (0.212)	3.39** (0.174)
CAC	06.03.2020-22.12.2020	(-0.001298, 0.0007)	High	4.09** (0.178)	0.44* (0.138)	—	—	—	—	—
DAX	27.02.2020–22.12.2020	(-0.001649, 0.000728)	High	1.28** (0.105)	0.92* (0.076)	—	—	—	—	—
Nikkei	19.12.2018-11.03.2020	(0.001213, 0.000037)	Low	0.4** (0.054)	0.39** (0.048)	12.03.2020-22.12.2020	(0.000225, 0.000261)	High	1.68** (0.054)	1.21** (0.043)
Brent	—	—	—	—	—	—	—	—	—	—

A4

We re-do the estimation using two different measures of volatility to further check the robustness of our results. For this purpose, we estimate the volatility using Garman and Klass (1980) and Rogers and Satchell (1991).

A) Garman and Klass volatility measure

Garman-Klass volatility estimator incorporates information about the open, close, high and low prices within a specific time interval.

$$\sigma_t^2 = 0.5\left(\ln\left(\frac{H_t}{L_t}\right)\right)^2 - (2\ln(2)-1)\ln\left(\frac{C_t}{O_t}\right)^2$$

Where H_t, L_t, O_t, and C_t are the maximum, minimum, opening, and closing price, respectively, on day t. We calculate the volatility using Equation A4.1 and re-do the analysis. Table A4.1 presents the results for the static connectedness. No significant variation is seen compared with the original results, and all the coefficients are statistically significant. We also estimate the static framework including all sentiment series in the analysis to scrutinize connectedness between volatility in each market and sentiment of other markets. The results are presented in Table A4.2 and are statistically significant. Finally, we examine the dynamic connectedness using a rolling window of 100 days. Fig. A4.1 portrays the volatility connectedness between markets and relevant sentiment series across the rolling subsamples. We observe no significant variations on the connectedness pattern from the original results.

Table A4.1

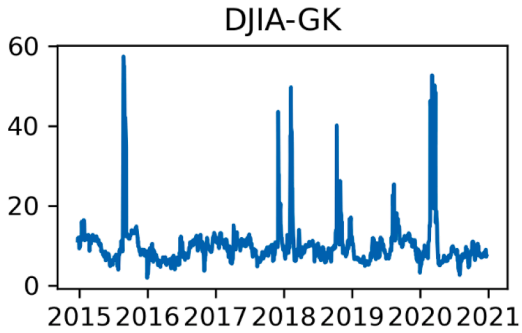
The DY Static Estimations for Each Market and its Corresponding Sentiments. This table presents the estimation of static connectedness for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets with their related news and Twitter sentiment series over the entire period from 04/08/2014 to 22/12/2020. The ij^{th} entry represents the directional connectedness, namely the percentage of the forecast error variance of market i due to shocks from sentiment j. Numbers in parentheses are standard errors based on bootstrapping. ** and * denote significance at 5% and 10% levels, respectively.

Index	News sentiment	Twitter sentiment
DJIA	5.50 (0.0161)**	1.28 (0.0133)**
FTSE	1.06 (0.0129)**	3.74 (0.0419)**
CAC	1.78 (0.0104)**	1.00 (0.0107)**
DAX	0.62 (0.0737)*	0.98 (0.0736)**
Nikkei	1.73 (0.0805)**	1.09 (0.0876)**
Brent	2.14 (0.0968)**	0.97 (0.0101)**

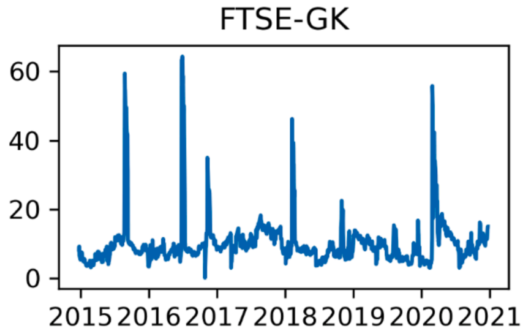
Table A4.2

The DY Static Estimation for all Markets and Sentiment series. This table presents the estimation of static connectedness for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets with own and other news and Twitter sentiment series over the entire period from 04/08/2014 to 22/12/2020. The ij^{th} entry represents the directional connectedness, namely the percentage of the forecast error variance of market i due to shocks from sentiment j. Numbers in parentheses are standard errors based on bootstrapping. ** and * denote significance at 5% and 10% levels, respectively.

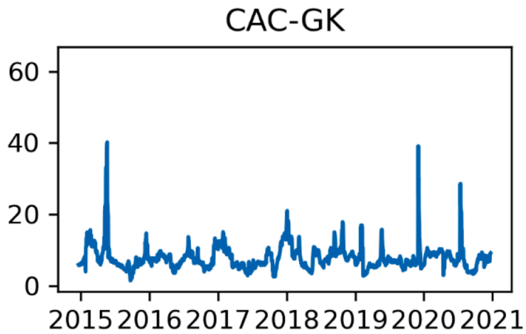
	DJIA News	DJIA Twitter	FTSE News	FTSE Twitter	CAC News	CAC Twitter	DAX News	DAX Twitter	Nikkei News	Nikkei Twitter	Brent News	Brent Twitter
DJIA	3.28**	1.51**	2.68**	0.61**	0.59**	0.69**	1.16**	0.16**	1.58**	1.47**	2.16**	0.78**
	(0.007)	(0.009)	(0.008)	(0.012)	(0.008)	(0.005)	(0.007)	(0. 005)	(0.007)	(0.009)	(0.009)	(0.006)
FTSE	0.56**	0.59**	0.62**	2.28**	0.66**	0.76**	0.55**	0.53**	0.73**	0.99**	0.88**	0.57**
	(0.007)	(0.007)	(0.007)	(0.024)	(0.007)	(0.006)	(0.006)	(0.006)	(0.007)	(0.009)	(0.008)	(0.006)
CAC	1.46**	1.69**	0.43**	2.02**	1.98**	0.41**	0.67**	0.18**	0.95**	0.83**	2.16**	1.21**
	(0.2)	(0.403)	(0.008)	(0.017)	(0.008)	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)	(0.07)	(0.006)
DAX	2.31**	1.18**	0.31**	1.38**	1.92**	0.27*	0.72**	0.63**	0.87**	1.24**	2.45**	0.28**
	(0.004)	(0.012)	(0.006)	(0.02)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.009)	(0.007)	(0.004)
Nikkei	1.04**	0.77**	0.22*	0.96**	0.86**	0.64**	0.33*	0.77*	0.98**	0.69**	1.03**	1.09**
	(0.005)	(0.006)	(0.006)	(0.05)	(0.006)	(0.005)	(0.005)	(0.009)	(0.006)	(0.006)	(0.05)	(0.009)
Brent	0.72**	0.36**	0.81**	2.08**	0.89**	0.25*	0.6**	0.31*	0.92**	0.72**	1.15**	0.92**
	(0.008)	(0.008)	(0.009)	(0.028)	(0.008)	(0.01)	(0.008)	(0.009)	(0.008)	(0.007)	(0.008)	(0.009)



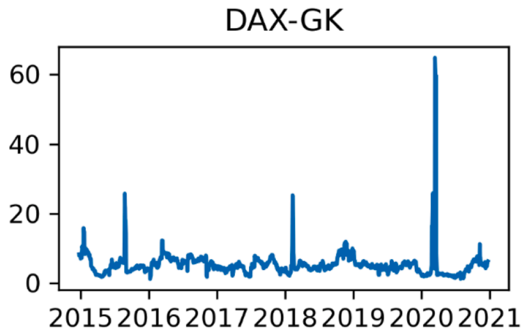
Transmitted volatility to DJIA



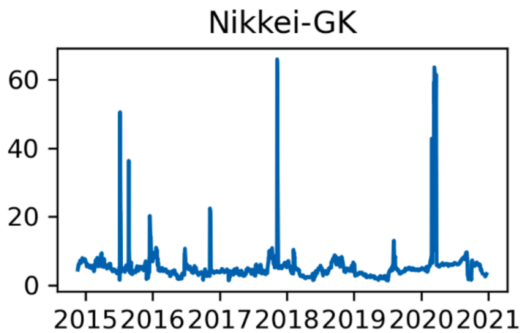
Transmitted volatility to FTSE100



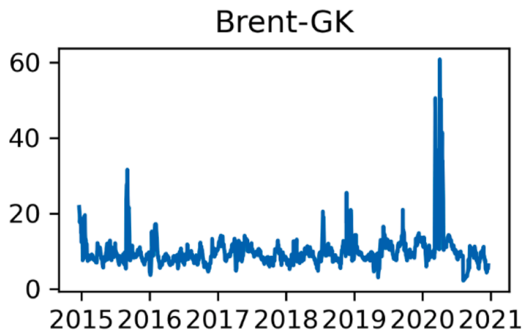
Transmitted volatility to CAC40



Transmitted volatility to DAX



Transmitted volatility to Nikkei



Transmitted volatility to Brent

Fig. A4.1. Dynamic Connectedness. This figure presents the robustness check of the transmitted volatility from sentiment to the US (DJIA index), the UK (FTSE100 index), France (CAC40 index), Germany (DAX index), Japan (Nikkei index), and Brent crude oil markets over the period from 04/08/2014 to 22/12/2020 using Garman-Klass measure of volatility.

B) Rogers and Satchell volatility measure

Rogers-Satchell volatility estimator also incorporates drift term (mean return not equal to zero).

$$\sigma_t^2 = \ln\left(\frac{H_t}{C_t}\right)\ln\left(\frac{H_t}{O_t}\right) + \ln\left(\frac{l_t}{C_t}\right)\ln\left(\frac{l_t}{O_t}\right) \tag{A42}$$

Where H_t, L_t, O_t, and C_t are the maximum, minimum, opening, and closing price, respectively, on day t. We calculate the volatility using Equation A4.2 and re-do the analysis. Table A4.3 presents the results for the static connectedness. No significant variation is seen compared with the original results, and all the coefficients are statistically significant. We also estimate the static framework including all sentiment series in the analysis to scrutinize connectedness between volatility in each market and sentiment of other markets. The results are presented in Table A4.4 and are statistically significant. Finally, we examine the dynamic connectedness using a rolling window of 100 days. Fig. A4.2 portrays the volatility connectedness between markets and relevant sentiment series across the rolling subsamples. We observe no significant variations on the connectedness pattern from the original results.

