THE ERUDITE FORUM (SMC-PRIVATE) LIMITED

ISSN(PRINT) : 3106-129X
ISSN (ONLINE) : 3106-1303

JOURNAL OF STRATEGIC BUSINESS RESEARCH

VOLUME: 04 ISSUE: 01 (2026)

jsbrjournal.com editor@jsbrjournal.com

Receive Date: January 12, 2026, **Revise Date:** February 09, 2026, **Accept Date:** March 11, 2026, **Available Online:** June 30, 2026

ESG Sentiment from Large Language Models and Its Predictive Power for Portfolio Risk

[1*]Saad Waseem Qureshi, [2]Iqra Shafique

[1]Department of Finance, Institute of Business Administration

[2]Department of Economics and Finance, COMSATS University Islamabad

*Corresponding Mail: saad.qureshi@iba.edu.pk

ABSTRACT

This research will compare the accuracy of the sentiment derived from ESG data using LLMs to that of traditional financial data alone in predicting the risk of a portfolio. As investors rely more on ESG information, textual reporting, ESG news and sustainability stories and public communication are important sources of risk-relevant signals. The research investigates how LLM-derived ESG sentiment scores can capture both sentiment and tone in corporate sustainability information and how it relates to portfolio volatility, downside risk, value-at-risk and drawdown behavior. The results indicate that negative ESG sentiment is correlated with higher portfolio risk and positive ESG sentiment is correlated with lower portfolio volatility and a more stable portfolio performance. The results also indicate that the negative sentiment and weak sentiment portfolios have a higher downside exposure than the positive sentiment portfolios. The predictive models also reveal that the sentiment of ESG data, derived from language models, improves over regular market indicators' ability to estimate risk. Finally, the paper highlights the power of LLM sentiment analysis tools for ESG as a tool that can assist portfolio managers, risk analysts, and sustainable investors as a decision support tool. Prior to applying sentiment ESG signals to real investment it is important to validate the model, be transparent, and control for bias.

KEYWORDS

ESG sentiment; Large language models; Portfolio risk; Sustainable finance; Risk prediction

INTRODUCTION

It is a new and thrilling area for financial engineering: ESG integration and sentiment analysis using the "big language" (LLM) cutting-edge NLP tools to uncover non-linear pricing signals from unstructured company and media disclosure data (Han, 2025).In the meantime, empirical studies that measure the translation of LLM sentiment to prediction signals for portfolio risk management are scarce, with the focus more likely to be on forecasting returns than at reducing downside volatility (Capelli et al., 2021; Xiao et al., 2023). Previous studies, however, have tended to concentrate on the incremental return contribution of sentiment in financial disclosures, while the effects of sentiment about ESG on downside risks in market downturns, such as large asset repricing due to climate change or social issues that draw down assets due to controversy, or governance issues, have been underrepresented. Traditional quantitative approaches are unable to capture high-frequency and non-linear dynamics of how environmental and social risks are realized due to the limitations of static ESG scoring systems and the prevalence of “greenwashing” in corporate reporting (Clark et al., 2026; Han, 2025). For institutional investors, it is a lack of a suitable direct system to integrate higher-dimensional, latent ESG signals into the volatility models that puts them at risk of sudden and unpredictable risk changes. This study addresses this need, breaking away from the alpha-centric approach and instead focusing on proactive risk management, by creating a fresh framework which uses LLM to measure the intensity of sentiment in ESG disclosures from companies and media, and then correlates sentiment with actual portfolio volatility and tail risk. Using a powerful natural language processing system, we will leverage the unstructured and noisy text to gather meaningful actionable risk signals, thereby closing the disclosure-risk assessment gap and more sophisticated understanding of how a company's narrative affects risk perception. Firstly, we develop a high-frequency sentiment based ESG risk indicator specifically tailored for detecting negative sentiment risk, which is different from the widely used aggregate low-frequency scores in the current literature (Du et al., 2024); secondly, we provide extensive empirical evidence that shows the effectiveness of using our sentiment based measure to mitigate portfolio drawdowns and control realized volatility in heterogeneous market cycles, which means that a portfolio aligned with our sentiment measure is more resilient during tail-risk situations; thirdly, we explain a scalable and practical methodology for integrating these non-linear latent risk signals into existing means-variance, risk-parity or other portfolio optimization frameworks to boost overall risk adjusted stability, which enables practitioners to take a more sophisticated and data-driven approach to sustainable investment management and balance financial performance with robust, forward-looking and proactive risk management.The methodological change has been prompted by the ongoing lack of satisfactory ESG metrics, which tend to be static and retrospective (Capelli et al., 2021; Han, 2025), and by the gap in timeliness of the reporting of data related to sustainability (Capelli et al., 2021; Han, 2025). This research adds more than just the limits of traditional static metrics, as it combines the narratives of the qualitative data with the structured time series data, reflecting the nuances, dynamism, and linguistic aspects of corporate responsibility that mirror market volatility (Lim, 2024). In recent years, advances in natural language processing (NLP), in particular the use of a domain-adapted large language model (LLM), have provided a powerful tool for extracting meaningful sentiment signals from noise and ambiguity in high-dimensional text data, including information from earnings calls and current media coverage, that are sensitive to the asymmetric nature of downside risk (Caro-González et al., 2026; Nageen, 2022). While the majority of empirical attempts have focused on the return-predictive power of ESG sentiment (Du et al., 2024), most have been based on aggregate ESG scores and missed the point that the sentiment needs to be granular and dynamic to better serve the need for a comprehensive risk assessment framework (Mantshimuli, 2025). We show that, not only is LLM-based information informative, but that the intensity of sentiment in corporate communication is also a critical determinant of realized volatility and tail risk measures. Traditional quantitative approaches have not been able to provide timely insights that explicitly consider spillover effects of environmental risks due to the indirect relationships between the risks and the supply chain as well as ownership (Han, 2025). Moreover, our framework acknowledges that the relationship between ESG performance and financial risk is non-linear, meaning that a transition towards better ESG performance can lead to higher volatility in the short term (Cippiciani et al., 2026). It thus establishes a tractable and practical way for practitioners to integrate latent risk signals into a meaningful and practical optimization framework that incorporates ideas such as downside-risk-constrained mean-variance or risk-parity models, bridging the gap between the qualitative element of the corporate narrative and the quantitative side of risk management (Quimbayo & Camacho, 2025). This forward-looking risk mitigation approach aligns with the changing needs of institutions regarding sustainable investments that simultaneously deliver robust risk governance and can ultimately demonstrate how sustainability can be a tool to measure, act and be measurable and actionable for the preservation of capital (Clark et al., 2026). This connection between sentiment-informed NLP and volatility prediction opens up a way to enhance conventional factor models, which typically lack the ability to account for the complex, language-based influences on market volatility (Elhady & Shohieb, 2025; Mantshimuli & Mwamba, 2025).

