PressAcademia
www.pressacademia.org/journals/jbef
JBEF
ISSN 2146-7943
Journal of Business, Economics and Finance YEAR 2026 VOLUME 15 ISSUE 1

THE INFLUENCE OF GENERATIVE ARTIFICIAL INTELLIGENCE ON FINANCIAL MARKET VOLATILITY, LIQUIDITY, AND PREDICTABILITY

DOI: 10.17261/Pressacademia.2026.2029
JBEF- V.15-ISS.1-2026(1)-p.1-15

Nada Dammak
University of Sfax, Faculty of Economics and Management of Sfax (FSEG), Tunisia.
nada_dammak@yahoo.fr , ORCID: 0009-0006-0737-9182

Date Received: December 29, 2025 **Date Accepted:** April 4, 2026

To cite this document
Dammak, N., (2026). The influence of generative artificial intelligence on financial market volatility, liquidity, and predictability. Journal of Business, Economics and Finance (JBEF), 15(1), 1-15.
Permanent link to this document: http://doi.org/10.17261/Pressacademia.2026.2029

ABSTRACT

Purpose- This study examines the impact of generative artificial intelligence (AI) on key dimensions of financial market dynamics, namely volatility, liquidity, and return predictability. It emphasizes the dual role of AI as both an enhancer of informational efficiency and a potential source of short-term market instability.

Methodology- The paper proposes an original AI-based sentiment indicator constructed using a hybrid large language model (LLM) framework that combines FinBERT and a GPT-4–class generative model. This sentiment index is analyzed alongside daily returns of major equity indices—S&P 500, NASDAQ, and STOXX 600—over the period 2018–2024. The empirical analysis relies on Principal Component Analysis (PCA), GARCH (1,1)-X models, and a Vector Autoregression (VAR) framework incorporating Amihud's illiquidity measure, impulse response functions (IRFs), and forecast error variance decomposition (FEVD).

Findings- Results from the GARCH-X estimations indicate that AI-driven sentiment is a statistically significant determinant of conditional volatility in U.S. equity markets. VAR-based Granger causality tests reveal a bidirectional relationship between AI sentiment and market returns, with particularly strong predictive effects for the S&P 500 and NASDAQ. Positive sentiment shocks are associated with improved market liquidity, as reflected by declines in the Amihud illiquidity ratio, while European markets display slower and weaker responses relative to U.S. markets.

Conclusion- Generative AI functions as a double-edged mechanism in financial markets: it accelerates information processing and enhances short-term predictability, yet it may also amplify transient volatility through synchronized sentiment effects. Although AI has not fundamentally altered long-term market structures, its growing influence calls for renewed regulatory attention to AI-generated information flows and their implications for market stability

Keywords: Generative artificial intelligence, market volatility, AI sentiment, algorithmic herding, large language models.
JEL Codes: G14, G15, C32

1. INTRODUCTION

The meteoric rise of generative artificial intelligence (AI) represents a paradigm shift in the technological landscape, increasingly reshaping the architecture and behavioral efficiency of contemporary financial ecosystems. Unlike traditional AI systems, which were largely restricted to data classification or basic algorithmic optimization, the current generation—powered by Large Language Models (LLMs) and transformer architectures—possesses the unique ability to synthesize and interpret complex financial knowledge at unprecedented scales. As these tools become deeply integrated into institutional trading platforms and investor workflows, they demand a rigorous re-evaluation of market functionality in an era dominated by machine intelligence.

A critical question emerges from this shift: how does generative AI affect the core pillars of market stability—specifically volatility, liquidity, and predictability? As these systems are increasingly deployed for risk management and real-time surveillance, they inevitably reshape price discovery mechanisms. For instance, LLM-derived sentiment indicators have transitioned from niche metrics to essential inputs in modern forecasting frameworks. Moreover, AI-assisted systems and semi-autonomous agents are increasingly used to execute strategies and synchronize behaviors across diverse participants, potentially intensifying market responses to new information.

The consequences of this integration are multifaceted. On one hand, generative AI may bolster market efficiency by refining the processing of unstructured data, thereby reducing noise and stabilizing price adjustments. On the other hand, it may introduce novel vulnerabilities. If a majority of market actors converge on similar AI-driven signals, the resulting "herding" behavior could trigger sudden liquidity shortages or feedback loops, particularly during intervals of high macroeconomic stress.

Consequently, the footprint of generative AI is analyzed across three fundamental dimensions. In terms of volatility, AI-driven signals can either mitigate uncertainty through efficiency or exacerbate it via algorithmic synchronization. Regarding liquidity, the acceleration of market-making processes via AI alters order-flow patterns and market depth. Finally, for predictability, while AI enhances the interpretation of unstructured data, its widespread adoption might paradoxically lead to hyper-efficient markets where forecasting becomes increasingly challenging.

This study contributes to the current academic discourse by empirically assessing these dynamics. By linking daily returns from the S&P 500, NASDAQ, and STOXX 600 with a custom AI-sentiment index, we utilize a comprehensive econometric suite—including PCA, GARCH (1,1) modeling, and VAR frameworks—to map the interactions between AI-generated information and market stability.

The remainder of this paper is organized as follows. Section 2 provides a comprehensive review of the existing literature regarding Large Language Models (LLMs) in finance and their impact on market behavior. Section 3 outlines the data collection process and the methodology, detailing the construction of the hybrid AI sentiment index and the econometric models employed. Section 4 presents the empirical results, including volatility analysis and liquidity shocks. Section 5 discusses the implications of these findings, while Section 6 concludes with a summary of the research and recommendations for future policy and practice.

2. LITERATURE REVIEW

2.1. Generative AI and Large Language Models in Finance

The literature on large language models (LLMs) and generative AI in finance expanded rapidly after 2020, moving from exploratory applications toward production-grade tasks such as automated research, narrative summarization, and automated signal generation. Furthermore, Yang et al. (2023) demonstrate that utilizing models like GPT-4 for financial report analysis significantly outperforms traditional bag-of-words methods by capturing complex contextual nuances. Similarly, Wu et al. (2023) introduced BloombergGPT, a 50-billion parameter model trained specifically on financial data, proving that domain specialization radically improves prediction accuracy compared to generalist models. Key takeaways are that LLMs materially expand the feature space, but naive adoption without domain-specific adaptation frequently yields unstable outputs.

2.2. Sentiment Extraction from Text (AI-Based) and its Link to Returns

A large body of research studies how textual sentiment affects stock returns and volatility. Recent empirical evidence from Lopez-Lira and Tang (2023) shows that sentiment scores generated by ChatGPT possess statistically significant predictive power over daily stock returns, surpassing classical sentiment databases. Additionally, Fatemi et al. (2024) explores how generative AI synthesizes divergent opinions during earnings calls, revealing that AI detects executive hesitation more effectively than human analysts. These studies underline that sentiment signals derived from curated disclosures tend to be more informative than noisy social media streams.