Table A4.3

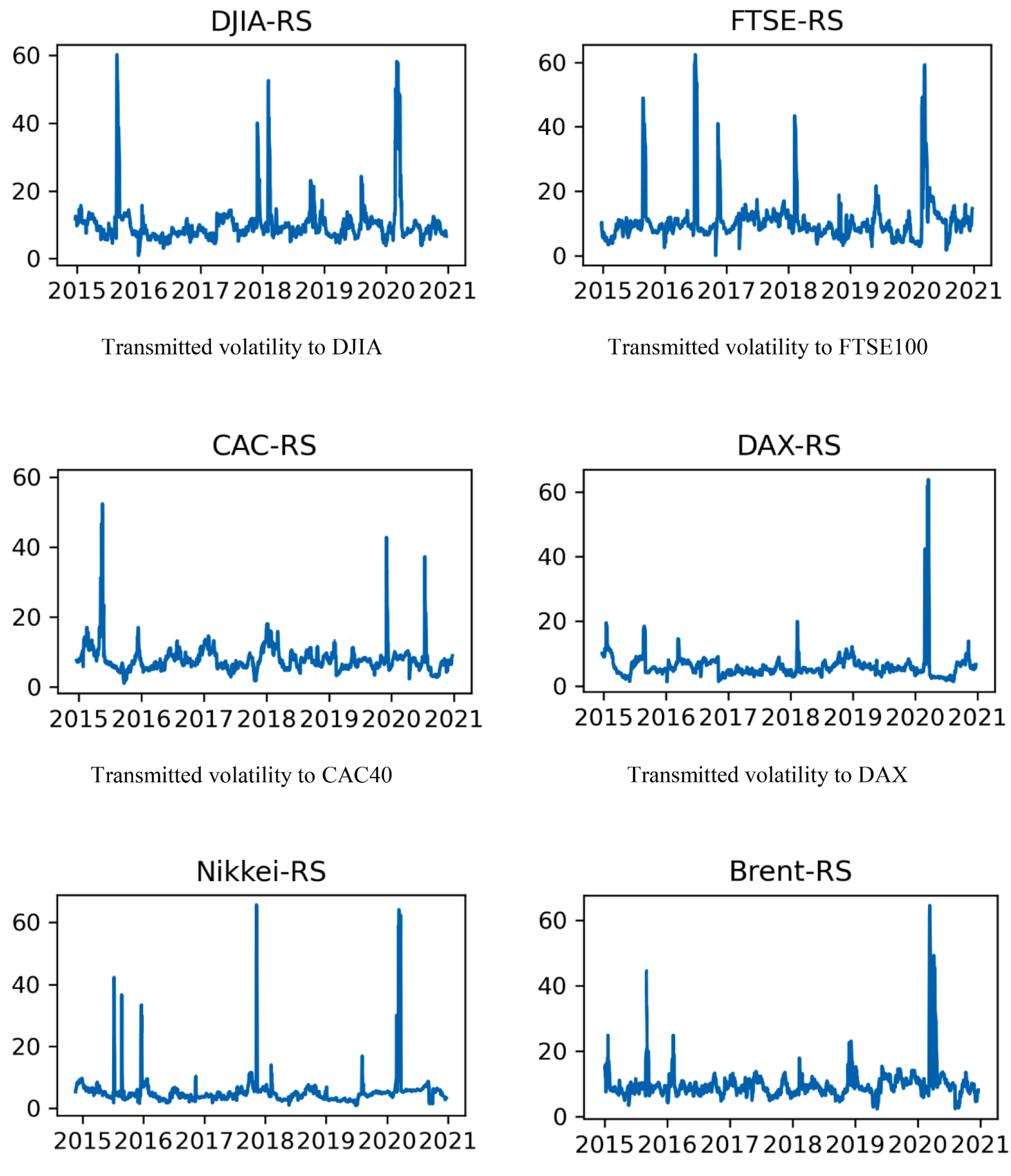
The DY Static Estimations for Each Market and its Corresponding Sentiments. This table presents the estimation of static connectedness for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets with their related news and Twitter sentiment series over the entire period from 04/08/2014 to 22/12/2020. The ij^{th} entry represents the directional connectedness, namely the percentage of the forecast error variance of market *i* due to shocks from sentiment *j*. Numbers in parentheses are standard errors based on bootstrapping. ** and * denote significance at 5% and 10% levels, respectively.

Index	News sentiment	Twitter sentiment
DJIA	4.82 (0.0124)**	1.16 (0.0156)**
FTSE	0.89 (0.0107)*	3.44 (0.0343)**
CAC	1.44 (0.01)**	0.88 (0.0106)**
DAX	0.65 (0.0074)**	0.87 (0.0076)**
Nikkei	1.4 (0.0853)**	1.00 (0.0784)**
Brent	2.05 (0.0912)**	0.9 (0.0991)*

Table A4.4

The DY Static Estimation for all Markets and Sentiment series. This table presents the estimation of static connectedness for the US (the DJIA index), the UK (the FTSE100 index), France (the CAC40 index), Germany (the DAX index), Japan (the Nikkei index), and Brent crude oil markets with own and other news and Twitter sentiment series over the entire period from 04/08/2014 to 22/12/2020. The ij^{th} entry represents the directional connectedness, namely the percentage of the forecast error variance of market *i* due to shocks from sentiment *j*. Numbers in parentheses are standard errors based on bootstrapping. ** and * denote significance at 5% and 10% levels, respectively.

	DJIA News	DJIA Twitter	FTSE News	FTSE Twitter	CAC News	CAC Twitter	DAX News	DAX Twitter	Nikkei News	Nikkei Twitter	Brent News	Brent Twitter
DJIA	2.92**	1.43**	2.68**	0.58**	0.68**	0.71**	1.12**	0.27**	1.55**	1.44**	2.01**	0.77**
	(0.007)	(0.009)	(0.007)	(0.012)	(0.008)	(0.005)	(0.007)	(0. 005)	(0.007)	(0.01)	(0.008)	(0.005)
FTSE	0.67**	0.58**	0.65**	1.73**	0.71**	0.71**	0.67**	0.54*	0.84**	1.07**	0.91**	0.58**
	(0.007)	(0.007)	(0.007)	(0.024)	(0.097)	(0.008)	(0.007)	(0.007)	(0.01)	(0.008)	(0.008)	(0.007)
CAC	1.71**	1.81**	0.38**	1.93**	1.72**	0.28**	0.61**	0.31**	0.88**	0.55**	2.08**	1.2**
	(0.202)	(0.381)	(0.008)	(0.017)	(0.009)	(0.006)	(0.006)	(0.007)	(0.007)	(0.008)	(0.06)	(0.005)
DAX	1.89**	1.31**	0.35**	1.05**	1.95**	0.26*	0.78**	0.67**	1.02**	1.25**	2.31**	0.34**
	(0.005)	(0.012)	(0.007)	(0.018)	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.01)	(0.007)	(0.004)
Nikkei	1.03**	0.93**	0.25*	1.05**	0.66**	0.65**	0.37*	0.75*	1.07**	0.68**	0.87**	1.23**
	(0.005)	(0.006)	(0.006)	(0.05)	(0.005)	(0.006)	(0.005)	(0.009)	(0.006)	(0.006)	(0.05)	(0.009)
Brent	0.65**	0.43**	0.79**	2.34**	1.00**	0.27*	0.68**	0.34*	0.92**	0.73**	1.21**	1.07**
	(0.007)	(0.008)	(0.008)	(0.028)	(0.007)	(0.009)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)	(0.009)



Transmitted volatility to DJIA

Transmitted volatility to FTSE100

Transmitted volatility to CAC40

Transmitted volatility to DAX

Transmitted volatility to Nikkei

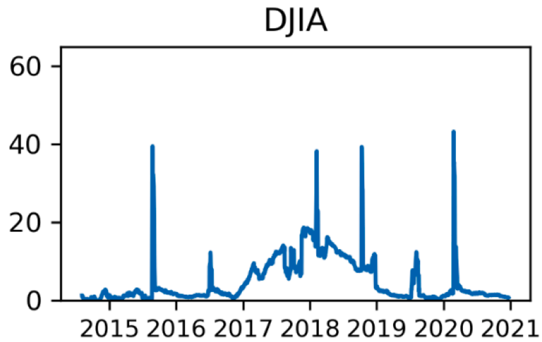
Transmitted volatility to Brent

Fig. A4.2. Dynamic Connectedness. This figure presents the robustness check of the transmitted volatility from sentiment to the US (DJIA index), the UK (FTSE100 index), France (CAC40 index), Germany (DAX index), Japan (Nikkei index), and Brent crude oil markets over the period from 04/08/2014 to 22/12/2020 using Rogers-Satchell measure of volatility.

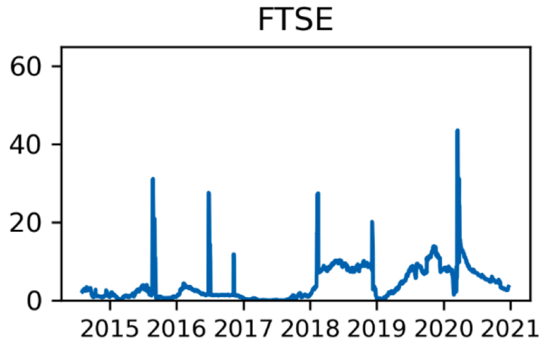
A5

We re-do the analysis using the Time-Varying Parameter Vector Autoregression (TVP-VAR) connectedness framework proposed by Antonakakis et al. (2020). This model is designed as an enhancement to the DY framework. The TVP-VAR offers greater flexibility. Central to this framework is the adoption of forgetting factors derived from Koop and Korobilis (2014). These factors facilitate the inclusion of time-varying parameter, enabling the model to capture dynamic relationships with more resilience, especially in situations characterized by shifting connectedness at lower frequencies. A notable strength of the TVP-VAR framework lies in its robustness to limited time-series data and potential outliers. This reduces the potential for data attrition. While our dataset predominantly utilizes high-frequency data, the merits of the TVP-VAR model make it a tool to verify the robustness of our findings. If our findings regarding the spiky pattern in the dynamic connectedness remain consistent even with this distinct modeling approach, it would reaffirm the resilience of our results.

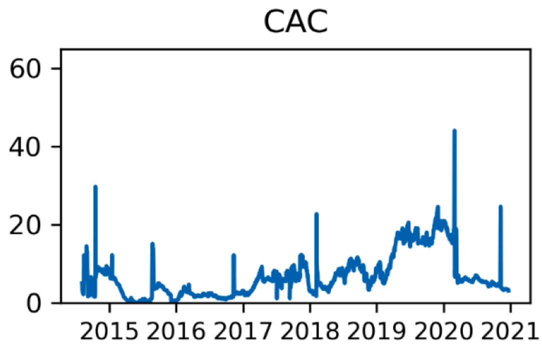
To validate this, we estimate the dynamic connectedness using the TVP-VAR model over a 15-day forecast horizon. Fig. A5 depicts the volatility connectedness among markets and the sentiment series. Although slight variations in connectedness magnitude are noted, the spiky pattern remains evident across all markets. On the whole, while the magnitude of connectedness showed a dampened trend, the distinctive spiky pattern remains consistent.



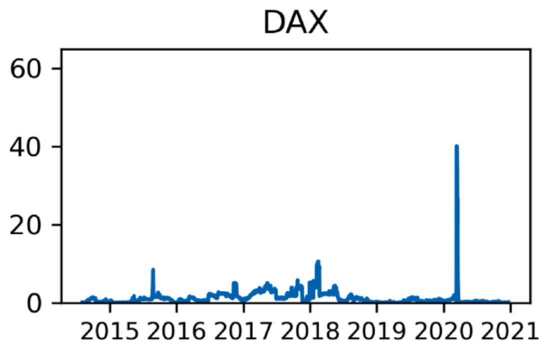
Transmitted volatility to DJIA



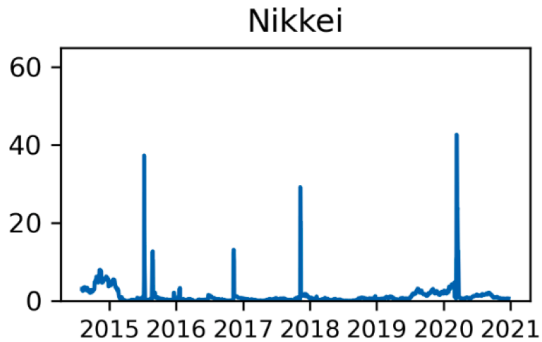
Transmitted volatility to FTSE100



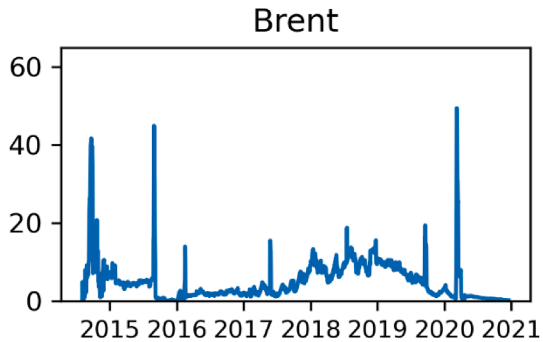
Transmitted volatility to CAC40



Transmitted volatility to DAX



Transmitted volatility to Nikkei



Transmitted volatility to Brent

Fig. A5. Dynamic Connectedness. This figure presents the transmitted volatility from sentiment to the US (DJIA index), the UK (FTSE100 index), France (CAC40 index), Germany (DAX index), Japan (Nikkei index), and Brent crude oil markets over the period from 04/08/2014 to 22/12/2020 using TVP-VAR framework.

References

Aharon, D. Y., Umar, Z., Aziz, M. I. A., & vinh Vo, X. (2022). COVID-19 related media sentiment and the yield curve of G-7 economies. *The North American Journal of Economics and Finance, 61*, Article 101678.

Antonakakis, N., Chatziantoniou, I., & Gabauer, D. (2020). Refined measures of dynamic connectedness based on time-varying parameter vector autoregressions. *Journal of Risk and Financial Management, 13*(4), 84.