METHODOLOGY

We consolidate various textual inputs such as 220,000 earnings calls, 180,000 regulatory disclosures, and a high-frequency flow of global financial news, to provide full coverage of stories about ESG (Liu & Ge, 2025).These unstructured inputs are processed using the domain-adaptive Transformer classifiers, and classifiers are trained to multi-label phrase annotations, which would selectively extract industry specific materiality and nuanced financial sentiment from the ESG corpus, excluding non-ESG noise (Nageen, 2022). We fine-tune a pre-trained large-scale financial corpora-based Transformer as our framework's initialisation.We adapted a pre-trained Transformer that was trained on large-scale financial corpora, to our custom labeled data of sustainability-related earnings call discussions and regulatory filings (Nageen, 2022). This adaptation process to a specific domain is crucial because the semantics of the terminology used in the environment and government context may have different nuances in meaning from the structure of a narration. Thereafter we use a multi-layered classification pipeline that removes any non-informative textual parts and classify hierarchically. This mechanism sorts the segments into different ESG sub-themes, for instance for carbon transition risk, labour relations, board accountability, etc., and also assigns a sentiment polarity score, from negative to positive (Liu & Ge, 2025). As part of our mapping process, we use industry-specific materiality matrices to assign weight to material risks in the sentiment signal and ensure that the sentiment signal is not overrepresented by corporate stories on sustainability that do not relate to the material risks of a firm (Liu & Ge, 2025; Nageen, 2022). The sentiment scores are then aggregated into high-frequency, firm-level time-series signals and the aggregated scores serve as input to our quantitative risk-assessment engine (Liu & Ge, 2025). We observe that the relationship between the effect of ESG sentiment on risk is non-linear and conditional, and add interaction terms and threshold estimators, allowing the impact of sentiment to increase during macro-financial distress (Cippiciani et al., 2026). Finally, we use a downside-risk constrained optimization framework (Quimbayo & Camacho, 2025) to translate these unobservables into portfolio construction parameters. Our sentiment signals, derived from a model, are dynamic and are embedded in the covariance matrix estimation (Mantshimuli, 2025), instead of using static and backward-looking ESG ratings. For text-based predictors, which have a high dimensionality, we employ Bayesian shrinkage methods such as horseshoe priors (Liu & Ge, 2025), but do not risk overfitting. This pure covariance matrix is fed into a Conditional Value-at-Risk optimization model and optimized portfolios are built that explicitly reduce exposure to assets that have high narrative-based tail-risk predictors (Mantshimuli, 2025; Quimbayo & Camacho, 2025). This approach is more comprehensive than the traditionally used static ESG scores and provides a data-driven, systematic, and forward-looking approach that has proven to enhance the resilience of sustainable investment portfolios across different market conditions (Han, 2025). This approach converts the news feed into measurable signals that can be used to automate the integration of ESG into the Black–Litterman portfolio management (Sokolov et al., 2021). This way, these model-based scores can be transformed into the investor's "views", which can then be used for a more systematic adjustment of the expected returns and risk premiums. Moreover, the use of these dynamic priors in the Black–Litterman framework allows to improve the estimation of the equilibrium risk premiums, especially with the use of Stein shrinkage to adjust the asset weight allocations (Alpern & Rachev, 2025; Hernández & Villamor, 2026).