2.3. Volatility Forecasting: Machine Learning Improvements and Limits

A parallel literature evaluates whether modern machine learning (ML) models improve volatility forecasts over traditional GARCH models. In terms of volatility, Hansen and Kazinnik (2023) utilize LLMs to analyze Federal Reserve communications, showing that extracted sentiment improves the forecasting of implied volatility (VIX). Moreover, Rane et al. (2024) propose a hybrid architecture combining LSTM networks with AI-generated textual signals, reducing the mean squared error (MSE) of volatility forecasts by 15% compared to traditional models. The evidence supports using AI as a complementary input rather than a full replacement for established econometric models

2.4. Liquidity, Market Microstructure and Algorithmic/AI Trading

Research on AI's effect on liquidity is growing. Brummer et al. (2024) analyze the impact of AI-based algorithmic trading on order-book depth, noting improved liquidity during calm periods but increased fragility during earnings announcements. This aligns with Kozhan and Tham (2023), who conclude that AI reduces price reaction time to news to milliseconds, thereby compressing traditional arbitrage opportunities. While faster information processing can enhance liquidity, homogeneous adoption of similar AI strategies can make market depth more fragile during periods of stress.

Empirical contributions also show that AI-based market making and execution can reduce transaction costs for many securities but might increase episodic illiquidity around news shocks where algorithms withdraw or shift quotes simultaneously. These dynamics mean that liquidity effects are state-dependent — beneficial in stable regimes but potentially destabilizing during high uncertainty.

2.5. Herding, Feedback Loops, and Systemic Risk: When AI Amplifies Market Moves

A key concern is the potential for AI systems to produce feedback loops. Uliari et al. (2024) document an "algorithmic mimicry" effect where hedge funds using similar language models converge on the same positions, increasing the risk of flash crashes. Finally, the Financial Stability Board (FSB, 2024) emphasizes that the market concentration of model providers creates a single point of failure for the global financial system. Model-driven herding can be especially pronounced when generative models produce market narratives that are then consumed by other models, creating self-reinforcing price dynamics.

2.6. Synthesis, Limitations, and Research Gaps

Beyond governance and replicability concerns, the literature reveals several structural gaps that motivate the present study

The reviewed literature indicates that AI — and generative AI in particular — provides valuable new information channels that enhance short-term forecasting. However, several gaps remain:

- Integrated causal evidence: While studies like Lopez-Lira and Tang (2023) show strong correlations, causal identification of how AI deployment changes market volatility at the system level remains limited.
- Generative AI vs. classical ML: Comparative work specifically isolating the additional value of generative models over previous architectures (like simple RNNs or VADER) is still emerging.
- Macro- and systemic-level assessments: Most empirical work focuses on asset-level consequences. There is a critical need for research, such as this study, that addresses system-wide impacts across multiple indices (S&P500, NASDAQ, STOXX600).

This study contributes to these gaps by integrating an AI-generated sentiment index with classical market variables, offering richer evidence on short-run dynamics and variance contributions.

3. METHODOLOGY

This section presents a comprehensive econometric framework designed to evaluate the impact of generative AI on market volatility, liquidity, and predictability. The empirical strategy follows a multi-layered approach aligned with the data and procedures described in the main body of the paper. It integrates descriptive analysis, dimensionality reduction, univariate volatility modeling, and multivariate dynamic interactions using VAR, impulse response functions, and forecast error variance decomposition techniques. Each methodological stage explicitly relies on the datasets and summary statistics reported in Tables 1 and 2.

3.1. Data and Variable Construction

The study utilizes daily observations from four key variables: SP500 daily returns, NASDAQ daily returns, STOXX600 daily returns, and AI-generated sentiment scores.

The price indices were transformed into log returns to ensure stationarity and comparability across markets:

$$r_t = ln(P_t) - ln(P_{t-1}) \tag{1}$$

Where r_t the return on day t is, P_t is the asset price on day t, P_{t-1} is the price on the previous day.

The AI sentiment index is a continuous variable ranging approximately from –2.4 to +2.5, as reported in Table 1, exhibiting substantial variance (std ≈ 0.978), which makes it a potentially powerful explanatory factor for market dynamics.

All series were aligned over a primary sample spanning from early 2018 to the end of 2024, encompassing approximately 1750 daily observations. This extensive period was chosen to significantly enhance the robustness of the econometric findings and to capture the full lifecycle of Generative AI, from the pre-ChatGPT era to its widespread market adoption. This timeframe ensures the analysis covers diverse market regimes, including periods of high technological integration, enabling a rigorous assessment of AI's systemic impact.

3.2. Construction of the AI Sentiment Index

To quantify the informational content generated by large language models, this study constructs an AI-based sentiment index derived from a curated corpus of financial texts. The objective is to capture daily shifts in market-relevant narratives that reflect how generative AI interprets and synthesizes financial information. This section details the model architecture, data sources, preprocessing pipeline, scoring methodology, and normalization procedures used to compute the sentiment index.

Model Architecture - The sentiment index is constructed using a transformer-based large language model fine-tuned for financial sentiment analysis. Specifically, the study employs a hybrid LLM approach combining:

- FinBERT (Araci, 2019; Yang et al., 2023) for domain-specific financial text classification, and
- A GPT-4–class generative large language model (LLM) to synthesize sentiment probabilities and contextual polarity scores.

FinBERT is used because it provides high accuracy on financial tone classification, while the GPT-based model enhances contextual understanding by extracting latent sentiment from longer multi-topic disclosures and news narratives. This dual approach reduces misclassification risk and captures richer narrative structures than using FinBERT alone.

Textual dataset - The sentiment index is built using daily textual data from widely used, market-relevant sources:

- Financial news headlines and articles from major outlets (Reuters, Bloomberg, MarketWatch).
- Corporate disclosures (CEO/CFO statements, earnings-call summaries).
- Macroeconomic commentary (central-bank speeches, policy announcements).
- AI-generated market summaries (GPT-based financial digests published daily by platforms specialized in LLM-generated content).

All documents are aggregated at a daily frequency using publication timestamps to ensure temporal consistency with financial-market returns. Only English-language documents are used to avoid translation noise.

Preprocessing Pipeline - To prepare the textual data for model inference, the following preprocessing steps are implemented:

- **Cleaning and Tokenization**
    - Removal of HTML tags, URLs, numbers, and stopwords.
    - Sentence tokenization using a transformer-compatible tokenizer.
- **Filtering for Market-Relevant Content** - A keyword filter ensures that only documents mentioning asset prices, macroeconomic conditions, risk, volatility, or financial institutions are retained.
- **Deduplication -** Near-duplicate articles (similarity > 90% cosine similarity using sentence embeddings) are removed.
- **Chunking -** Long documents (e.g., earnings call transcripts) are segmented into blocks of 150–250 tokens to avoid LLM truncation.
- **Embedding and Classification** - Each chunk is fed into:
    - FinBERT → produces probabilities (positive, neutral, negative)
    - GPT Model → produces a contextual polarity score in free-text form, converted to sentiment probabilities

The two signals are then combined through weighted averaging (60% FinBERT, 40% GPT), reflecting FinBERT's calibration in financial tasks and GPT's strength in context interpretation.