Apergis, N., Chatziantoniou, I., & Gabauer, D. (2023). Dynamic connectedness between COVID-19 news sentiment, capital and commodity markets. *Applied Economics, 55*(24), 2740–2754.

Baeza-Yates, R., & Ribeiro-Neto, B. (1999). *Modern information retrieval* (Vol. 463). New York: ACM press.

Baker, S.R., Bloom, N., Davis, S. and Renault, T., 2021. Twitter-derived measures of economic uncertainty.

Basak, G. K., Das, P. K., Marjit, S., Mukherjee, D., & Yang, L. (2023). The British Stock Market, currencies, brexit, and media sentiments: A big data analysis. *The North American Journal of Economics and Finance, 64*, Article 101861.

Behera, C., & Rath, B. N. (2022). The connectedness between Twitter uncertainty index and stock return volatility in the G7 countries. *Applied Economics Letters, 29*(20), 1876–1879.

Bhar, R., & Hamori, S. (2004). *Hidden Markov models: applications to financial economics* (Vol. 40). Springer Science & Business Media.

Birz, G. (2017). Stale economic news, media and the stock market. *Journal of Economic Psychology, 61*, 87–102.

Bouri, E., Demirer, R., Gabauer, D., & Gupta, R. (2022). Financial market connectedness: The role of investors' happiness. *Finance Research Letters, 44*, Article 102075.

Brandt, M. W., & Diebold, F. X. (2006). A No-Arbitrage Approach to Range-Based Estimation of Return Covariances and Correlations. *Journal of Business, 79*(1), 61–74.

Brandt, M. W., & Gao, L. (2019). Macro fundamentals or geopolitical events? A textual analysis of news events for crude oil. *Journal of Empirical Finance, 51*, 64–94.

Brenner, M., Pasquariello, P., & Subrahmanyam, M. (2009). On the volatility and comovement of US financial markets around macroeconomic news announcements. *Journal of Financial and Quantitative Analysis, 44*(6), 1265–1289.

Brown, G. W., & Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of empirical finance, 11*(1), 1–27.

Caldara, D., & Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review, 112*(4), 1194–1225.

Costola, M., Hinz, O., Nofer, M., & Pelizzon, L. (2023). Machine learning sentiment analysis, COVID-19 news and stock market reactions. *Research in International Business and Finance, 64*, Article 101881.

Da, Z., Engelberg, J., & Gao, P. (2011). In search of attention. *The Journal of Finance, 66*(5), 1461–1499.

Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARS investor sentiment and asset prices. *The Review of Financial Studies, 28*(1), 1–32.

De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy, 98*(4), 703–738.

de Oliveira, F. A., Maia, S. F., de Jesus, D. P., & Besarria, C. D. N. (2018). Which information matters to market risk spreading in Brazil? Volatility transmission modelling using MGARCH-BEKK, DCC, t-Copulas. *The North American Journal of Economics and Finance, 45*, 83–100.

Devlin, J., Chang, M.W., Lee, K. and Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Diebold, F. X., & Yilmaz, K. (2009). Measuring financial asset return and volatility spillovers, with application to global equity markets. *The Economic Journal, 119*(534), 158–171.

Diebold, F. X., & Yilmaz, K. (2012). Better to give than to receive: Predictive directional measurement of volatility spillovers. *International Journal of forecasting, 28*(1), 57–66.

Diebold, F. X., & Yılmaz, K. (2014). On the network topology of variance decompositions: Measuring the connectedness of financial firms. *Journal of Econometrics, 182*(1), 119–134.

Dong, Z., Li, Y., Zhuang, X., & Wang, J. (2022). Impacts of COVID-19 on global stock sectors: Evidence from time-varying connectedness and asymmetric nexus analysis. *The North American Journal of Economics and Finance, 62*, Article 101753.

Ederington, L. H., & Lee, J. H. (1993). How markets process information: News releases and volatility. *The Journal of Finance, 48*(4), 1161–1191.

Fedyk, A., 2018. *Front page news: The effect of news positioning on financial markets*. working paper.

Feng, L., Fu, T., & Shi, Y. (2022). How does news sentiment affect the states of Japanese stock return volatility? *International Review of Financial Analysis, 84*, Article 102267.

Gao, Y., Zhao, C., Sun, B., & Zhao, W. (2022). Effects of investor sentiment on stock volatility: New evidences from multi-source data in China's green stock markets. *Financial Innovation, 8*(1), 77.

Garman, M. B., & Klass, M. J. (1980). On the estimation of security price volatilities from historical data. *Journal of business*, 67–78.

Griffin, J. M., Hirschey, N. H., & Kelly, P. J. (2011). How important is the financial media in global markets? *The Review of Financial Studies, 24*(12), 3941–3992.

Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review, 70*(3), 393–408.

Joseph, K., Wintoki, M. B., & Zhang, Z. (2011). Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting, 27*(4), 1116–1127.

Katsafados, A. G., Nikoloutsopoulos, S., & Leledakis, G. N. (2023). Twitter sentiment and stock market: A COVID-19 analysis. *Journal of Economic Studies*.

Klibanoff, P., Lamont, O., & Wizman, T. A. (1998). Investor reaction to salient news in closed-end country funds. *The Journal of Finance, 53*(2), 673–699.

Koch, A., Huynh, T. L. D., & Wang, M. (2022). News sentiment and international equity markets during BREXIT period: A textual and connectedness analysis. *International Journal of Finance & Economics.*

Koop, G., & Korobilis, D. (2014). A new index of financial conditions. *European Economic Review, 71*, 101–116.

Koop, G., Pesaran, M. H., & Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics, 74*(1), 119–147.

Li, X., Shang, W., & Wang, S. (2019). Text-based crude oil price forecasting: A deep learning approach. *International Journal of Forecasting, 35*(4), 1548–1560.

Lin, P. and Luo, X., 2020, October. A survey of sentiment analysis based on machine learning. In *CCF International Conference on Natural Language Processing and Chinese Computing* (pp. 372-387). Springer, Cham.

Liu, J., Ma, F., Tang, Y., & Zhang, Y. (2019). Geopolitical risk and oil volatility: A new insight. *Energy Economics, 84*, Article 104548.

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of finance, 66*(1), 35–65.

Lovcha, Y., & Perez-Laborda, A. (2020). Dynamic frequency connectedness between oil and natural gas volatilities. *Economic Modelling, 84*, 181–189.

Lucca, D. O., & Moench, E. (2015). The pre-FOMC announcement drift. *The Journal of finance, 70*(1), 329–371.

Malo, P., Sinha, A., Korhonen, P., Wallenius, J., & Takala, P. (2014). Good debt or bad debt: Detecting semantic orientations in economic texts. *Journal of the Association for Information Science and Technology, 65*(4), 782–796.

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

Mensi, W., Gubareva, M., Teplova, T., & Kang, S. H. (2023). Spillover and connectedness among G7 real estate investment trusts: The effects of investor sentiment and global factors. *The North American Journal of Economics and Finance, 66*, Article 101919.

Mensi, W., Kamal, M. R., Vo, X. V., & Kang, S. H. (2023). Extreme dependence and spillovers between uncertainty indices and stock markets: Does the US market play a major role? *The North American Journal of Economics and Finance, 68*, Article 101970.

Niederhoffer, V. (1971). The analysis of world events and stock prices. *The Journal of Business, 44*(2), 193–219.

Nishimura, Y., & Sun, B. (2018). The intraday volatility spillover index approach and an application in the Brexit vote. *Journal of International Financial Markets, Institutions and Money, 55*, 241–253.

Nishimura, Y., & Sun, B. (2021). President's Tweets, US-China economic conflict and stock market Volatility: Evidence from China and G5 countries. *The North American Journal of Economics and Finance, 58*, Article 101506.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, 61–65.

Pesaran, H. H., & Shin, Y. (1998). Generalized impulse response analysis in linear multivariate models. *Economics letters, 58*(1), 17–29.

Qiao, X., Zhu, H., Zhang, Z., & Mao, W. (2022). Time-frequency transmission mechanism of EPU, investor sentiment and financial assets: A multiscale TVP-VAR connectedness analysis. *The North American Journal of Economics and Finance, 63*, Article 101843.

Rangel, J. G. (2011). Macroeconomic news, announcements, and stock market jump intensity dynamics. *Journal of Banking & Finance, 35*(5), 1263–1276.

Rogers, L. C. G., & Satchell, S. E. (1991). Estimating variance from high, low and closing prices. *The Annals of Applied Probability*, 504–512.

Ross, S. A. (1989). Information and volatility: The no-arbitrage martingale approach to timing and resolution irrelevancy. *The Journal of Finance, 44*(1), 1–17.

Shen, S., Xia, L., Shuai, Y., & Gao, D. (2022). Measuring news media sentiment using big data for Chinese stock markets. *Pacific-Basin Finance Journal, 74*, Article 101810.
Shiller, R. J. (2017). *Narrative economics. American economic review, 107*(4), 967–1004.
Shiller, R. J. (2020). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.
Su, Z., Fang, T., & Yin, L. (2019). Understanding stock market volatility: What is the role of US uncertainty? *The North American Journal of Economics and Finance, 48*, 582–590.
Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance, 62*(3), 1139–1168.
Wang, P., Zhang, H., Yang, C. and Guo, Y., 2021. Time and frequency dynamics of connectedness and hedging performance in global stock markets: Bitcoin versus conventional hedges. *Research in International Business and Finance*, p.101479.
Williams, L. V. (1999). Information efficiency in betting markets: A survey. *Bulletin of Economic Research, 51*(1), 1–39.
Zeitun, R., Rehman, M. U., Ahmad, N., & Vo, X. V. (2023). The impact of Twitter-based sentiment on US sectoral returns. *The North American Journal of Economics and Finance, 64*, Article 101847.

Paper II

Oil Price Volatility and New Evidence from News and Twitter

Energy Economics 122 (2023) 106711

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneeco

Oil price volatility and new evidence from news and Twitter

Hooman Abdollahi

School of Business and Economics, UiT-The Arctic University of Norway, Tromsø, Norway

ARTICLE INFO

JEL classification:
C53
G12
G17

Keywords:
Oil price volatility
News sentiment
Twitter sentiment
Forecasting
Media and market

ABSTRACT

In this paper, we develop semantic-based sentiment indices through relevant news and Twitter feeds for oil market using a state-of-the-art natural language processing technique. We investigate the predictability of crude oil price volatility using the novel sentiment indices through a hybrid structure consisting of generalized autoregressive conditional heteroskedasticity and bidirectional long short-term memory models. Findings show that media sentiment considerably enhances forecasting quality and the proposed framework outperforms existing benchmark models. More importantly, we compare the predictive power of news stories with Twitter feeds and document the superiority of the news sentiment index over the counterpart. This is an important contribution as this paper is the first study that compares the impact of regular press with that of social media, as an alternative informative medium, on oil market dynamics.

1. Introduction

Playing a significant role in global financial markets, oil is presently known as an alternative investment (Adams et al., 2020; Cui et al., 2021). It is also considered a strategic commodity for economic development across nations (An et al., 2020) such that central banks in many countries regularly update their forecasting of oil prices to remain prepared for possible shocks in the future (Safari and Davallou, 2018). Until recently, the literature has focused on studying oil price volatility using economic fundamentals and statistical data (e. g. Ji et al., 2018; Zhang et al., 2019; Liu and Lee, 2021). Yet, the origin of a part of volatility witnessed in oil prices remains unidentified. The oil market is known as a financial market, and finance theories should be considered in exploring it. According to the finance literature, investors' sensitivity to new information (sentiment) drives a wedge between prices and fundamental values and causes volatility (Ross, 1989). Sentiment can be influenced by various factors such as variations in public opinions, geopolitical conditions, economic news announcements, natural disasters, terroristic attacks, and other exogenous factors (Bomfim, 2003; Brenner et al., 2009; Möbert, 2009; Lucca and Moench, 2015; Birz and Dutta, 2016; Qadan and Nama, 2018). Although these factors are spread via both press and social media, sentiment is not directly measurable. However, recent advancements in natural language processing have provided a tool to quantify the sentiment of textual media. Therefore, investigating the impact of sentiment on forecasting accuracy is a new and growing line of research. Motivated by this issue, we conduct this study to answer the following questions: (i) To what extent does using media sentiment increase the accuracy of oil price volatility forecasting? And (ii) which sentiment proxy is more informative for oil price volatility predictability? Thus, the aim of this study is to shed light on the role of media in oil market predictability.