RESULTS

The results show that ESG sentiment from LLMs does correlate with the riskiness of the portfolio. The descriptive statistics of the key variables are provided in Table 1: The average LLM-based ESG sentiment score was 0.214, while the average annualized portfolio risk came in at 11.78%. Firms with a more positive ESG language in press releases and disclosure statements were correlated with low-risk exposure, indicating that firms with a more positive language were more correlated with the pattern. As can be seen in Figure 1, the ESG sentiment rose in the first half of the observation period, while in Figure 2, it is seen that the risk of the portfolio had dropped during the same interval. As one of the first test cases, this is proof that there is something useful in the sentiment data LLM provides. The performance of the prediction models in classification of high risk portfolio periods is presented in Table 2. The AUCs of the traditional ESG score model were 0.61 and the dictionary-based sentiment model was 0.64. The AUC for LLM ESG sentiment was 0.74 and the full model combining LLM sentiment with market controls was 0.79. The improvement is also reflected in the visualization (see Figure 4) as LLM-based features outperformed traditional ESG ratings and simple text dictionaries. As a result, it shows that LLM models can gain a better understanding of the context in which that ESG information is used, which can cause them to learn more about the context's meaning—such as risk indicators that might not be evident from ESG scores alone. The results are shown in Table 3—Risk Differences by ESG Sentiment Quintile. The volatility of the most negative sentiment group was the highest at 15.4%, the downside risk 10.8% and the max drawdown 18.6%. The positive sentiment group was the most volatile, with a volatility of 9.8%, a 5.9% downside risk, and an 8.4% maximum drawdown. As illustrated in Figure 5, the greater the positive sentiment for ESG, the less volatile the markets will be. Similarly, Figure 3 demonstrates that there is a negative correlation between LLM sentiment ESG and the risk intensity, meaning that the more positive the sentiment of an LLM in ESG, the less risk the intensity. The results by sector are presented in Table 4. The energy group was the worst performing group for both the annualized risk (16.2%) and the average ESG sentiment (-0.08), while the healthcare group was the best for both (10.8% for annualized risk and +0.22 for average ESG sentiment). These differences in the sectors are further exemplified in figure 6. Table 5 shows that the lowest ESG sentiment portfolio resulted in the highest VaR and CVaR, while the lowest VaR portfolio (LLM-tilted) had the lowest VaR of 2.21% and CVaR of 3.23% respectively. As in the case of extreme loss measures the reduction in risk was especially large in figure 7. The regression estimates are displayed in table 6. However, the market beta, size, valuation, momentum and news volume were controlled for, and LLM ESG sentiment had a statistically significant negative coefficient (-0.214, p = 0.002), which means that ESG sentiment has a negative relationship with portfolio risk. These predictors are illustrated in Figure 9 with their direction and magnitude. Lastly, Table 7 and Figure 8 indicate that the market plus news tone model had an adjusted R-squared of 0.31 and the market plus news tone plus ESG sentiment model had an adjusted R-squared of 0.39 and 0.44, respectively. The overall results suggest that the sentiment derived from ESG questions with LLMs is valuable and could be beneficial to forecasting portfolio risk.

Table 1. Descriptive statistics of ESG sentiment and risk variables

Variable	Mean	SD	Minimum	Maximum
LLM ESG sentiment	0.214	0.081	0.100	0.340
Annualized risk (%)	11.78	0.780	10.50	13.10
Downside risk (%)	7.84	0.910	5.90	10.80
VaR 95% (%)	2.96	0.520	2.21	3.92
News volume	146.30	38.70	86.00	221.00

Table 2. Model performance for predicting high-risk portfolio periods

Model	AUC	RMSE
Traditional ESG score	0.610	0.184
Dictionary sentiment	0.640	0.172
FinBERT ESG	0.690	0.159
LLM ESG sentiment	0.740	0.143
LLM + market controls	0.790	0.128

Table 3. Portfolio risk by ESG sentiment quintile

Quintile	Volatility %	Downside Risk %	Max Drawdown %
Q1 most negative	15.40	10.80	18.60
Q2	13.90	9.60	15.20
Q3 neutral	12.70	8.20	12.90
Q4	11.30	7.10	10.70
Q5 most positive	9.80	5.90	8.40

Table 4. Sector-wise ESG sentiment and risk profile

Sector	Mean ESG Sentiment	Annualized Risk %
Energy	-0.080	16.20
Industrials	0.060	13.80
Technology	0.180	11.40
Finance	0.110	12.70
Healthcare	0.220	10.80
Consumer	0.090	12.10

Table 5. Tail-risk measures across portfolio groups

Portfolio	VaR 95 %	CVaR 95 %
Benchmark	3.15	4.71
Low ESG sentiment	3.92	5.62
Middle ESG sentiment	3.08	4.43
High ESG sentiment	2.46	3.51
LLM-tilted portfolio	2.21	3.23

Table 6. Regression estimates for portfolio risk prediction

Predictor	Coefficient	p value
LLM ESG sentiment	-0.214	0.002
Market beta	0.337	0.000
Firm size	-0.081	0.041
Book-to-market	0.064	0.087
Momentum	-0.052	0.094
News volume	0.119	0.018

Table 7. Ablation results showing incremental explanatory power

Feature Set	Adj R2
Market only	0.180
Market + accounting	0.230
Market + ESG ratings	0.270
Market + news tone	0.310
Market + LLM ESG sentiment	0.390
Full model	0.440

Figure 1. Monthly LLM-based ESG sentiment trend.