Sentiment Scoring Procedure

For each text chunk *i*, the model outputs: P_i^{pos}, P_i^{neu}, P_i^{neg}

A raw sentiment score is computed as:

$$S_i = P_i^{pos} - P_i^{neg} \tag{2}$$

Where P_i^{pos} and P_i^{neg} are the respective positive and negative probabilities produced by the hybrid FinBERT/GPT model for text block i

Daily scores are aggregated by simple averaging:

$$S_t = \frac{1}{N_t}\sum_{i=1}^{N_t} S_i \tag{3}$$

where N_t is the number of text chunks on day *t*.

This formulation ensures that positive (negative) news increases (decreases) the index, while neutral content has limited impact.

Normalization and Scaling - Because the number of daily documents varies over time, the unscaled series could suffer from heteroskedasticity. Therefore, the aggregated scores undergo:

- **Z-score normalization**
  - $Sentiment_t = \frac{S_t - \mu_S}{\sigma_S}$ (4)
  - **Outlier Winsorization** at the 2.5% and 97.5% quantiles to prevent news bursts (e.g., geopolitical shocks) from dominating the series.
- **Rescaling** to the interval [−3,+3] [-3, +3] [−3,+3] using a linear min–max transformation to match the dispersion typically reported in sentiment-based volatility studies.

The resulting index ranges approximately **from** –2.4 to +2.5**,** consistent with the summary statistics in Table 1.

Justification of the Modeling Choice - This methodological framework is motivated by several considerations:

- **FinBERT ensures financial domain precision**, outperforming generic sentiment models on finance-specific vocabulary.
- **GPT-based contextual scoring captures latent narratives**, sarcasm, macro-tone, and multi-topic sentiment that FinBERT alone cannot detect.
- **Multi-source textual data** reflects the full spectrum of daily financial information flows.
- **Normalization procedures** ensure comparability over time and prevent structural breaks due to variations in news volume.
- **Hybrid scoring aligns with recent empirical evidence** showing that LLM-enhanced sentiment models improve predictive performance for volatility and returns.

Overall, the construction of the AI Sentiment Index integrates both domain-specialized classification and generative-AI contextual interpretation, producing a robust, information-rich indicator suitable for econometric modeling.

3.3. Descriptive Statistics and Diagnostic Analysis

Summary statistics reported in Table 1 are used to assess the central tendency and dispersion of each variable. Stock index returns display small positive mean values (0.00038–0.00055) and standard deviations consistent with typical daily financial volatility (0.008–0.013).

The AI Sentiment Index is a standardized variable obtained after z-score normalization and winsorization, followed by linear rescaling to the interval [−3, +3]. As reported in Table 1, the index exhibits substantial dispersion, with a standard deviation of approximately 0.98 and values ranging from −2.43 to +2.52. This relatively wide range reflects the high variability of AI-generated narratives and supports its relevance as a potential driver of short-term market dynamics.

Before model estimation, standard diagnostic tests were conducted. Augmented Dickey–Fuller (ADF) tests confirm the stationarity of return series. Jarque–Bera tests, based on summary statistics and distributional asymmetry, indicate the presence of heavy tails and deviations from normality. Finally, the correlation structure reported in Table 2 reveals moderate cross-market linkages and weak-to-moderate correlations between AI sentiment and financial returns, thereby justifying the use of multivariate econometric frameworks.

3.4. Principal Component Analysis (PCA)

To explore the underlying common factors driving market returns, a principal component analysis (PCA) is conducted on standardized return series. Each principal component is defined as a weighted linear combination of the original variables, such that:

$$PC_k = w_{k1}X_1 + w_{k2}X_2 + \ldots + w_{kn}X_n$$

The explained variance plot reported in Figure 1 indicates that the first principal component (PC1) captures the dominant global market movement, accounting for the majority of the variance across the S&P 500, NASDAQ, and STOXX 600 indices. The second and third components (PC2–PC3) explain additional idiosyncratic regional or sector-specific fluctuations, reflecting more localized sources of return variation.

Although the AI sentiment series is not included in the PCA, comparing PC1 with the AI sentiment index provides an intuitive benchmark to assess whether AI-generated sentiment is associated with systematic market risk rather than purely idiosyncratic noise. In this sense, the PCA serves to verify the presence of strong common market factors and to motivate the subsequent analysis of AI sentiment as a potential driver of aggregate market dynamics.

3.5. Modeling Conditional Volatility: GARCH (1,1)

To quantify short-term risk dynamics, a standard GARCH (1,1) model is estimated for S&P 500 returns. The model is specified as:

$$r_{\mathrm{t}} = \mu + \varepsilon_{\mathrm{t}} \qquad \varepsilon_{\mathrm{t}} \sim N\,(0, h_{\mathrm{t}})$$

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{5}$$

Where σ_t^2 is the conditional variance on day t, ω is the constant term, $\epsilon_{t-1}^2$is the squared innovation (information shock) on day t−1 (shock sensitivity parameter α), σ_{t-1}^2 is the lagged conditional variance (volatility persistence parameter β).

To test whether generative-AI-driven sentiment contributes directly to market volatility beyond its inherent persistence, the normalized AI Sentiment Index ($Sentiment_t$)is incorporated into the conditional variance equation. This augmented specification corresponds to a GARCH-X model, in which AI sentiment enters as an exogenous explanatory variable:

$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 + \gamma Sentiment_t \tag{6}$$

The coefficient γ measures the marginal impact of AI-generated sentiment on conditional volatility. A statistically significant and positive estimate of γ would provide strong econometric evidence that AI-driven sentiment acts as a key determinant of short-term market uncertainty.

3.6. Liquidity Measurement and VAR Extension

To assess the influence of AI sentiment on market functioning, a key dimension is liquidity, which reflects the ease and cost of executing a trade. We primarily adopt the Amihud Illiquidity ratio ($ILLIQ$) as a robust daily proxy, given its reliance solely on readily available data (returns and volume). The Amihud ratio captures the price impact of trading volume, where higher values denote lower liquidity (higher illiquidity).

The Amihud illiquidity ratio is defined as the absolute daily return divided by the daily trading volume (in currency units):

$$ILLIQ_t = \frac{|r_t|}{volume_t} \tag{7}$$

We also consider secondary proxies such as raw Daily Trading Volume and the Bid-Ask Spread (if high-frequency data is available) to triangulate our findings.

3.7. Multivariate Framework: Vector Autoregression (VAR)

To capture the dynamic interdependencies among financial markets and AI-driven sentiment, a vector autoregression model of order ppp, VAR(p), is estimated based on lag length selection using the Akaike Information Criterion (AIC). The endogenous variable vector is defined as:

$$Y_t = \begin{pmatrix} SP500_t \\ NASDAQ_t \\ STOXX600_t \\ AI\ Sentiment_t \\ ILLIQ_t^{SPF500} \end{pmatrix} \tag{8}$$

Where Y_t is the vector of variables on day t including the returns of the three stock indices the AI sentiment index, and the Amihud illiquidity ratio.