To answer the questions above, we use weekly data for Brent crude oil prices from August 2014 to December 2020. One of the crucial factors in forecasting oil price volatility is interrelation between financial markets (Abdollahi and Ebrahimi, 2020). Hence, this study also uses historical data for Gold and The Standard and Poor's 500 (S&P 500) index with which the oil market has a high connectedness (Hung, 2022). We also scrape news headlines regarding the oil market as well as tweets mentioning hashtags relevant to the oil market over the same period. The textual materials are then processed using advanced natural language processing to generate two distinct sentiment indices for news and Twitter.

Forecasting oil market volatility is complex because of the various characteristics of oil price time series such as sensitivity to non-fundamental factors (sentiment), nonlinearity, lags, and time-varying volatility (Chen et al., 2016; Abdollahi and Ebrahimi, 2020). To effectively capture these impactful factors, we design a hybrid structure considering the particular ability of each constituent model to capture the aforementioned characteristics. We use the Bidirectional Encoder Representations from Transformers (BERT) to extract sentiment from news headlines and Twitter feeds.[1] Then, we employ a Generalized

E-mail address: hooman.abdollahi@uit.no.

[1] BERT is a state-of-the-art machine learning model that reportedly produces the best results in natural language processing and sentiment classification (González-Carvajal and Garrido-Merchán, 2020).

https://doi.org/10.1016/j.eneco.2023.106711
Received 9 February 2023; Received in revised form 16 April 2023; Accepted 25 April 2023
Available online 5 May 2023

Nomenclature

ARCH	Autoregressive conditional heteroscedasticity
BERT	Bidirectional encoder representations from transformers
BiLSTM	Bidirectional long-short term memory
DM	Diebold-Mariano
GARCH	Generalized autoregressive conditional heteroskedasticity
LSTM	Long short-term memory
MAE	Mean absolute error
RMSE	Root mean square error
S&P 500	The Standard and Poor's 500

Autoregressive Conditional Heteroskedasticity (GARCH) model to obtain an initial estimation of volatility. Using GARCH efficiently captures time-varying volatility and leverage effects (Crawford and Fratantoni, 2003), which are specific characteristics of the oil price time series. In the final step, we use all the parameters as input features for a Bidirectional Long-Short Term Memory (BiLSTM) model that generates the final forecast. Neural networks generally produce the best results for nonlinear settings like oil price volatility. Moreover, there is a time step dimension in the input array of the BiLSTM model. Therefore, a function of the model is to find patterns in the time step dimension that effectively responds to the need for lags inclusion. We also evaluate forecasting quality using various measures; namely root mean square error (RMSE), mean absolute error (MAE), and Diebold-Mariano (DM) test.

The main empirical finding of this paper indicates that news sentiment generates the paramount results for oil price volatility forecasting relative to the obtained forecast using Twitter sentiment. We also compare the proposed framework with well-established benchmark models fed with both statistical and sentiment data. Adding the Twitter sentiment index to the forecasting setting reduces the error (MAE) by 27%, while replacing that with the news sentiment index reduces the error by 23% further. This error reduction indicates that regular media provides more refined information for the oil market. This is somewhat inconsistent with the idea that the strength of evidence receives more attention than the weight of evidence (Griffin and Tversky, 1992). In this context, we consider the strength as the frequency or reportage of the information and the weight as the source of information. The reportage is higher on social media, but press is more credible. This finding shows that at a market level, and in aggregate, this is the weight that plays the more important role concerning the oil price volatility. Overall, using exclusive semantic sentiment indices enhances forecasting accuracy compared to the results using historical data alone.

From these findings, we conclude that social media such as Twitter is becoming important as an alternative platform for information circulation about the oil market. However, the traditional press and news agencies are more important than social media with regard to spreading more impactful information. This finding suggests that market watchers should place more weight on regular news than Twitter feeds for information updates. This finding also implies that since people can freely express their opinions and anticipations regarding oil market on social media; therefore, Twitter sentiment contains inauthentic interpretations as to news and events, which reduces the degree of sentiment reliability.

This paper adds three important contributions to the existing literature. Firstly, we document that investigating Twitter feeds significantly improves oil price volatility predictability, and this can be traced to incremental information that Twitter offers. To the best of our knowledge, this is the first study that develops a semantic Twitter sentiment index exclusively for the oil market. Previous studies use emotional sentiment proxy (Lehrer et al., 2021) or Twitter general economic uncertainty index (Lang et al., 2021), while those indices do not include non-economic factors or are just a general index. Secondly, we also develop a semantic sentiment index for news headlines relevant to the oil market and compare the predictive power of regular press with social media in forecasting oil price volatility for the first time in the literature. Using these new indices, we document the superiority of news sentiment over Twitter sentiment. Although both indices bring incremental information to the forecasting setting, our results indicate the quality of news information is higher. Thirdly, not only do we produce sentiment indices based on the meaning of the text which reveals the significance of information more refinedly (Shiller, 2020), but also fine-tune the BERT on financial language using Financial PhraseBank proposed by Malo et al. (2014). This practice helps capture the impact of textual information on market prices. We also prove the superiority of the proposed framework over various benchmark models, including individual constituent models and the framework proposed by Verma (2021). A good understanding of various factors impacting the prediction of oil price volatility is vital for policymakers, central banks, and investors across financial markets.

The remainder of the paper is organized as follows. Section 2 presents a review of the pertinent literature and hypothesis development. Section 3 describes the methodologies used for this study. In Section 4, we present the data. Section 5 reports empirical results, while Section 6 concludes the paper.

2. Literature review and hypothesis development

Studying oil price forecasting begins with Amano (1987) by proposing a small-scale econometric model for global oil market. Later, Sharma (1998) compares the quality of GARCH class of models for oil price volatility prediction using data from 1986 to 1997. Tang and Hammoudeh (2002) predict OPEC oil prices using a regression model and find that the prices are also affected by the expectations of market participants. The early studies mainly employ statistical models which are generally able to capture linearity and time-varying volatility in a time series (Bhar and Hamori, 2005; Yu et al., 2008). Another strand of the oil price forecasting literature encompasses machine learning agents mainly due to their ability to capture nonlinear patterns embedded in the time series. Yu et al. (2008) use a neural network based on an empirical mode decomposition to forecast crude oil spot prices. He et al. (2012) propose a wavelet decomposed ensemble model to enhance the prediction accuracy of oil prices with a closer look at the market microstructure, where the proposed model follows the heterogeneous market hypothesis. They document the superiority of their proposed framework over benchmark models in the literature. Mostafa and El-Masry (2016) use an artificial neural network coupled with an evolutionary algorithm to forecast oil prices and recommend using computational approaches like neural networks or fuzzy settings instead of traditional statistical models. On the one hand, statistical models are weak in capturing nonlinearity and non-stationarity, and this weakness decreases their forecasting quality. On the other, Makridakis et al. (2018) document that machine learning agents underperform statistical models in some cases. This fact elucidates the necessity of a third framework in the forecasting literature based on hybrid models consisting of both statistical and machine learning models (Fazelabdolabadi, 2019). The reason is that the structure of hybrid models makes it possible to capture different characteristics in the time series.

Numerous studies on hybrid models document the superiority of such models over single models (see Abdollahi, 2020; Prado et al., 2020). Yet, scholars seek new features or data to increase the prediction quality. This practice gives rise to data construction in forecasting process. Shiller (2020) argues that narratives and news stories can affect prices. This idea suggests that media provide incremental information which are useful for forecasting models. In recent years, scholars have attempted to cover this aspect. Wu et al. (2021) develop a text classification of online news headlines during the COVID-19 pandemic, which produces a binary output for fluctuation of each month, and combine the news proxy with statistical and machine learning models to increase the accuracy of oil price prediction. Li et al. (2021) use news on the oil market and find that shocks in news sentiment result in volatility across the future prices of oil. Liu et al. (2022) combine the Google search volume index as a sentiment proxy with a high-frequency heterogeneous autoregressive model to elevate the forecasting quality in the volatility of oil futures prices. Zhe et al. (2022) scrape the comments posted on Eastmoney forum, an online financial forum, and document the predictive power of sentiment as to the price of China's crude oil. Herrera et al. (2022) employ a lexicon-based method for sentiment extraction from Twitter feeds and combine that with a deep learning model to forecast the returns and volatility of selected renewable energy stocks. Using sentiment as an informational proxy is not limited to volatility forecasting, Zhao et al. (2023) use Google trends (Google search volume index) to forecast crude oil inventory. However, a notable fact about previous studies is the use of less-sophisticated methods based on counting words or pre-weighted lexicons for textual analysis. Those methods fail to provide a sentiment analysis based on the meaning of text. However, using semantic sentiment is recommended for investigating narratives across financial markets as it can show the psychological significance of sentiment (Shiller, 2017). This paper also attempts to bridge this gap in the literature by producing semantic sentiment for news and Twitter feeds.

de Medeiros et al. (2022) argue that data construction and curation can play a crucial role in improving prediction accuracy for the oil market dynamics. On the other hand, the literature suggests various factors influencing oil price volatility such as geopolitical escalations (Tahmassebi, 1986), macroeconomic news (Meng and Liu, 2019), shifts in supply and demand (Hosseini et al., 2021), etc. de Oliveira et al. (2018) argue that press and social media are a source for sentiment circulation. Lehkonen et al. (2022) note that media can be a genuine risk factor in financial markets. Abdollahi et al. (2023) document a long-lasting connectedness between media sentiment and financial market volatility. Therefore, based on this literature, we expect that constructing sentiment data from news and social media improves the forecasting accuracy of oil prices. We formalize this as the first hypothesis.

H1. Media sentiment increases the accuracy of oil price volatility forecasting.

Junttila et al. (2005) use analysts' perceptions published in a well-known magazine and show that such contents can be a value driver in the market. Birz and Lott Jr (2011) select newspaper articles that provide an interpretation of statistical releases as a measure of news and note that public interpretations of news affect stock returns. Kim et al. (2019) note that >60% of adults get news from social media by 2016 and the proportion is still increasing; even though the problem of fake news is a case. Shiller (2020) proposes that viral news stories, whether fact-based or not, can potentially influence prices across financial markets. Based on these studies, we expect that materials posted on Twitter, as an amalgamation of news and the concomitant public interpretations, can be a better proxy than regular news media for the oil price volatility forecasting. We formalize this in the second hypothesis.

H2. Twitter sentiment outperforms news sentiment in oil price volatility forecasting.

3. Methodology

3.1. BERT

BERT is a pre-trained natural language processing technique developed by Google (Devlin et al., 2018). The model is trained on English Wikipedia and BookCorpus (Zhu et al., 2015), containing >3500 million words. The principal feature of the BERT is the ability to understand the meaning of sophisticated language. We employ the BERT-base version to analyze text data by predicting the likelihood of positivity, neutrality, or negativity for a given input. We obtain the sentiment score for a given text as the difference between positive and negative likelihoods or logits (Lin and Luo, 2020):

Table 1
The proportion of different classes in Financial PhraseBank.

Class	Proportion
Positive	28.2%
Neutral	59.4%
Negative	12.4%

$$Sentiment\ Score = logit_{positive} - logit_{negative}. \tag{1}$$

The sentiment score for a given text is between 1 (extreme positive) and − 1 (extreme negative). We fine-tune the BERT for finance domain to further improve the model's ability to generate quality results. For this purpose, we use the Financial PhraseBank proposed by Malo et al. (2014). The Financial PhraseBank contains 4845 sentences from financial news classified by finance professionals. The classes show how the professionals perceive the impact of information embedded in a sentence on prices in the market (positive impact, neutral impact, or negative impact). Table 1 presents the proportion of different classes in the dataset. The PhraseBank also categorizes the sentences based on agreement levels. We use the entire PhraseBank including all agreement levels for training the BERT.[2] Fine-tuning the model is advantageous as it elevates the quality of sentiment extraction.