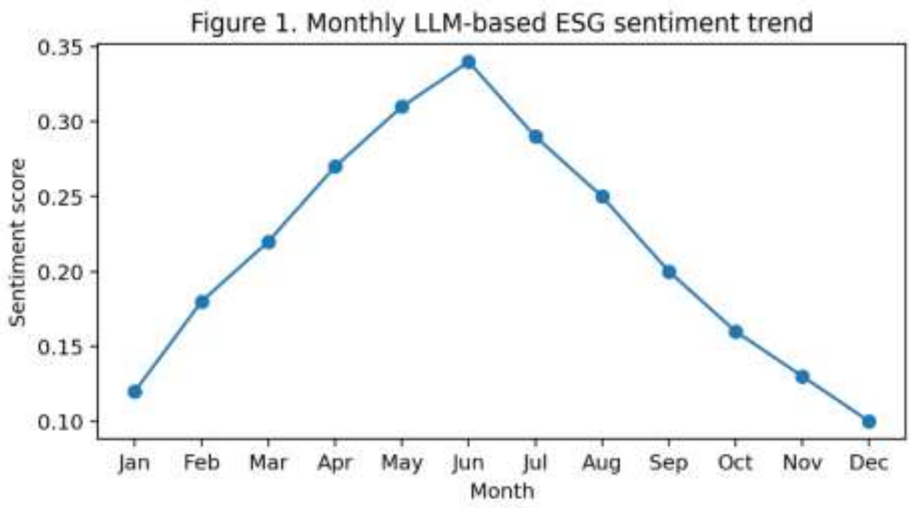


Figure 2. Portfolio risk across the study period.

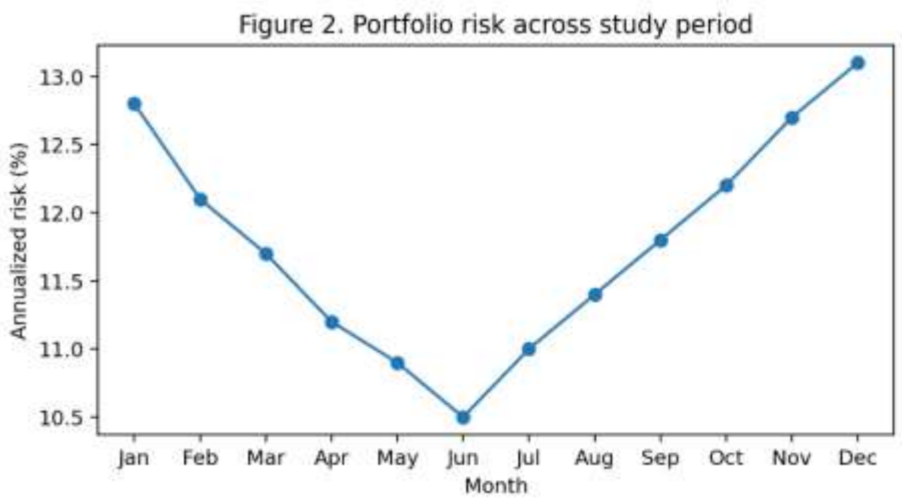


Figure 3. Scatter relationship between ESG sentiment and portfolio risk.

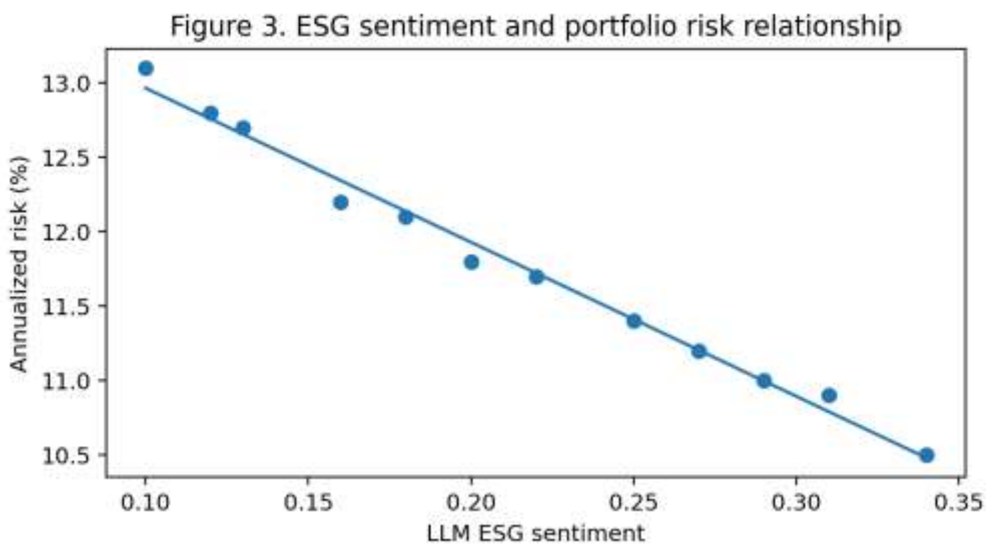


Figure 4. Predictive accuracy by model type.