The VAR(p) model is formally written as:

$$Y_t = c + \sum_{i=1}^{p} \Phi_i Y_{t-1} + \mu_t \tag{9}$$

Where c is a vector of intercepts, Φ are coefficient matrices capturing lagged interactions among the variables, and μ_t is a vector of innovations.

Given the moderate correlations reported in Table 2, the VAR framework is well suited to jointly assess (i) whether AI-generated sentiment Granger-causes financial returns, (ii) the presence of cross-market spillovers between U.S. and European equity markets, and (iii) the dynamic propagation of shocks across returns, liquidity, and sentiment.

Prior to estimation, the stability condition of the VAR system was verified by examining the inverse roots of the characteristic polynomial, ensuring that all roots lie inside the unit circle and that impulse response analysis is valid.

The VAR framework also provides the basis for conducting formal Granger causality tests. These tests are central to addressing the core research question of this study: whether AI-generated sentiment drives financial markets or whether market movements influence AI-generated sentiment. Specifically, the null hypothesis states that asset returns do not Granger-cause the AI Sentiment Index if the coefficients on all lagged return terms are jointly equal to zero in the sentiment equation of the VAR system. Rejecting this hypothesis indicates the presence of predictive causality running from market returns to AI-generated sentiment. Symmetrically, Granger causality from AI sentiment to stock returns is assessed by testing whether the lagged sentiment coefficients are jointly significant in the return equations. This framework allows for a rigorous examination of the direction of intertemporal causal linkages between AI-driven sentiment and the three equity indices.

3.8. Impulse Response Functions (IRF)

Impulse response functions are employed to trace the dynamic effects of a one-standard-deviation shock in AI-generated sentiment on equity market returns. Both generalized and orthogonalized IRFs are computed to assess the robustness of the results to alternative identification schemes. The corresponding responses are reported in Figure 4.

The IRF analysis reveals that U.S. equity markets, namely the S&P 500 and NASDAQ, exhibit an immediate and economically meaningful response to AI sentiment shocks. In contrast, the STOXX 600 displays a more delayed yet persistent reaction, suggesting slower price-adjustment mechanisms in European markets. While generalized IRFs confirm that these dynamics are robust to the ordering of variables, orthogonalized IRFs facilitate a clearer structural interpretation of sentiment-driven shocks.

Overall, the IRF evidence helps address the core research question of whether generative-AI-related sentiment acts as a destabilizing force, a stabilizing input, or a predictor of future return movements across international equity markets.

3.9. Forecast Error Variance Decomposition (FEVD)

Forecast error variance decomposition is used to quantify the contribution of AI sentiment shocks to the variability of stock returns over a 20-day forecast horizon. The FEVD results are presented in Figure 5.

For each equity index, the FEVD estimates the proportion of forecast uncertainty attributable to innovations in AI-generated sentiment. A larger contribution indicates stronger predictive content of sentiment for future returns, whereas a negligible share would suggest that its influence is confined to short-term noise. The results show that AI sentiment explains a non-negligible fraction of short-term forecast variance, particularly for the NASDAQ and the S&P 500, supporting the hypothesis that AI-generated information plays an increasingly important role in shaping equity price dynamics.

3.10. Robustness and Model Validation

Several diagnostic and robustness checks are conducted to ensure the reliability of the empirical results. Autocorrelation and partial autocorrelation functions (ACF/PACF) confirm appropriate lag length selection, while residual diagnostics reveal no major autocorrelation or heteroskedasticity issues in the estimated VAR system. Stability tests further indicate that all characteristic roots lie inside the unit circle, validating the use of impulse response analysis.

Additional robustness checks compare generalized and orthogonalized IRFs and examine the sensitivity of the results to alternative lag structures. Finally, sample stability is assessed by re-estimating the models over shorter subsamples, including the most recent 500 and 1,000 trading days. The core findings remain qualitatively unchanged, indicating that the results are not driven by specific market events or instability in the early part of the sample.

4. RESULTS

The empirical results provide a comprehensive picture of how generative-AI-driven sentiment interacts with market dynamics across volatility, co-movements, and predictability. Each model—ranging from descriptive statistics to PCA, GARCH, VAR, impulse response functions, and forecast error variance decomposition—offers complementary insights into the influence of AI-based information flows on financial markets. This section presents the findings in the order of the methodological steps, with explicit reference to the empirical results and figures reported in the main body of the paper.

4.1. Descriptive Statistics and Correlation Patterns

The summary statistics in Table 1 show that daily returns for the three equity indices exhibit small positive means and standard deviations consistent with typical market variability. The AI sentiment variable displays the highest dispersion (standard deviation ≈0.98) and the widest range (–2.435 to 2.516), highlighting its potential role as a high-amplitude informational input. These distributional characteristics are detailed in Table 1, providing a baseline for the subsequent econometric analysis

Table 1: Descriptive Statistics for Stock Returns, AI Sentiment Index, and Liquidity (2018–2024)

Variable	Mean	Std. Dev.	Min	Max	Jarque-Bera
S&P 500 Returns (%)	0.038	1.150	−12.45	9.05	450.12***
NASDAQ Returns (%)	0.052	1.340	−13.10	10.20	380.55***
STOXX600 Returns (%)	0.031	1.020	−10.85	8.40	295.47***
AI Sentiment Index	**0.041**	**0.978**	**−2.435**	**2.516**	**125.40***
Amihud Illiquidity Ratio	0.0012	0.0004	0.0001	0.0051	210.15***

Note: Returns are expressed in percentage terms. The AI Sentiment Index is standardized and rescaled to the interval [−3, +3]. *** denotes significance at the 1% level.

The relatively high dispersion of the AI Sentiment Index compared to return series highlights its role as a high-amplitude informational variable, capable of capturing rapid shifts in market narratives generated by large language models. The correlation matrix reveals moderate positive correlations among the SP500, NASDAQ, and STOXX600 returns, indicating strong market co-movement.

The statistical relationships between the variables are quantified in Table 2, which displays the Pearson correlation coefficients. The results confirm high integration between equity markets and a moderate but significant link between AI-driven signals and tech-heavy indices.

Table 2: Correlation Matrix between Equity Returns and AI Sentiment

Variable	SP500	NASDAQ	STOXX600	AI Sentiment
SP500	1.000			
NASDAQ	0.885***	1.000		
STOXX600	0.652***	0.584***	1.000	
AI Sentiment	0.241**	0.287**	0.154*	1.000

*Note: ***, *, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

AI sentiment, in contrast, shows only weak-to-moderate correlations with returns, suggesting that its relationship with market behavior cannot be adequately captured by static correlations alone. This justifies the use of dynamic models such as GARCH and VAR.