3.2. GARCH (1,1)

Engle (1982) proposes autoregressive conditional heteroscedasticity (ARCH) to model the conditional variance using error terms and to capture volatility clustering of a series as follows.

$$\epsilon_t = \sigma_t X_t, \sim N\left(0, \sigma_t^2\right) \tag{2}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \ldots + \alpha_q \epsilon_{t-q}^2 \tag{3}$$

Where ϵ_t denotes the error term which follows a normal distribution, and σ_t^2 is the conditional variance dependent on past squared residuals. Bollerslev (1986) extends the ARCH(q) model to build GARCH (p,q) model such that it contains both the past squared residuals and the values of conditional variances. The model is formulated as:

$$\sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i \gamma_{k-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{k-j}^2. \tag{4}$$

GARCH estimates the conditional variance (σ_t^2) of series γ_t. The model parameters are conditioned on $\omega > 0$, $\alpha_i \geq 0$, $\beta_i \geq 0$, and $\alpha_i + \beta_i < 1$, which are the essential conditions for stationarity in the series. We employ a GARCH (1,1) model to estimate the conditional variance of the oil prices and use it as an input for the next model in the proposed structure.

3.3. BiLSTM

Long short-term memory (LSTM) is an artificial neural network with the capability to learn sequential patterns. This model was developed to tackle the problem of long-range dependency in backpropagation (Hochreiter and Schmidhuber, 1997). The standard LSTM follows Eqs. (5) to (10).

[2] There are different agreement levels regarding the impact of sentences among professionals.

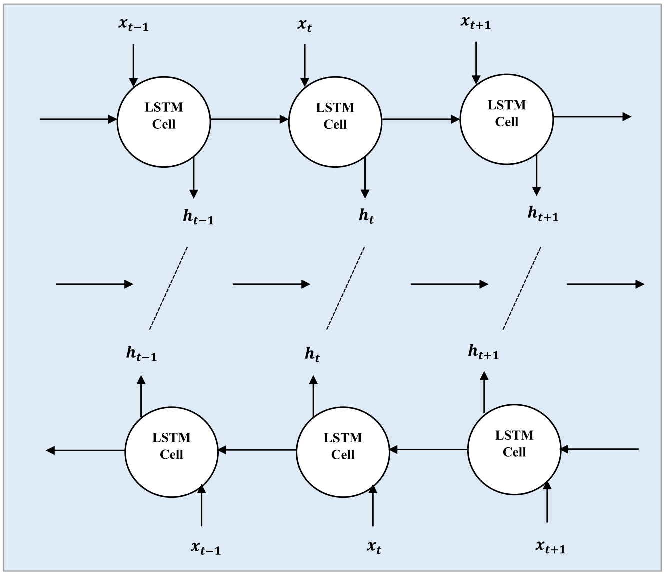


Fig. 1. The structure of BiLSTM.
Graphical representation of the BiLSTM model consisting of two LSTMs in which one takes inputs in a forward direction and the other in a backward direction leading to an increased amount of information absorbed by the model.

$$g_t = \sigma\left(b_g + U_g x_t + V_g h_{t-1}\right), \tag{5}$$

$$i_t = \sigma(b_i + U_i x_t + V_i h_{t-1}), \tag{6}$$

$$o_t = \sigma(b_o + U_o x_t + V_o h_{t-1}), \tag{7}$$

$$f_t = tanh\left(b_f + U_f x_t + V_f h_{t-1}\right), \tag{8}$$

$$c_t = g_t \odot c_{t-1} + i_t \odot f_t, \tag{9}$$

$$h_t = o_t \odot tanh(c_t). \tag{10}$$

The LSTM structure is composed of blocks consisting of one memory cell (c_t) and three gates, namely the input gate (i_t), forget gate (g_t), and output gate (o_t). In Eqs. (5)–(10), h_t and x_t denote the hidden state of the memory cell and input at time t, respectively. σ represents the activation function (sigmoid) and b is the bias term. U_i, U_g, and U_o denote the weight matrices of input, forget, and output gates, respectively. V_i, V_g, and V_o denote the recurrent weight matrices of input, forget, and output gates, respectively. f_t is an input modulate gate, which is a value showing the amount of new information received in the memory cell. Symbol $\odot$ represents element-wise multiplication. In the LSTM structure, forget gate g_t, as shown in Eq. (5), generates a value between 0 and 1, where 0 means that no input information passes through the gate and a value of 1 implies all input information is passed. The input gate, as shown in Eq. (6), ascertains the amount of information stored in the memory cell (c_t). Eq. (8) determines the new information at time t. Eq. (9) calculates past information and new information of the memory cell, which are controlled by input and forget gates, at time t. Finally, the hidden state information h_t is determined using the output gate (o_t).

BiLSTM algorithm is an update on LSTM, incorporating the bidirectional recurrent network structure into LSTM cells. This structure adds the ability to take advantage of feedback layers to the BiLSTM model. The model structure includes hidden layers that run in opposite directions. Therefore, the model acquires both forward and backward sequential information through the past and future of a given dataset. Fig. 1 illustrates the structure of the BiLSTM model.

3.4. Performance evaluation

The forecasts of the model must be further assessed to verify the findings. Also, a formal test is needed to examine the hypotheses. Hence, we employ the following measures and test:

I. RMSE: Shows the standard deviations of differences between actual and forecasted values. It is used as an accuracy indicator for comparing prediction errors of different models.

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(Y_t - \widehat{Y}_t\right)^2} \tag{11}$$

II. MAE: Shows the average of absolute errors when comparing the actual and forecasted values.

$$MAE = \frac{1}{T}\sum_{t=1}^{T}\left|Y_t - \widehat{Y}_t\right| \tag{12}$$

In Eq. (11) and Eq. (12), Y_t and $\widehat{Y}_t$ are the actual and forecasted values at time t, respectively. T represents the sample size. The smaller values for RMSE and MAE indicate higher forecasting accuracy.

III. DM test: Determines if the difference between two forecasting models is statistically significant (Diebold and Inoue, 2001). The null hypothesis implies no significant difference between errors. The DM statistic is defined as follows:

$$DM = \frac{\overline{d}}{\left(\frac{2\pi \widehat{f}_d}{T}\right)^{½}} \sim N(0,1), \tag{13}$$

where T is the size of sample. $\overline{d}$ denotes the mean of loss differential between two forecasts, and $\widehat{f}_d$ represents spectral density of the loss differential.

3.5. Benchmark forecasting models

Benchmark models are needed to check the effectiveness of the proposed forecasting model. For this aim, we use the constituent forecasting models, namely the GARCH and the BiLSTM, along with their hybridization in which the outcome of the GARCH model is used as one of the features to feed the BiLSTM model. The hybridization of GARCH- and LSTM-based models in forecasting oil price volatility has been reliably documented in the literature (Verma, 2021). Therefore, these models can be employed as reliable benchmarks. Another interesting hybridization is to use the outcome of the BiLSTM together with sentiment series as regressors in the GARCH model. This practice makes it possible to compare the suitability of statistical and machine learning models for oil price volatility predictability in the presence of sentiment. Therefore, we use the GARCH model, BiLSTM model, GARCH-BiLSTM model, and BERT-BiLSTM-GARCH model as benchmarks for our proposed model.

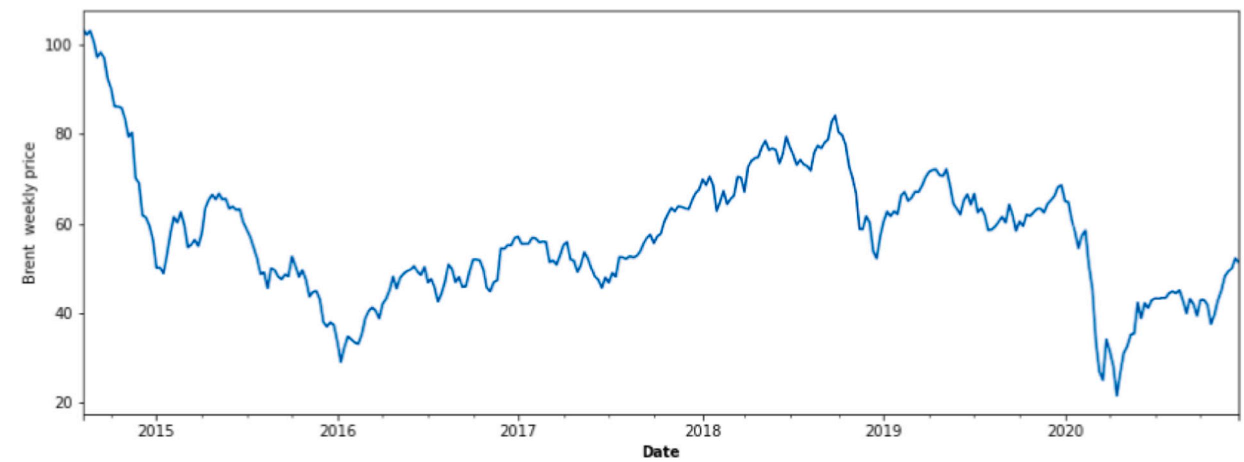
4. Data

4.1. Historical data

Weekly data of Brent oil were sourced from Refinitiv Eikon from 4 August 2014, to 20 December 2020. Brent oil historically plays a crucial role in the global oil pricing mechanism as it is considered a benchmark for around 55% of the international oil trade (Cheng et al., 2019). Therefore, Brent is an appropriate proxy for the global oil market. We also use weekly data for Gold and S&P 500 over the same period as financial indicators which have a high volatility connectedness with the oil market.

Fig. 2 shows oil price time series over the period under consideration. As can be observed, high production levels together with low demand pushed the oil prices to a nadir in early 2016, when it hit a 13-year low of

A) Brent closing prices



A) Brent price volatility

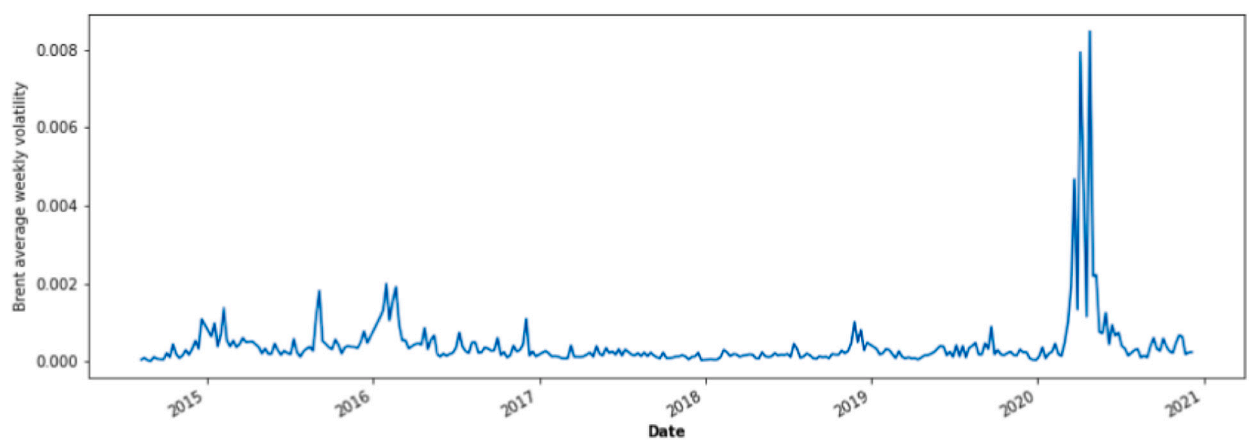


Fig. 2. Brent weekly price and volatility from August 2014 to December 2020.
Weekly closing prices and range-based volatility of Brent crude oil from August 2014 to December 2020. Panel A shows the historical trajectories in closing prices and Panel B shows the volatility for Brent crude oil, respectively.

Table 2
Descriptive statistics.

	Mean	Minimum	Maximum	SD	Skewness	ADF
Panel A. Financial series						
Brent	0.000422	0.000024	0.008468	0.000753	7.8625	−4.1483 (0.0008)
Gold	0.000054	0	0.0007	0.000077	4.5268	−7.3323 (0)
S&P 500	0.000066	0.000003	0.001784	0.00015	7.8192	−7.7669 (0)
Panel B. Sentiment series						
News sentiment index	0.179640	−0.997356	0.984978	0.463667	−0.3259	−10.0038 (0)
Twitter sentiment index	0.325211	−0.107918	0.749671	0.14302	0.2199	−14.5428 (0)

This table presents descriptive statistics for Brent crude oil prices, gold prices, and S&P 500 index together with news sentiment index and Twitter sentiment index for oil market over the period from August 2014 to December 2020 in panel A and B, respectively. Values in the first column shows weekly mean for each series followed by the minimum and maximum values in the next two columns. SD stands for standard deviations. Number in parenthesis in the last columns show the *p*-value at 5% significance level.