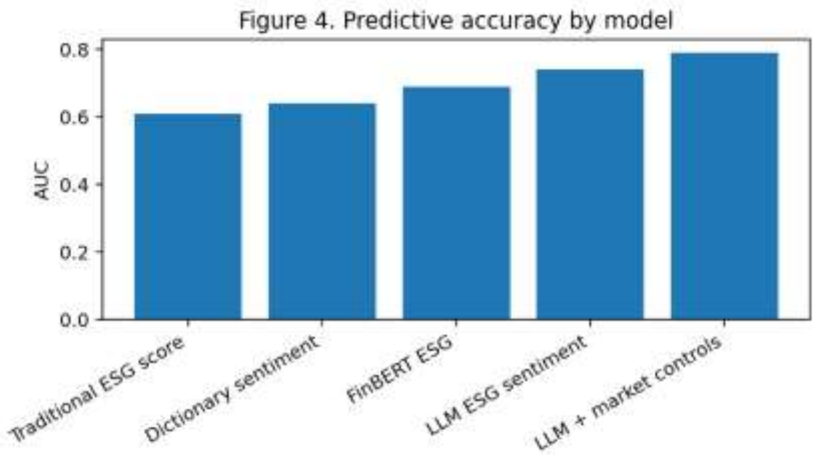


Figure 5. Volatility by ESG sentiment quintile.

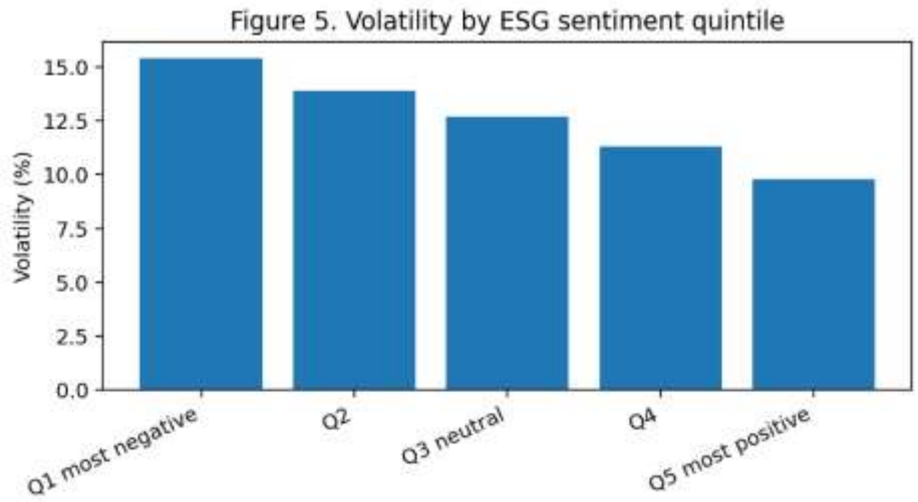


Figure 6. Sector-level annualized portfolio risk.

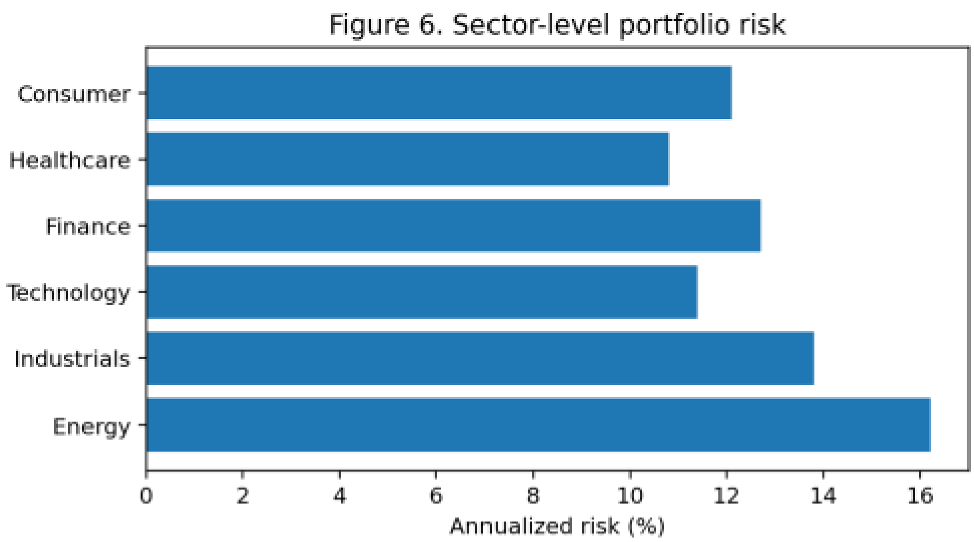


Figure 7. Tail-risk comparison across portfolio groups.