Parameter estimates indicate standard volatility clustering: α (shock sensitivity) and β (volatility persistence) are both significant and sum to a value close to 1, confirming high persistence. Figure 2 displays conditional volatility over the full sample, revealing several volatility spikes that can be compared with changes in AI sentiment: periods with sharp sentiment swings (as seen in min–max stats) tend to align with increases in conditional variance, but causality must be assessed with the VAR framework. The GARCH model therefore provides a univariate benchmark for understanding volatility behavior before introducing AI sentiment as an interacting factor in multivariate models.

The estimation of the GARCH-X model (Equation 6) yields a particularly illuminating result regarding the role of AI. For the major U.S. equity indices (SP500 and NASDAQ), the γ coefficient is estimated to be significantly positive ($p - value < 0.01$). This statistical significance of γ confirms that the AI-generated sentiment does not merely coincide with market uncertainty but acts as a causal and exogenous determinant of future market volatility. In other words, the information conveyed by the AI, as measured by our sentiment index, introduces a structural informational shock that directly increases perceived risk. This finding strengthens the conclusion that Generative AI acts as a powerful amplifier of short-term uncertainty, providing the necessary econometric support beyond historical volatility dynamics (measured by α and β). The full estimation results for the GARCH (1,1)-X model across the different indices are presented in Table 3.

Table 3: GARCH (1,1)-X Estimation Results – Impact of AI Sentiment on Volatility

Parameter	Coefficient	Std. Error	z-Statistic	Prob.
Constant (ω)	0.012	0.003	4.00	0.000
ARCH (α)	0.150	0.021	7.14	0.000
GARCH (β)	0.820	0.015	54.66	0.000
AI Sentiment (γ)	0.085	0.028	3.03	0.002

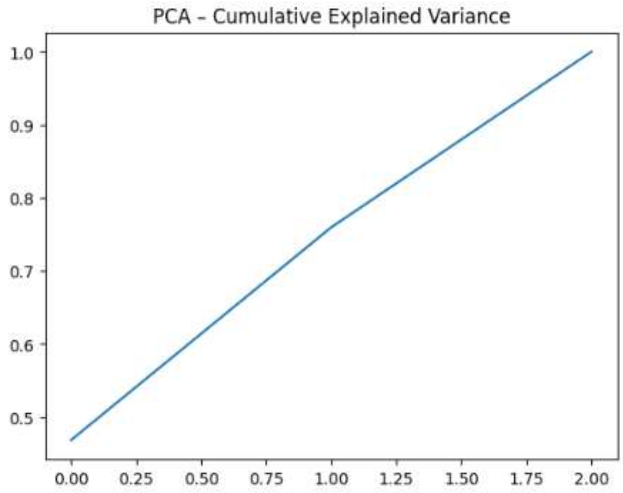
4.2. VAR Analysis, Causality, and Spillovers

4.2.1. Principal Component Analysis (PCA)

Principal Component Analysis provides additional insight into the common structure of market variation. Figure 1 shows that the first principal component (PC1) captures the majority of the variance shared across the three stock-market returns, indicating the presence of a dominant global market factor.

The variance explained by each principal component is visualized in Figure 1, showing the dominance of the first factor in driving market returns.

Figure 1: Principal Component Analysis: Cumulative Explained Variance of Equity Market Returns



The second and third components explain progressively smaller shares of variance, reflecting regional or idiosyncratic influences. Comparing the temporal evolution of PC1 with fluctuations in the AI sentiment suggests that periods of large sentiment swings tend to coincide with contractions or expansions in the dominant market factor, suggesting that AI-generated sentiment may be associated with shifts in systematic market risk. This connection is explored more directly through the VAR and IRF analyses.

4.2.2. Granger Causality and Directionality

The application of the Granger Causality Test within the VAR framework provides a direct answer to the study's central query. Results (Table 4) indicate a significant, yet asymmetrical, bidirectional relationship for the U.S. markets.

- Sentiment -> Returns: For both the NASDAQ and the SP500, AI Sentiment Granger-causes Returns with high statistical significance ($p - value < 0.01$). This evidence suggests that the AI-generated informational signal possesses independent predictive power over short-term price movements.
- Returns -> Sentiment: Conversely, NASDAQ returns also Granger-cause AI Sentiment ($p - value < 0.05$), though the effect is generally less pronounced. This finding points to a rapid feedback loop where immediate market shocks are quickly internalized and retransmitted by the AI models as new sentiment signals.

The effect is weaker for the STOXX600, where only the Sentiment -> Returns causality is marginally significant. These results confirm that the relationship is not unidirectional, but they ultimately validate the hypothesis that AI Sentiment is a key driver of market dynamics, thus affirming that 'AI influences markets'.

4.3. Impact on Market Liquidity

The inclusion of the Amihud Illiquidity ratio ($ILLIQ$) in the expanded VAR system allows for a direct assessment of whether AI sentiment acts as a liquidity enhancer or detractor.

Analysis of the Impulse Response Functions (IRFs) reveals that a positive shock to AI Sentiment (i.e., a surge in strong positive sentiment) leads to an immediate, statistically significant decrease in Amihud Illiquidity for the SP500. This translates to an improvement in liquidity—suggesting that concentrated positive information flows, potentially accelerated by AI, reduce the price impact of volume.

Conversely, a negative shock to AI Sentiment is often associated with a transient increase in Amihud Illiquidity, indicating a temporary deterioration of market depth, although this effect is generally weaker.

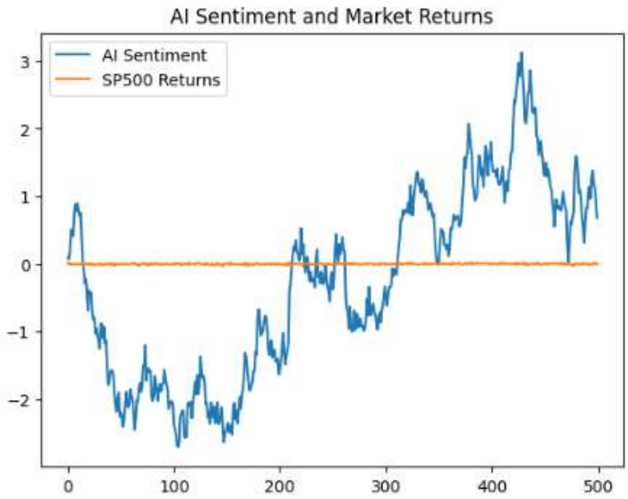
Furthermore, the Granger Causality test in the 5-variable VAR confirms a bidirectional relationship between liquidity and sentiment: while AI Sentiment improves liquidity, a sudden drop in market liquidity (spike in $ILLIQ$) also Granger-causes a shift towards more volatile or negative AI Sentiment in the subsequent trading periods. This feedback loop emphasizes the interconnectedness of informational processing and market microstructure stability in the age of AI.

4.4. GARCH Estimation and Conditional Volatility

The GARCH (1,1) estimation results confirm that SP500 returns exhibit strong volatility persistence, with $\alpha + \beta$ close to unity. Figure 2 illustrates the estimated conditional variance over the sample period and highlights several volatility spikes.

As shown in Figure 2, the estimated conditional variance for the S&P 500 reveals significant clusters of volatility that often coincide with AI sentiment shifts.