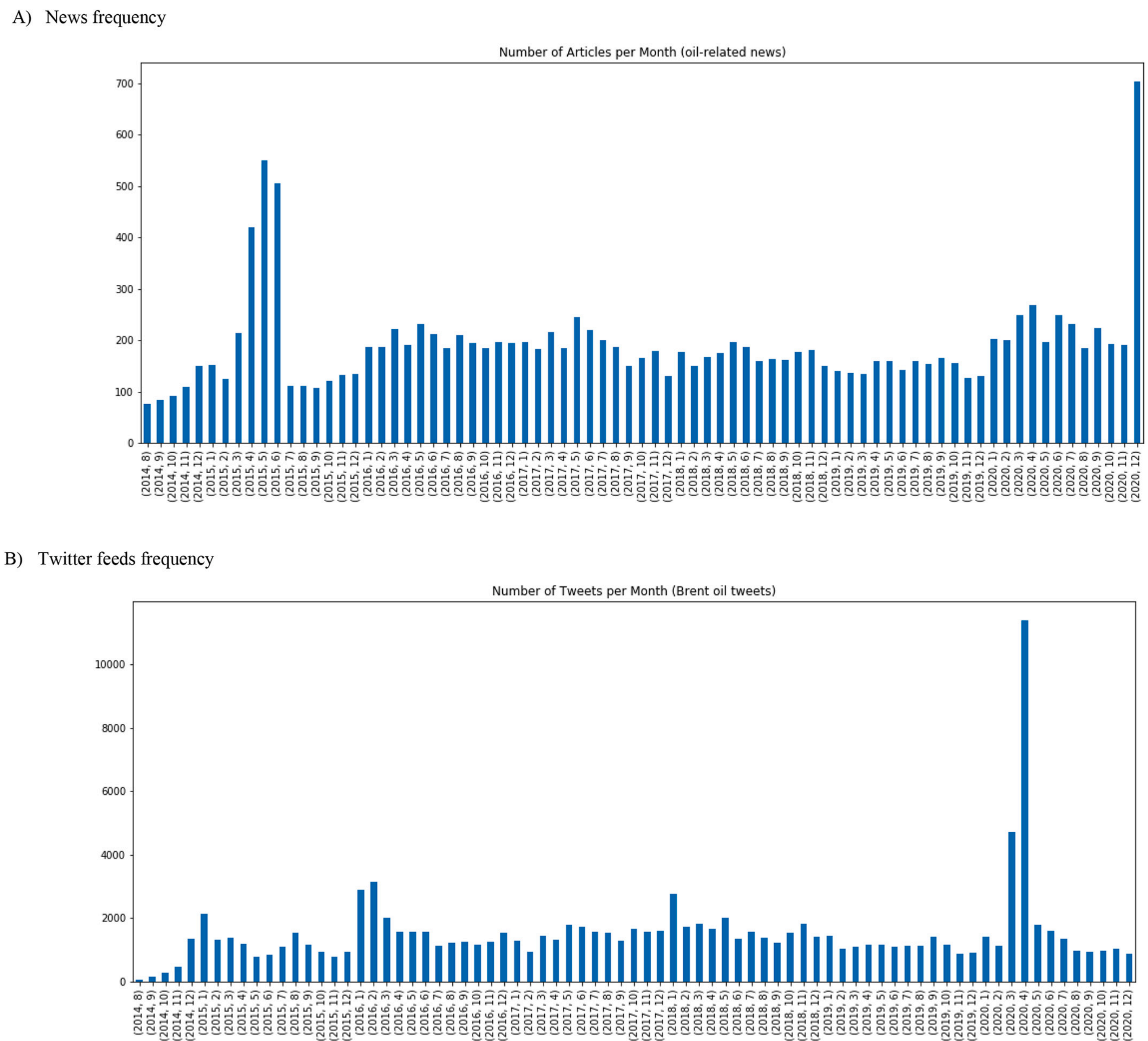


Fig. 3. Number of news and tweets relevant to the oil market from August 2014 to December 2020.
Panel A shows the frequency of online news about Brent crude oil across all news services on Investing.com from August 2014 to December 2020. Panel B sows the frequency of Twitter feeds using one of the hashtags: #oilmarket, #brentoil, #oilprice, #wtioilprice from August 2014 to December 2020.

$27.10. Later, the trend reversed with a simultaneous increase in global demand and a drop in output levels. Also, a series of geopolitical issues took place over 2017–18, like the imposition of the U.S. sanctions on Iran, or Russia and Saudi Arabia's decisions to curb production levels. All these events sent oil prices to a four-year high of more than $80 in late 2018. Throughout 2019, rise in the U.S. oil production put downward pressure on prices. Moreover, geopolitical issues such as the attack on Saudi Arabia's petroleum installations and production cut announcements by OPEC contributed to a lower average of oil prices. Finally, the fall in 2020 was related to a remarkable drop in demand coupled with weakened economic prospects during the Covid-19 pandemic even though it reversed later. The oil price is both volatile and trending at some points over this period in which media shocks play a role; therefore, using this period is beneficial to investigating the reliability of the proposed model.

Panel A in Table 2 presents descriptive statistics for the financial series. For all markets, the standard deviation values are greater than the average. This shows that Brent (0.000753) has the highest volatility level, while gold (0.000077) is of the lowest volatility. The skewness values indicate that all financial series are highly skewed. We check the stationarity of the series using the augmented Dickey-Fuller (ADF) test, where the null hypothesis implies the existence of a unit root in series. The last column of Table 2 presents the results of the ADF test, confirming the stationarity of all series at the 5% significance level.

4.2. Textual data

Using the archival news of *investing.com*[3] Website, a total of 14,833

[3] Investing.com is a financial markets platform providing exclusive and eclectic news and other facilities regarding multifarious exchanges and commodities around the world.

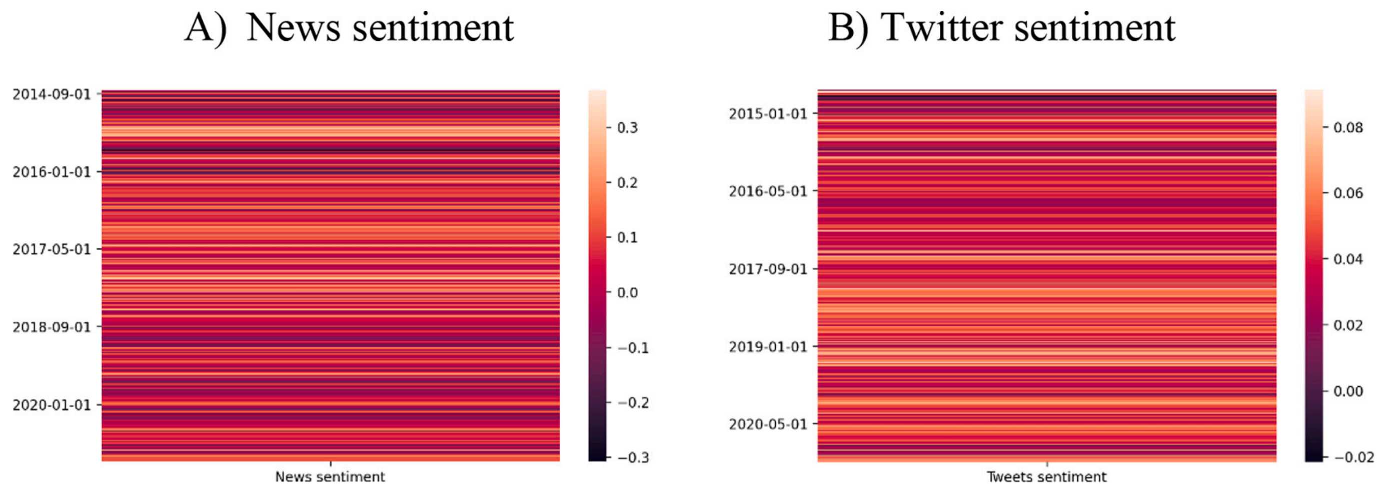


Fig. 4. Heatmaps of sentiment indices.
Heatmaps of the semantic indices through news and Twitter feeds exclusively for the oil market from August 2014 to December 2020. The indices rely on real-time events causing the sentiment to be varying. Time-varying characteristic is more pronounced in the case of news sentiment, while Twitter sentiment is a bit more positive.

news headlines about the oil market was collected from 4 August 2014, to 22 December 2020. We use news headlines as headlines efficiently convey the gist of the whole article (Li et al., 2019). Panel A in Fig. 3 shows the number of news articles per month. There is a sharp increase in the quantity of news in December 2020 which can relate to the OPEC meeting, OPEC and non-OPEC ministerial meeting, and announcement of Covid-19 vaccine production.

We also collected 116,463 tweets for different hashtags related to the oil market over the same period.[4] Panel B in Fig. 3 illustrates the number of tweets per month. As can be seen, there is a dramatic increase in March and April 2020 when the oil price unprecedently plunged and, simultaneously, there was a serious conflict between OPEC+ members where the U.S. president mediated between them and repeatedly updated the public via his Twitter account.

5. Empirical analysis

To check Hypothesis 1, which predicts media sentiment elevates the accuracy of oil price volatility forecasting, we commence the analysis by sentiment extraction using the BERT model trained on the Financial PhraseBank. We use 80% of the dataset for training and validation and 20% for test. This practice boosts the model accuracy to 0.83%. The weighted average of model's precision and recall equals 0.82, and model's loss is 0.38. We then process the textual contents through the fine-tuned model. Fig. 4 represents the heatmaps for sentiment indices. As can be seen, the sentiment indices are very dynamic and varying as they mostly depend on exogenous news like real-time events. Twitter sentiment also includes public interpretations, which seemingly add a tinge of pessimism to the sentiment as most of the index values fall within the positive area. Also, From Panel B in Table 2, we see that Twitter sentiment index is positively skewed, while the news sentiment index is negatively skewed. Standard deviations also indicate that the news sentiment index (0.463667) is more volatile than the Twitter sentiment index (0.14302).

We use the sentiment indices to feed the BiLSTM model in the last step of the proposed model. Before that, we estimate the GARCH model, whose parameters are presented in Table 3, to obtain an initial forecast of oil prices volatility. Not only do we assess this forecast against the actual volatility, but we also use it as an input feature for the final forecast.

Table 3
The estimated parameters of GARCH model.

	Estimate	SE	t-statistic	P-value
ω	13×10^{-9}	0	304.875	(0)
α	0.2	0.0922	2.169	(0.03006)
β	0.78	0.0331	23.568	(0)

Table presents the estimation of GARCH (1,1) parameters where ω is the constant coefficient of the variance equation, α represents the value of the autoregressive coefficient, and β denotes the value of variance coefficient. SE: standard error.

Finally, we feed the BiLSTM model with the sentiment indices, GARCH estimation, and financial data as input features. As for the sentiment data, we use the Twitter index and news index each at a time. The structure of the BiLSTM model consists of two layers, having 12 and 1 neurons, and the dense layer. We also use Adam optimizer (Kingma and Ba, 2014) to optimize the training process. The learning rate is set at 0.001. All BiLSTM settings in this study are trained for 60 epochs. The prediction horizon is 22 weeks, and we use 12-week window of explanatory variables as input.

Fig. 5 illustrates the final prediction for Brent oil volatility by the proposed hybrid model using the news sentiment index (News-GARCH-BiLSTM) in Panel A and using the Twitter sentiment index (Twitter-GARCH-BiLSTM) in Panel B. Fig. 5 also shows forecasting results obtained by benchmark models including News-BiLSTM-GARCH (Panel C), Twitter-BiLSTM-GARCH (Panel D), GARCH-BiLSTM using financial data (Panel E), BiLSTM using financial data (Panel F), and GARCH using oil historical data (Panel G).