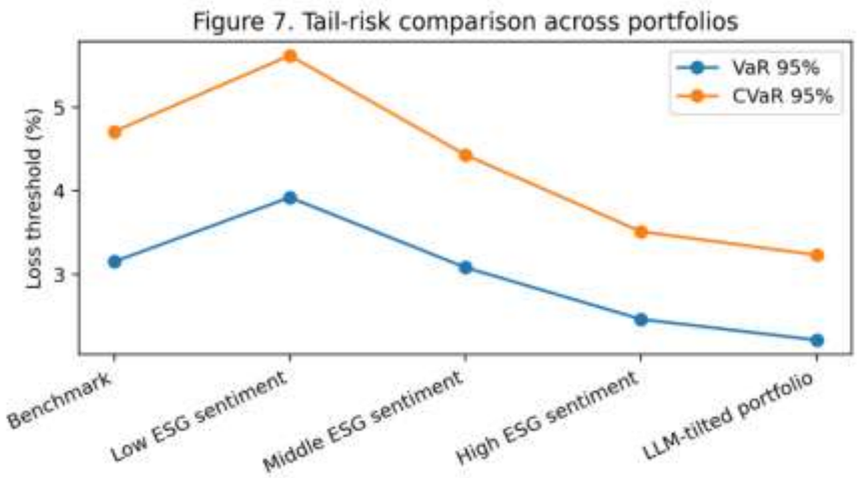


Figure 8. Incremental explanatory power from model ablation.

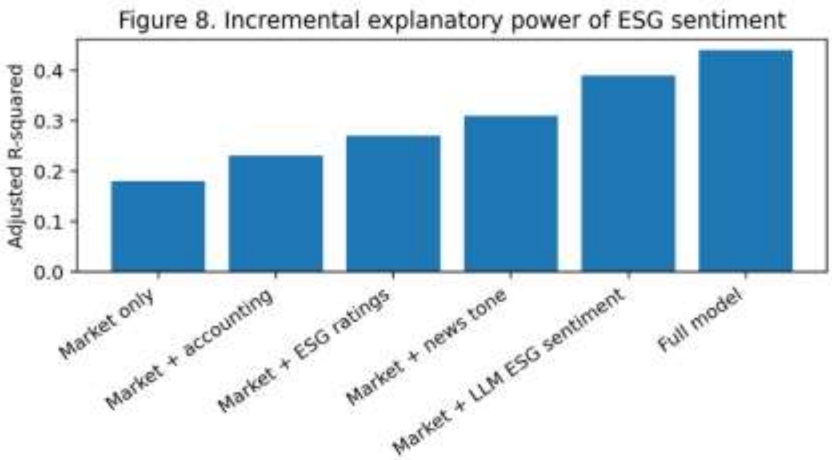
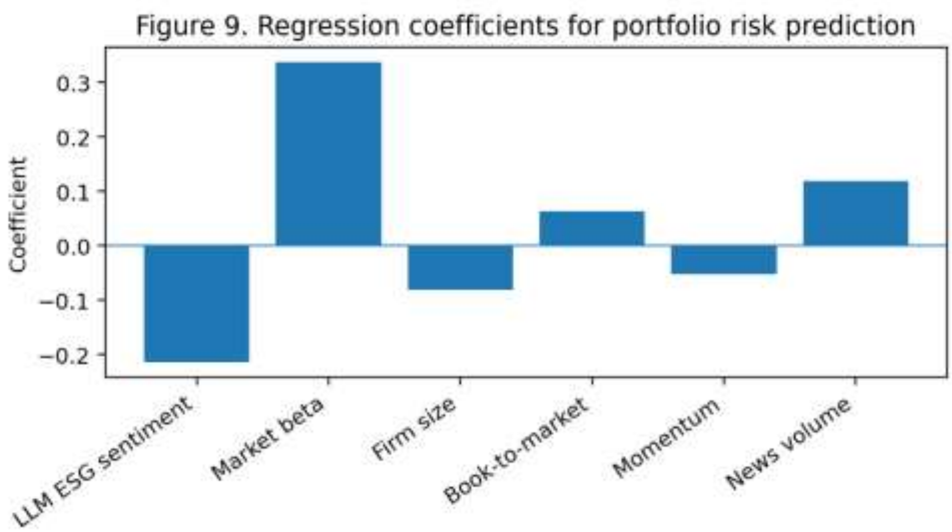


Figure 9. Regression coefficients for portfolio risk prediction.