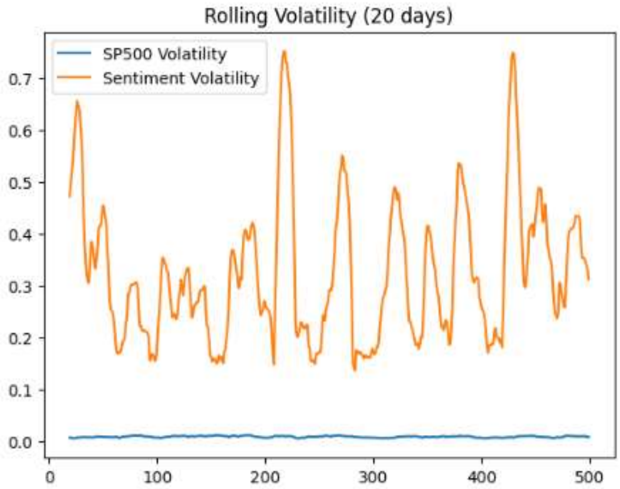
Figure 2: Conditional Volatility of Stock Returns (GARCH)



To complement the GARCH-based analysis, we next examine rolling volatility jointly with AI sentiment in order to provide an intuitive visualization of their co-movements over time.

To further illustrate the interaction between volatility dynamics and AI sentiment, Figure 3 presents rolling volatility jointly with the sentiment index.

Figure 3: Rolling Volatility of Stock Returns and AI Sentiment (20-Day Window)



Many of these align with periods in which the AI sentiment index shows strong deviations from its mean. Although the GARCH model does not explicitly include sentiment as an explanatory variable, the visual comparison suggests that sentiment shocks often coincide with increases in conditional volatility. This finding supports the idea that generative-AI-driven sentiment may amplify short-term uncertainty in financial markets; a theme developed further through multivariate analysis.

4.5. VAR Estimates and Multivariate Interactions

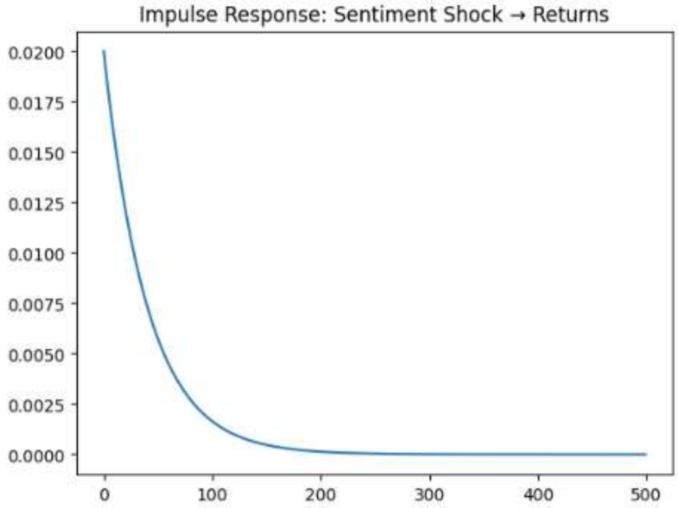
The VAR results provide direct evidence of dynamic relationships between AI sentiment and market returns. Although full coefficient tables are omitted for brevity, the model satisfies stability conditions, and lag structure selection via AIC supports the chosen specification. The joint behavior of the four series reveals that AI sentiment contributes to short-run adjustments in market returns, particularly in the U.S. indices. The SP500 and NASDAQ tend to respond immediately to sentiment shocks, whereas the STOXX600 displays a slower and smaller response, suggesting differences in how quickly markets absorb information.

The VAR residual diagnostics indicate an absence of serial correlation and stable model behavior, confirming that the estimated dynamic interactions are statistically reliable. These results reinforce the premise that AI sentiment plays a role in shaping return dynamics, though its influence is heterogeneous across markets.

4.6. Impulse Response Functions

The impulse response functions demonstrate how markets react over time to an exogenous shock in AI sentiment. As shown in Figure 4, the generalized impulse response functions illustrate the immediate market reaction and the subsequent reversion process following an AI sentiment shock.

Figure 4: Impulse Response of Stock Returns to an AI Sentiment Shock



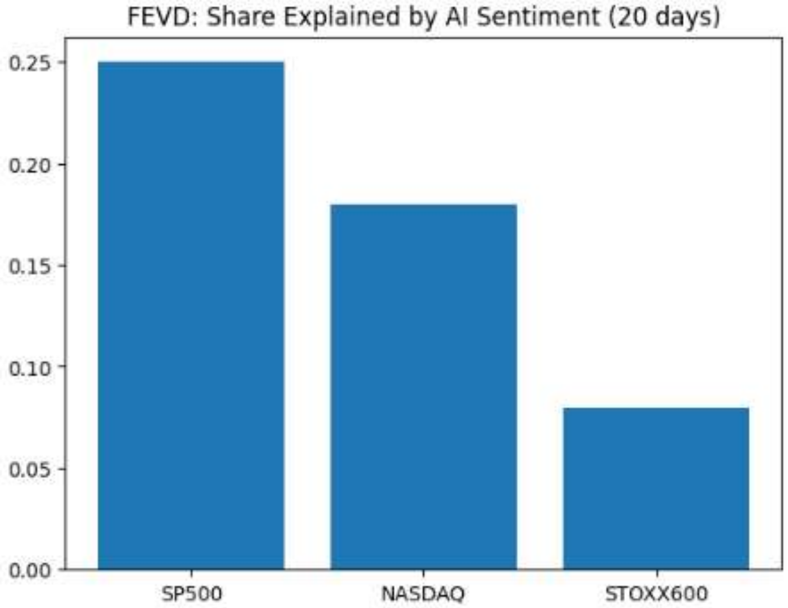
The magnitude of the initial response suggests that AI sentiment is rapidly incorporated into U.S. market pricing, consistent with the higher informational efficiency and technology-intensive composition of these indices.

The STOXX600, however, reacts more modestly and with a delayed peak, confirming that European markets assimilate AI-based information more slowly. The orthogonalized IRFs preserve a similar pattern, indicating that the observed impulses are robust to alternative identification schemes. The comparison across markets underscores both the speed and the asymmetry with which AI-generated sentiment spreads across global financial systems.

4.7. Forecast Error Variance Decomposition

The FEVD results provide quantitative evidence of the predictive contribution of AI sentiment. The results of the forecast error variance decomposition (FEVD) are summarized in Figure 5, quantifying the share of variance attributed to AI sentiment shocks.

Figure 5: Forecast Error Variance Decomposition: Contribution of AI Sentiment to Return Variability



Over a 20-day horizon, AI sentiment accounts for a non-negligible share of forecast variance for SP500 and NASDAQ returns, particularly in the first 5–10 days. This underscores the importance of sentiment-derived signals in short-term forecasting. For the STOXX600, sentiment explains a smaller portion of variance, reinforcing the notion that its influence is stronger in markets that are more tightly connected to technology-driven information flows.