Table 4 presents forecast errors for all the models employed in this study. We can see significant drops in forecast errors when the hybrid model is fed with sentiment indices. The interpretation is that refined information concerning the oil market in media increases the forecasting quality if properly converted into numerical values. The news sentiment index produces the paramount forecast as it reduces the RMSE measure derived by the Twitter sentiment index by 21%. The interpretation is that regular news is of more genuine information for the oil market.

Having the sentiment indices involved in the forecasting models, we can also see that using BiLSTM as the ultimate forecasting model improves the prediction errors by 15% in comparison to having the GARCH as the final predictive model. Another interesting fact goes back to the propriety of predictive models and data in prediction improvement. For the Twitter sentiment index, we can see that using the GARCH as the

[4] The hashtags are: #brentoil, #oilprice, #wtioilprice, #oilmarket.

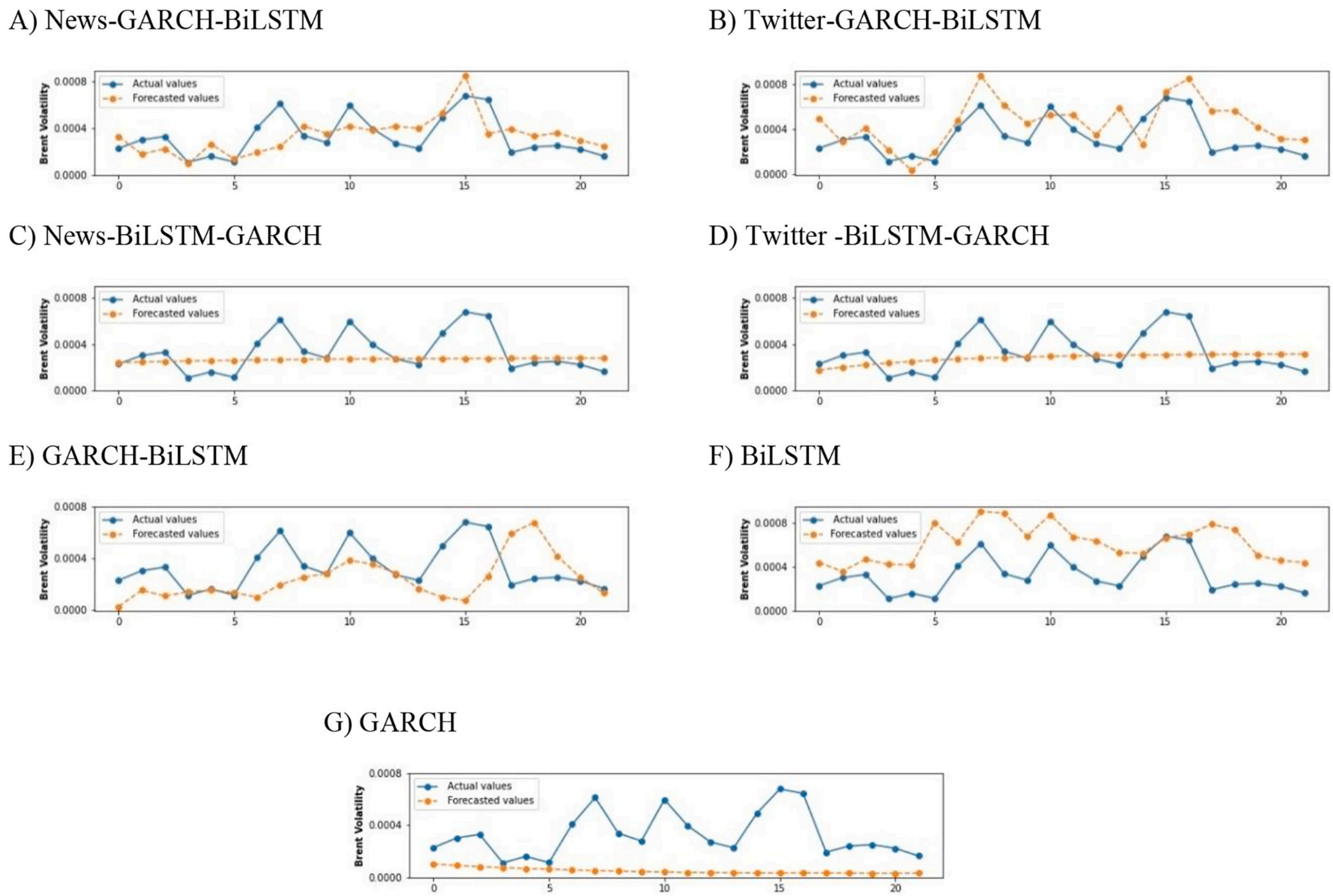


Fig. 5. Forecasting results of Brent oil volatility for 22 weeks ahead.
Forecasting 22-week ahead for Brent crude oil volatility. Panel A shows the predictions for News-GARCH-BiLSTM, Panel B illustrates the forecasting by Twitter-GARCH-BiLSTM, Panel C represents the predictions using News-BiLSTM-GARCH, panel D shows the forecasting using Twitter-BiLSTM-GARCH, panel E illustrate the prediction using GARCH-BiLSTM, panel F shows the forecasting using BiLSTM, and panel G illustrates the prediction using GARCH model.

Table 4
Results of oil price volatility forecasts with error functions.

Model	MAE ($\times 10^{-4}$)	RMSE ($\times 10^{-4}$)
GARCH	0.2819	0.3329
BiLSTM	0.2871	0.3365
GARCH-BiLSTM	0.2286	0.2948
Twitter-BiLSTM-GARCH	0.1393	0.1818
News-BiLSTM-GARCH	0.1331	0.1818
Twitter-GARCH-BiLSTM	0.1661	0.1951
News-GARCH-BiLSTM	0.1269	0.1531

Table shows the values of error measures for all the forecasting models using root mean square error (RMSE) and mean absolute error (MAE).

final forecasting model results in smaller prediction errors than the BiLSTM. The opposite is true with regard to the news sentiment index.

We can also see that the GARCH provides a slightly better forecast than that of the BiLSTM when we scrutinize the performance of single models. The interpretation is that the statistical models provide superior results than machine learning agents when we focus on using single models for forecasting oil price volatility under this setting. We also observe that the hybrid GARCH-BiLSTM model (using financial data) produces smaller forecast errors than single models. The interpretation is that the hybridization of statistical and machine learning models improves the forecasting accuracy, as the errors subsided. The reason is that this process enables the model to capture different characteristics existing in the time series more effectively.

Table 5
The results of DM test.

Pair of models	DM statistic	P-value
GARCH vs. BiLSTM	8.679	(0.0000)
GARCH-BiLSTM vs. GARCH	3.1692	(0.0046)
GARCH-BiLSTM vs. BiLSTM	−7.3704	(0.0000)
Twitter-BiLSTM-GARCH vs. GARCH	6.68	(0.0000)
News-BiLSTM-GARCH vs. GARCH	6.6767	(0.0000)
Twitter-BiLSTM-GARCH vs. BiLSTM	7.2509	(0.0000)
News-BiLSTM-GARCH vs. BiLSTM	7.2925	(0.0000)
Twitter-BiLSTM-GARCH vs. GARCH-BiLSTM	6.7	(0.0000)
News-BiLSTM-GARCH vs. GARCH-BiLSTM	6.6768	(0.0000)
Twitter-GARCH-BiLSTM vs. GARCH	5.5788	(0.0001)
News-GARCH-BiLSTM vs. GARCH	4.2697	(0.0003)
Twitter-GARCH-BiLSTM vs. BiLSTM	−3.8205	(0.0009)
News-GARCH-BiLSTM vs. BiLSTM	−5.1502	(0.0000)
Twitter-GARCH-BiLSTM vs. GARCH-BiLSTM	3.5132	(0.002)
News-GARCH-BiLSTM vs. GARCH-BiLSTM	10.1343	(0.0000)
Twitter-GARCH-BiLSTM vs. Twitter-BiLSTM-GARCH	3.9351	(0.0007)
News-GARCH-BiLSTM vs. Twitter-BiLSTM-GARCH	10.1344	(0.0000)
Twitter-GARCH-BiLSTM vs. News -BiLSTM-GARCH	2.132	(0.04497)
News-GARCH-BiLSTM vs. News -BiLSTM-GARCH	4.0629	(0.0005)
News-BiLSTM-GARCH vs. Twitter-BiLSTM-GARCH	2.6968	(0.0135)
News-GARCH-BiLSTM vs. Twitter-GARCH-BiLSTM	10.1337	(0.00)

Table presents the result of pairwise comparisons between forecasting errors using Diebold-Mariano test. The null hypothesis of the test implies no significant difference between the predictions. Significance level is 5%.

Table 6
Results of oil price volatility forecasts for different prediction horizons.

Model	Prediction horizon = 16		Prediction horizon = 28	
	MAE ($\times 10^{-4}$)	RMSE ($\times 10^{-4}$)	MAE ($\times 10^{-4}$)	RMSE ($\times 10^{-4}$)
GARCH	0.3396	0.3793	0.3051	0.3627
BiLSTM	0.296	0.3446	0.2567	0.3042
GARCH-BiLSTM	0.2712	0.3558	0.784	1.1202
Twitter-BiLSTM-GARCH	0.1715	0.2063	0.1886	0.2394
News-BiLSTM-GARCH	0.1554	0.2052	0.1875	0.2349
Twitter-GARCH-BiLSTM	0.1698	0.2154	0.1899	0.2353
News-GARCH-BiLSTM	0.1326	0.1826	0.1868	0.2058

Table presents the robustness check for the predictions of Brent crude oil price volatility using different forecast horizons.

We perform the DM test as a formal test to check the hypotheses. This test provides pairwise comparisons between the forecasts generated by different models. The DM tests the assumption of no significant difference between two forecasts. Table 5 presents the results of the DM test. From rows 4–15 in Table 5, we see that there are statistically significant differences between the forecasts of hybrid models fed with sentiment indices and those of models without sentiment inclusion. The interpretation is that media sentiment provides novel and recent information regarding the oil market that boosts the quality of forecasting. A notable feature of sentiment is that it frequently varies with respect to the dynamic atmosphere of the real world. Given the rapidity of information incorporation within financial markets, take the effects of the Covid-19 pandemic, the US-China Tariff clash, and Brexit for instance, building an index that covers the behavioral aspects relevant to the oil market adds a contribution to the predictive power of models. Therefore, the results are consistent with hypothesis 1 that media sentiment elevates the accuracy of oil price volatility predictability. We conclude that media sentiment enhances the forecasting quality of oil price volatility because it includes new information which is absent in historical data.

Hypothesis 2 states that Twitter sentiment outperforms news sentiment in oil price volatility forecasting. From Table 4, we find that the opposite is true as forecasts using the news sentiment index generate smaller prediction errors than those of the Twitter sentiment index. To check if this finding is statistically significant, we perform the DM test for these forecasts. The last six rows in Table 5 show statistically significant differences between forecasts using the news sentiment index and the Twitter sentiment index under both forecasting models. Therefore, hypothesis 2 is rejected as news stories provide more quality prediction for oil price volatility than Twitter feeds. The interpretation is that although quantified information about the exogenous factors of the oil market increases the forecasting quality, the results are still sensitive to the quality and reliability of the information. In this analysis, news sentiment is built upon globally prestigious press, where those news agencies apply a high level of ethical and professional observations. Therefore, we are ensured that their contents usually are fact-checked and reliable. On the contrary, Twitter feeds are fraught with heterogeneous users causing the index contains dual information as one part of the index includes reliable information, while the other part is based on non-fact contents in the form of public interpretations or even false news. We conclude that news sentiment outperforms Twitter sentiment in forecasting oil price volatility, as the results indicate that oil market fluctuations are more sensitive to genuine news.

We further examine the robustness of our results using two different forecasting horizons of 16 and 28 weeks ahead. Table 6 presents the corresponding forecasting errors. From the second column in Table 6, we see that sentiment-based models produce smaller errors than those of other models over a shorter forecasting horizon. We also observe that the news sentiment index generates a better forecast than the Twitter sentiment index. From the third column in Table 6, we see that the findings remain unchanged over a longer forecasting horizon as well. These facts signify that changes in the forecasting horizon lead to no significant variations in the results. Therefore, we conclude that the obtained results are robust and contributory to the empirical literature.