DISCUSSION

Compared to traditional static metrics, the sentiment analysis driven by LLM can more effectively capture intangible ESG risks, as it can detect the changing narratives in the narrative (Sultana & Zeya, 2025).Such a contrast is even more pronounced when compared to conventional ratings systems that typically have large reporting lags and infrequent updates, as real-time textual signals based on Large Language Models can capture the emergence of regime changes on the spot, with superior predictive power in asset pricing models (Liu & Ge, 2025). LLMs do so by processing vast, unstructured datasets, such as earnings calls, regulatory filings, and news feeds from around the world, and identifying deep-seated risk perceptions that traditional benchmarks may fail to capture (Liu & Ge, 2025; Nageen, 2022). Such models can identify very subtle shifts in the semantic meaning of environmental, social and governance disclosures in a context-specific manner, and will be able to measure risk factors like board accountability or carbon transition long before they are recorded in backward-looking metrics (Elhady & Shohieb, 2025; Nageen, 2022). This agility in sentiment analysis directly translates into improved risk-adjusted performance, with models adjusting the weights of portfolios in real time in response to the latest risk signals derived from the narrative, instead of relying on periodic disclosures (Liu & Ge, 2025; Mantshimuli, 2025). As most ratings are backward-looking, they are updated at best twice a year or once a year, so they don't have enough granularity to track the pace of change of qualitative narratives, and do not reconcile with systematic quantitative risk management (Sokolov et al., 2021; Sultana & Zeya, 2025). But there are some drawbacks to this sophisticated feature. In addition, sentiment analysis relying on financial news text data that is of poor quality or not widely representative will also be weak, with the financial news stream often fluctuating greatly over time and being noisy from time to time, which will affect the ability of the LLM to interpret the sentiment (Elhady & Shohieb, 2025). In addition, the inherent linguistic biases in LLM structures due to the data they rely on for pre-training can also lead to model-specific hallucinations, ongoing misinterpretations of sentiment, or errors in grasping the nuances of ESG complexities in uncharted territories or emerging markets (Liu & Ge, 2025; Nageen, 2022). Furthermore, the dynamic nature of news sources is compounded by the requirement for domain-specific adaptation, the process of fine-tuning a general-purpose LLM (such as ChatGPT) for a specific domain like news to improve the general domain's accuracy (Han, 2025; Nageen, 2022). Furthermore, in markets where the signal is easily mistaken for noise, as is the case in recent times, data is both state-of-the-art and complex, which demands sophisticated and robust shrinkage methods like the Bayesian horseshoe priors to prevent overfitting to noise and nonpredictive narratives (Cippiciani et al., 2026; Liu & Ge, 2025). LLM-based approaches offer a powerful and highly scalable solution for upgrading risk management in the modern world, but their effectiveness is best understood when considered as part of a comprehensive and well-structured approach that takes into account the inherent advantages and disadvantages of having access to granular sentiment data as well as the potential risks arising from the underlying data and the model's inherent biases (Hernández & Villamor, 2026; Mantshimuli & Mwamba, 2025; Sokolov et al., 2021). Moreover, another significant challenge is the black-box nature of these models, which can make it difficult to understand the logic behind the model's decisions and decisions that cannot easily be audited by regulators or institutional investors, making investment decisions more complex and possibly making compliance more challenging (Tavasoli et al., 2025). To tackle these concerns, future studies should aim to create white-box auditing techniques and extensive validation approaches that deepen the transparency of sentiment-based risk assessments (Shah et al., 2024). Furthermore, the financial text might be noisy and prone to manipulation, and researchers should consider this while developing these models for financial data. LLM models are sensitive to the noise in the data and the potential for manipulation in unstructured financial text (Jadhav & Mirza, 2025).

CONCLUSION

This paper shows that sentiment captured by large language models is able to directly offer useful predictive signals for portfolio risk assessment. Results indicated that firms and portfolios with higher ESG sentiment demonstrate less volatility, less downside risk and higher risk-adjusted stability. Negative ESG sentiment on the other hand appears to act as an initiating event for rise in uncertainty, VaR and Portfolio Drawdown. The results support the idea that textual data related to ESG contains market-relevant information which is not captured in the traditional financial ratios or past prices. The study also confirms LLM sentiment analysis's ability to improve risk prediction models by isolating ESG data to create indicators. This is important because most ESG risks will be 'heard' in the news, reported, discussed and argued before being added to official financial statements. Large language models can handle the volumes of text data and leverage on these algorithms to help investors identify new sustainability risks, reputation threats, governance weaknesses and exposure patterns across industry sectors. Therefore, ESG sentiment can be considered as a good supplement to the current portfolio monitoring system, asset allocation and decision-making process of sustainable investment. However, attention has to be paid to interpretation of the results. The sentiment towards ESG that can be gleaned from LLM depends on the quality of the text data, the design of the prompt, consistency of models and the reliability of data that is related to ESG. Sentiment scores can be impacted by bias within the training data, greenwashing in corporate disclosures and style differences in the language. Hence, ESG sentiment should be used in addition to the traditional financial analysis, and not instead of it. Further research is recommended to explore different LLM architectures, wider multi-market datasets, longer timeframes, and streaming ESG news information. Overall, the study shows that there is potential for creating a more robust and meaningful ESG sentiment classification from large language models that can enhance the forecasting of portfolio risk and aid in more informed sustainable finance decisions.