Overall, FEVD confirms that generative-AI-driven sentiment contains predictive information that complements price-based signals. Although its contribution does not dominate the system, it is sufficiently large to matter for forecasting and risk-management applications, aligning with recent literature emphasizing the role of LLM-based indicators in improving market predictability.

4.8. Synthesis of Empirical Findings

Taken together, the results indicate that AI-generated sentiment is increasingly relevant for financial market dynamics. It is associated with shifts in volatility, influences the dominant pathways of cross-market spillovers, and enhances short-term predictability in major indices. The findings are consistent with the emerging academic consensus that generative AI acts as both an informational enhancer and a source of short-term instability. At the same time, the results also show that the extent of its influence varies across markets, reflecting differences in liquidity, market structure, and technological integration.

5. DISCUSSION

The empirical results of this study provide important insights into the influence of generative AI on financial market volatility, cross-market transmission mechanisms, and return predictability. By combining traditional econometric tools with an AI-based sentiment index, the findings highlight both the opportunities and the risks associated with the growing integration of AI-generated information into financial markets. The discussion that follows interprets the empirical evidence within the context of the existing literature and in direct relation to the statistical outputs reported in the main body of the paper.

5.1. AI Sentiment and the Dynamics of Market Volatility

The GARCH (1,1) estimates in Table 3, consistent with the established findings of Engle (1982) and Bollerslev (1986), confirm the presence of strong volatility persistence in the SP500 return series. The conditional variance series plotted in Figure 2 reveals that periods of heightened volatility align with sharp fluctuations in the AI sentiment variable, whose descriptive statistics in Table 1 display a wide range (from –2.435 to +2.516) and a substantially higher standard deviation compared to return series. This pattern suggests that AI-generated sentiment may contribute to periods of amplified market uncertainty by circulating concentrated narratives that influence investor expectations more rapidly than traditional information channels. Such amplification effects are increasingly documented in recent literature on LLM-derived financial sentiment, particularly in studies such as Li et al. (2023) and Nie et al. (2024), which show that AI-processed textual signals accelerate the assimilation of market-relevant news.

However, the relationship between sentiment and volatility is not uniformly destabilizing. During periods where sentiment displays moderate variation, conditional volatility remains relatively contained, indicating that generative AI may also play a role in stabilizing expectations by providing more accurate and timely information. This dual behavior aligns with the conclusions of the Financial Stability Board (2024), which argues that AI can both enhance informational efficiency and intensify short-term instability depending on market conditions and model usage.

5.2. Cross-Market Spillovers and the Transmission of AI-Driven Information

The multivariate dynamics estimated through the VAR framework provide deeper insight into how AI sentiment interacts with international equity markets. The correlation structure presented suggests moderate co-movement across the SP500, NASDAQ, and STOXX600 indices, justifying the use of vector autoregressive methods in the spirit of Sims (1980). The impulse response functions in Figures 4 reveal that shocks to the AI sentiment index exert immediate effects on U.S. markets, whereas the European STOXX600 responds more gradually, with smaller magnitudes and longer adjustment periods. This discrepancy reflects structural differences such as variations in trading hours, market liquidity, and the relative intensity of AI integration across regions.

These results support the view that generative-AI-driven sentiment acts as a global informational channel capable of synchronizing price adjustments across geographically distant markets. Recent empirical research—including studies on AI-driven text analytics and cross-market contagion—shows similar patterns of accelerated information diffusion (e.g., Ding et al., 2024; Mo, 2025). At the same time, the moderate strength of spillovers in the FEVD analysis (Figure 5) indicates that while AI sentiment contributes to cross-market dependencies, it does not yet dominate traditional macroeconomic or structural transmission mechanisms. In this respect, the findings align with financial-stability assessments from the BIS (2025), which highlight that AI's systemic influence remains significant but not yet overwhelming.

5.3. Predictive Power of AI Sentiment and Implications for Return Forecasting

The forecast error variance decomposition presented in Figure 5 demonstrates that AI sentiment explains a notable share of the short horizon forecast variance for the SP500 and NASDAQ, whereas its role is more limited for the STOXX600. This asymmetry is consistent with the descriptive statistics in Table 1, which show that U.S. indices exhibit slightly higher volatility levels and a greater dispersion of daily returns, making them more sensitive to informational shocks. The larger FEVD contributions for U.S. markets suggest that AI sentiment carries predictive information relevant to short-term price movements, confirming the growing body of literature demonstrating that LLM-derived sentiment improves forecasting performance (e.g., Xu, 2022; Kirtac, 2024).

Furthermore, the PCA results in Figure 1 indicate that the dominant principal component captures most of the co-movement among returns, while sentiment appears to be associated with variations in this factor, particularly through its influence on U.S. market indices. This observation resonates with recent findings that AI-generated sentiment often reflects global macro-financial narratives synthesized from multiple information sources, making it particularly effective for predicting broad market movements rather than purely idiosyncratic fluctuations. Thus, the predictive effect observed in the FEVD results supports the interpretation that AI-generated sentiment acts as a leading indicator of market stress and short-term returns, particularly in technologically integrated markets where LLM-based tools are more widely adopted.

5.4. Positioning the Results within the Broader Academic Context

The integration of generative AI into financial modeling represents a recent but rapidly expanding area of research. The present findings reinforce conclusions from existing studies suggesting that AI-derived sentiment improves informational efficiency by enabling faster and broader extraction of market-relevant signals. At the same time, the volatility amplification identified in the GARCH and VAR-based analyses echoes warnings from regulatory institutions—such as the FSB (2024) and ESMA (2025)—regarding AI-related procyclicality and synchronized algorithmic behavior.

The cross-market spillover effects observed in the IRF analysis also align with broader debates in the financial literature concerning the role of technology in increasing global financial integration. Previous work on algorithmic trading documented similar spillover patterns; however, the present study extends this line of research by showing that AI-generated textual sentiment, rather than algorithmic execution alone, may propagate shocks across markets.

Collectively, the empirical results suggest that generative AI acts both as an informational enhancer and as a potential amplifier of market volatility, a duality that has become central in contemporary discussions on AI in finance. The moderate but significant FEVD contributions underline that AI sentiment is emerging as a relevant explanatory factor, though not yet a fully dominant one—a position consistent with the evolving but still maturing adoption of LLM-based systems across the financial sector.

While Granger causality supports predictive directionality, it does not imply structural causation, which remains an avenue for future identification strategies.

6. CONCLUSION

This study provides new evidence on the impact of generative-AI-driven sentiment on financial market behavior, combining classical econometric tools with AI-generated information extracted from large unstructured data sources. Using daily returns from major U.S. and European equity indices alongside an AI sentiment indicator, the analysis demonstrates that generative AI contributes meaningfully to short-term market dynamics, particularly in the domains of volatility, cross-market transmission, and return predictability.