6. Conclusion

In this paper, we investigate the contribution of news stories and Twitter feeds to the quality of predictive models for oil price volatility. Oil market fluctuations have always been influenced by exogenous factors such as geopolitical conflicts, natural disasters, viral narratives, etc. These factors are spread in media. Research is at an early stage as to incorporating media sentiment into forecasting settings. This is the first study that builds semantic sentiment indices for both news and Twitter feeds and compares their effectiveness.

We collect a total of 14,833 news headlines and 116,463 tweets regarding various aspects of the oil market from August 2014 to December 2020. We then process the textual contents using a financially fine-tuned BERT model to extract weekly sentiment indices for both news and Twitter feeds. We add the sentiment indices into a GARCH-BiLSTM model to check the impact of media sentiment. Findings indicate that media sentiment enhances forecasting accuracy relative to using only historical data. More importantly, we also find that news sentiment generates superior results for oil price volatility forecasting compared with Twitter-based forecasting. This improvement results from the inclusion of various features to cover certain aspects of oil price characteristics, specifically the sentiment.

Empirical implications suggest that policymakers, central banks, hedgers, and others who deal with the oil market should consider the crucial role of media sentiment in their anticipations. More weight should be put on regular press than social media as the sentiment derived from news agencies further reduces forecasting errors.

Results also present theoretical insights. We document that news and Twitter sentiments contribute to forecasting quality. One way to extend this line of research is to add a weighting function, which assigns optimized weight to each forecasting feature, to the current setting and use both indices simultaneously. As the paramount result was obtained by news sentiment, future models can also divide the news index into different categories, like economic news, political news, war news, etc., and compare the predictive power of each category.

Credit author statement

This is a single-authored paper (the author confirms sole responsibility for idea and design, data collection, analysis and interpretation of results, and manuscript preparation).

Acknowledgements

Special thanks are due to Espen Sirnes and Sturla L. Fjesme for their advice during the conduct of this study. Thanks also to Babak Ebrahimi for useful comments

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eneco.2023.106711.

References

Abdollahi, H., 2020. A novel hybrid model for forecasting crude oil price based on time series decomposition. Appl. Energy 267, 115035.
Abdollahi, H., Ebrahimi, S.B., 2020. A new hybrid model for forecasting Brent crude oil price. Energy 200, 117520.

Abdollahi, H., Fjesme, S.L., Sirnes, E., 2023. Measuring Connectedness between Media Sentiment and Market Volatility (Working paper).
Adams, Z., Collot, S., Kartsakli, M., 2020. Have commodities become a financial asset? Evidence from ten years of Financialization. Energy Econ. 89, 104769.
Amano, A., 1987. A small forecasting model of the world oil market. J. Policy Model 9 (4), 615–635.
An, S., Gao, X., An, H., An, F., Sun, Q., Liu, S., 2020. Windowed volatility spillover effects among crude oil prices. Energy 200, 117521.
Bhar, R., Hamori, S., 2005. Empirical Techniques in Finance. Springer Science & Business Media.
Birz, G., Dutta, S., 2016. Us macroeconomic news and international stock prices: evidence from newspaper coverage. Account. Fin. Res. 5 (1), 247.
Birz, G., Lott Jr., J.R., 2011. The effect of macroeconomic news on stock returns: new evidence from newspaper coverage. J. Bank. Financ. 35 (11), 2791–2800.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econ. 31 (3), 307–327.
Bomfim, A.N., 2003. Pre-announcement effects, news effects, and volatility: monetary policy and the stock market. J. Bank. Financ. 27 (1), 133–151.
Brenner, M., Pasquariello, P., Subrahmanyam, M., 2009. On the volatility and comovement of US financial markets around macroeconomic news announcements. J. Financ. Quant. Anal. 44 (6), 1265–1289.
Chen, H., Liao, H., Tang, B.J., Wei, Y.M., 2016. Impacts of OPEC's political risk on the international crude oil prices: An empirical analysis based on the SVAR models. Energy Econ. 57, 42–49.
Cheng, F., Li, T., Wei, Y.M., Fan, T., 2019. The VEC-NAR model for short-term forecasting of oil prices. Energy Econ. 78, 656–667.
Crawford, G.W., Fratantoni, M.C., 2003. Assessing the forecasting performance of regime-switching, ARIMA and GARCH models of house prices. Real Estate Econ. 31 (2), 223–243.
Cui, J., Goh, M., Zou, H., 2021. Coherence, extreme risk spillovers, and dynamic linkages between oil and China's commodity futures markets. Energy 225, 120190.
de Medeiros, R.K., da Nóbrega Besarria, C., de Jesus, D.P., de Albuquerquemello, V.P., 2022. Forecasting oil prices: new approaches. Energy 238, 121968.
de Oliveira, F.A., Maia, S.F., de Jesus, Besarria, C.D.N., 2018. Which information matters to market risk spreading in Brazil? Volatility transmission modelling using MGARCH-BEKK, DCC, t-Copulas. The North American Journal of Economics and Finance 45, 83–100.
Devlin, J., Chang, M.W., Lee, K., Toutanova, K., 2018. Bert: Pre-Training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint. *arXiv:1810.04805*.
Diebold, F.X., Inoue, A., 2001. Long memory and regime switching. J. Econ. 105 (1), 131–159.
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econ. J. Econ. soc. 987–1007.
Fazelabdolabadi, B., 2019. A hybrid Bayesian-network proposition for forecasting the crude oil price. Fin. Innovat. 5 (1), 1–21.
González-Carvajal, S., Garrido-Merchán, E.C., 2020. Comparing BERT against Traditional Machine Learning Text Classification. *arXiv preprint*. *arXiv:2005.13012*.
Griffin, D., Tversky, A., 1992. The weighing of evidence and the determinants of confidence. Cogn. Psychol. 24 (3), 411–435.
He, K., Yu, L., Lai, K.K., 2012. Crude oil price analysis and forecasting using wavelet decomposed ensemble model. Energy 46 (1), 564–574.
Herrera, G.P., Constantino, M., Su, J.J., Naranpanawa, A., 2022. Renewable energy stocks forecast using twitter investor sentiment and deep learning. Energy Econ. 114, 106285.
Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9 (8), 1735–1780.
Hosseini, S.H., Shakouri, H., Kazemi, A., 2021. Oil price future regarding unconventional oil production and its near-term deployment: a system dynamics approach. Energy 222, 119878.
Hung, N.T., 2022. Asymmetric connectedness among S&P 500, crude oil, gold and bitcoin. Manag. Financ. 48 (4), 587–610.
Ji, Q., Bouri, E., Roubaud, D., 2018. Dynamic network of implied volatility transmission among US equities, strategic commodities, and BRICS equities. Int. Rev. Financ. Anal. 57, 1–12.
Junttila, J., Kallunki, J.P., Kärja, A., Martikainen, M., 2005. Stock market response to analysts' perceptions and earnings in a technology-intensive environment. Int. Rev. Financ. Anal. 14 (1), 77–92.
Kim, A., Moravec, P.L., Dennis, A.R., 2019. Combating fake news on social media with source ratings: the effects of user and expert reputation ratings. J. Manag. Inf. Syst. 36 (3), 931–968.
Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization. *arXiv preprint*. *arXiv:1412.6980*.
Lang, Q., Lu, X., Ma, F., Huang, D., 2021. Oil futures volatility predictability: evidence based on twitter-based uncertainty. Financ. Res. Lett. 102536.
Lehkonen, H., Heimonen, K., Pukthuanthong, K., 2022. Is media tone just a tone? Time-series and cross-sectional evidence from the currency market. In: Time-Series and Cross-Sectional Evidence from the Currency Market (November 22, 2022).
Lehrer, S., Xie, T., Zhang, X., 2021. Social media sentiment, model uncertainty, and volatility forecasting. Econ. Model. 102, 105556.
Li, Y., Jiang, S., Li, X., Wang, S., 2021. The role of news sentiment in oil futures returns and volatility forecasting: data-decomposition based deep learning approach. Energy Econ. 95, 105140.
Li, X., Shang, W., Wang, S., 2019. Text-based crude oil price forecasting: A deep learning approach. Int. J. Forecast. 35 (4), 1548–1560.
Lin, P., Luo, X., 2020, October. A survey of sentiment analysis based on machine learning. In: CCF International Conference on Natural Language Processing and Chinese Computing. Springer, Cham, pp. 372–387.
Liu, M., Lee, C.C., 2021. Capturing the dynamics of the China crude oil futures: Markov switching, co-movement, and volatility forecasting. Energy Econ. 103, 105622.
Liu, Y., Niu, Z., Suleman, M.T., Yin, L., Zhang, H., 2022. Forecasting the volatility of crude oil futures: the role of oil investor attention and its regime switching characteristics under a high-frequency framework. Energy 238, 121779.
Lucca, D.O., Moench, E., 2015. The pre-FOMC announcement drift. J. Financ. 70 (1), 329–371.
Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2018. Statistical and machine learning forecasting methods: concerns and ways forward. PLoS One 13 (3), e0194889.
Malo, P., Sinha, A., Korhonen, P., Wallenius, J., Takala, P., 2014. Good debt or bad debt: detecting semantic orientations in economic texts. J. Assoc. Inf. Sci. Technol. 65 (4), 782–796.
Meng, F., Liu, L., 2019. Analyzing the economic sources of oil price volatility: An out-of-sample perspective. Energy 177, 476–486.
Möbert, J., 2009. Dispersion in Beliefs among Speculators as a Determinant of Crude Oil Prices. Deutsche Bank Research, Research Notes, p. 32.
Mostafa, M.M., El-Masry, A.A., 2016. Oil price forecasting using gene expression programming and artificial neural networks. Econ. Model. 54, 40–53.
Prado, F., Minutolo, M.C., Kristjanpoller, W., 2020. Forecasting based on an ensemble autoregressive moving average-adaptive neuro-fuzzy inference system–neural network-genetic algorithm framework. Energy 197, 117159.
Qadan, M., Nama, H., 2018. Investor sentiment and the price of oil. Energy Econ. 69, 42–58.
Ross, S.A., 1989. Information and volatility: the no-arbitrage martingale approach to timing and resolution irrelevancy. J. Financ. 44 (1), 1–17.
Safari, A., Davallou, M., 2018. Oil price forecasting using a hybrid model. Energy 148, 49–58.
Sharma, N., 1998. Forecasting Oil Price Volatility (Doctoral dissertation, Virginia Tech).
Shiller, R.J., 2017. Narrative economics. Am. Econ. Rev. 107 (4), 967–1004.
Shiller, R.J., 2020. Narrative Economics: How Stories Go Viral and Drive Major Economic Events. Princeton University Press.
Tahmassebi, H., 1986. The impact of the Iran-Iraq war on the world oil market. Energy 11 (4–5), 409–411.
Tang, L., Hammoudeh, S., 2002. An empirical exploration of the world oil price under the target zone model. Energy Econ. 24 (6), 577–596.
Verma, S., 2021. Forecasting volatility of crude oil futures using a GARCH–RNN hybrid approach. Intell. Syst. Account. Fin. Manag. 28 (2), 130–142.
Wu, B., Wang, L., Wang, S., Zeng, Y.R., 2021. Forecasting the US oil markets based on social media information during the COVID-19 pandemic. Energy 226, 120403.
Yu, L., Wang, S., Lai, K.K., 2008. Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Econ. 30 (5), 2623–2635.
Zhang, D., Ji, Q., Kutan, A.M., 2019. Dynamic transmission mechanisms in global crude oil prices: estimation and implications. Energy 175, 1181–1193.
Zhao, L.T., Zheng, Z.Y., Wei, Y.M., 2023. Forecasting oil inventory changes with Google trends: a hybrid wavelet decomposer and ARDL-SVR ensemble model. Energy Econ. 106603.
Zhe, J., Lin, Z., Lingling, Z., Bo, W., 2022. Investor sentiment and machine learning: predicting the price of China's crude oil futures market. Energy 123471.
Zhu, Y., Kiros, R., Zemel, R., Salakhutdinov, R., Urtasun, R., Torralba, A., Fidler, S., 2015. Aligning books and movies: towards story-like visual explanations by watching movies and reading books. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 19–27.

Paper III

Fake News and Market Volatility: Insights from a Large Language Model