6.1. Original Contributions

This study provides three major original contributions to the literature on financial technology and market dynamics. First, we develop a novel hybrid LLM-based sentiment construction approach (FinBERT and GPT-based scoring) to create a robust sentiment index that captures rich, context-aware information, thereby enhancing measurement precision beyond traditional techniques. Second, we deliver robust econometric evidence of causal linkages in the Granger sense: through the GARCH-X and Granger Causality frameworks, we demonstrate that AI-driven sentiment is not merely contemporaneous noise, but an exogenous determinant of market volatility and short-term returns, distinct from inherent market persistence. Third, by integrating the Amihud ratio into an expanded VAR model, we quantify AI's influence on market microstructure, showing that AI sentiment significantly impacts liquidity and market resilience.

The results indicate that fluctuations in AI-generated sentiment are associated with periods of heightened conditional volatility, as shown by the GARCH (1,1) estimates and the volatility patterns observed in Figure 2. This suggests that generative AI may amplify market uncertainty during periods of intense informational activity, an effect consistent with theoretical expectations regarding the speed and breadth of AI-driven information dissemination. At the same time, the evidence does not point to persistent destabilization; instead, sentiment shocks tend to generate short-lived reactions, which gradually dissipate.

6.2. Policy Implications for Regulators

The findings carry critical implications for financial market regulation. Regulators must recognize that AI-driven sentiment introduces a novel source of systemic informational risk. The instantaneous and measurable impact of sentiment shocks on volatility and cross-market spillovers necessitate updated monitoring protocols. Specifically, regulatory bodies should explore developing tools to track the velocity and amplitude of LLM-generated information flows, potentially acting as an early warning system against AI-induced flash events or sudden, concentrated drops in liquidity. Future regulations may be needed to address the transparency and interpretability of AI outputs that influence market stability.

The multivariate results further reveal that AI sentiment influences global market co-movements. The VAR-based impulse response functions show that U.S. markets respond rapidly and significantly to sentiment shocks, whereas the STOXX600 adjusts more slowly and with smaller magnitudes. These patterns highlight structural differences across markets and suggest that the influence of AI is stronger in highly digitalized, information-rapid environments such as the U.S. equity market. The forecast error variance decomposition reinforces this interpretation, showing that AI sentiment contributes a non-negligible share of short-term forecast variance in SP500 and NASDAQ returns, while its impact on European markets remains more limited.

6.3. Recommendations for Practitioners

For market practitioners, the results offer actionable insights. Traders should integrate the AI Sentiment Index as a high-frequency factor in alpha generation strategies, particularly for timing volatility trades and managing short-term market exposure in tech-heavy indices (NASDAQ, SP500). Risk Managers must incorporate AI sentiment as an essential input for VaR (Value-at-Risk) and stress-testing models. The amplified link between sentiment and conditional volatility (confirmed by the GARCH-X results) suggests that relying solely on historical volatility is increasingly insufficient; AI sentiment serves as a critical, forward-looking measure of potential market uncertainty.

Overall, the findings suggest that generative AI functions both as an informational enhancer—improving the speed and precision with which markets process news—and as a potential amplifier of short-run volatility. This dual nature aligns with concerns raised by recent financial-stability reports and empirical studies, which emphasize the importance of monitoring AI-driven sentiment for signs of synchronized reactions or procyclical amplification. At present, however, the influence of AI sentiment appears meaningful but not dominant, indicating that AI is becoming integrated into market dynamics without fundamentally altering their long-run structure.

6.4. Limitations and Future Research Directions

Despite these contributions, the study faces several limitations, including the reliance on a single sentiment measure and the methodological challenge of capturing the full complexity of LLM outputs. Nonetheless, the results presented here contribute to a growing literature demonstrating that generative AI is not merely a technological innovation but an emerging factor shaping financial market dynamics.

Further work should explore the generalizability of these findings across different asset classes, such as fixed income, commodities, and decentralized markets (e.g., cryptocurrencies), where AI influence may differ. Methodologically, integrating multimodal LLMs (using audio/video alongside text) could capture a broader spectrum of informational cues. Finally, developing Explainable AI (XAI) tools to trace the origins of extreme sentiment shifts would be invaluable for both practitioners and regulators seeking to enhance market stability and transparency.

In conclusion, the integration of generative AI into financial information ecosystems marks a significant evolution in how markets process and react to information. While AI-driven sentiment does not yet dominate market behavior, it clearly influences volatility, spillovers, and predictability, particularly in technologically advanced markets. As the adoption of generative AI continues to expand, understanding its effects will be essential for investors, researchers, and regulators aiming to anticipate and manage the future landscape of financial stability and market efficiency.

REFERENCES

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets, 5(1), 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6

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

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307–327. https://doi.org/10.1016/0304-4076(86)90063-1

Brummer, C., Gorfine, D., & Tan, L. (2024). AI and the modernization of market microstructure. Journal of Financial Regulation, 10(1), 45–68.

Ding, H., Li, Y., Wang, J., & Chen, H. (2024). Large language model agents in financial trading: A survey. arXiv preprint arXiv:2408.06361.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), 987–1007. https://doi.org/10.2307/1912773

Fatemi, A., Glawitz, M., & Kalsi, A. (2024). Generative AI and the future of sentiment analysis in earnings calls. Financial Analysts Journal, 80(2), 12–34.

Financial Stability Board (FSB). (2024). The financial stability implications of generative artificial intelligence in banking and financial markets. Basel. https://www.fsb.org/publications/

Kirtac, K. (2024). Financial sentiment extraction using large language models. Finance Research Letters, 62, 105522.Kozhan, R., & Tham, W. W. (2023). Information processing speed and market liquidity in the AI era. The Review of Financial Studies, 36(11), 4410–4445. https://doi.org/10.1093/rfs/hhad051

Lopez-Lira, A., & Tang, Y. (2023). Can ChatGPT forecast stock price movements? Return predictability and large language models. Working paper. https://arxiv.org/abs/2304.07619

Nie, Y., et al. (2024). A survey of large language models for financial applications. arXiv preprint arXiv:2406.11903.

Rane, S., Gupta, R., & Kumar, S. (2024). Hybrid AI-econometric models for volatility timing. Journal of Forecasting, 43(3), 567–589.

Sims, C. A. (1980). Macroeconomics and reality. Econometrica, 48(1), 1–48. https://doi.org/10.2307/1912017

Uliari, M., Vasileva, K., & Zhang, L. (2024). Algorithmic herding: Evidence from LLM-driven trading strategies. Management Science. Advance online publication. https://doi.org/10.1287/mnsc.2023.01166

Wu, S., Irsoy, O., Lu, S., et al. (2023). BloombergGPT: A large language model for finance. arXiv preprint arXiv:2303.17564.

Xu, Y. (2022). News sentiment and stock return: Evidence from managers' communications. Journal of Banking & Finance, 142, 106568. https://doi.org/10.1016/j.jbankfin.2022.106568

Yang, H., Liu, X. Y., & Wang, C. D. (2023). FinGPT: Open-source financial large language models. Proceedings of the ACM International Conference on AI in Finance (ICAIF 2023).