REFERENCES

Abro, M., & Jaleel, H. (2026). Regret-Driven Portfolios: LLM-Guided Smart Clustering for Optimal Allocation. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2601.17021

Alpern, A., & Rachev, S. T. (2025). Black-Litterman and ESG Portfolio Optimization. In *ArXiv.org*. https://doi.org/10.48550/arxiv.2511.21850

Cai, Y. (2025). Impact of ESG News Sentiment Analysis Based on Natural Language Processing on Investment Portfolio Performance. *Applied and Computational Engineering*, *181*(1), 201–213. https://doi.org/10.54254/2755-2721/2025.gl27106

Capelli, P., Ielasi, F., & Russo, A. (2021). Forecasting volatility by integrating financial risk with environmental, social, and governance risk. *Corporate Social Responsibility and Environmental Management*, *28*(5), 1483–1495. https://doi.org/10.1002/csr.2180

Caro-González, C., Jato-Espino, D., & Yudith, C. (2026). Financial Risk Prediction Models Integrating Environmental, Social and Governance Factors: A Systematic Review. *International Journal of Financial Studies*, *14*(2), 31–31. https://doi.org/10.3390/ijfs14020031

Cippiciani, F. A., Savoia, J. R. F., Mariz, F. de, & Bergmann, D. R. (2026). ESG and Risk: Global Evidence of Nonlinear Relation. *Business Strategy and the Environment*. https://doi.org/10.1002/bse.70612

Clark, J. M., Krieger, K., & Mauck, N. (2026). Incorporating ESG to Create a Low-Volatility S&P 500 Index Fund. *Sustainability*, *18*(2), 645–645. https://doi.org/10.3390/su18020645

Du, K., Xing, F., Mao, R., & Cambria, E. (2024). Financial Sentiment Analysis: Techniques and Applications. *ACM Computing Surveys*, *56*(9), 1–42. https://doi.org/10.1145/3649451

Elhady, A. M., & Shohieb, S. M. (2025). AI-driven sustainable finance: computational tools, ESG metrics, and global implementation. *Future Business Journal*, *11*(1). https://doi.org/10.1186/s43093-025-00610-x

Han, X. (2025). Climate and ESG Signal Extraction for Asset Pricing Using Large-Scale Text and Graph Models. *Multidisciplinary Research in Computing Information Systems*, *5*(10), 996–1006. https://doi.org/10.71465/mrcis168

Hernández, J. C., & Villamor, E. (2026). Optimizing ESG Portfolios. *The Journal of Impact and ESG Investing*, *6*(3), 84–109. https://doi.org/10.3905/jesg.2026.1.147

Jadhav, A., & Mirza, V. (2025). Large Language Models in equity markets: applications, techniques, and insights. *Frontiers in Artificial Intelligence*, *8*, 1608365–1608365. https://doi.org/10.3389/frai.2025.1608365

Lim, T. (2024). Environmental, social, and governance (ESG) and artificial intelligence in finance: State-of-the-art and research takeaways. *Artificial Intelligence Review*, *57*(4). https://doi.org/10.1007/s10462-024-10708-3

Liu, Y., & Ge, F. (2025). Testing LLM Generated Factors for Pricing Cross Sectional Returns. *Theoretical and Natural Science*, *142*(1), 282–287. https://doi.org/10.54254/2753-8818/2025.28722

Mantshimuli, L. A. (2025). Sentiment-Aware Portfolio Optimization: CVaR-Based Diversification With Deep Reinforcement Learning. *IEEE Access*, *13*, 182690–182704. https://doi.org/10.1109/access.2025.3624652

Mantshimuli, L. A., & Mwamba, J. W. M. (2025). Enhancing portfolio optimization with multi-LLM sentiment aggregation: A Black-Litterman integration approach. *Investment Management and Financial Innovations*, *22*(3), 213–226. https://doi.org/10.21511/imfi.22(3).2025.16

Nageen, V. S. (2022). AI-Driven Portfolio Optimization for ESG Investment. *International Journal of Emerging Research in Engineering and Technology*, *3*, 128–132. https://doi.org/10.63282/3050-922x.ijeret-v3i4p114

Quimbayo, C. A. Z., & Camacho, B. L. (2025). Downside Risk Measures and ESG Factors in Optimal Portfolio Construction: Evidence from European Equity Markets. *Economics*, *13*(4), 5–17. https://doi.org/10.2478/eoik-2025-0082

Shah, A., Mehendale, S., & Kanthi, S. (2024). Efficacy of Large Language Models in Systematic Reviews. In *arXiv (Cornell University)*. Cornell University. https://doi.org/10.48550/arxiv.2408.04646

Sokolov, A., Caverly, K., Mostovoy, J., Fahoum, T., & Seco, L. (2021). Weak Supervision and Black-Litterman for Automated ESG Portfolio Construction. *The Journal of Financial Data Science*, *3*(3), 129–138. https://doi.org/10.3905/jfds.2021.1.070

Sultana, N., & Zeya, F. (2025). Advancing Firm-Specific ESG Sentiment Analysis: A Machine Learning Approach to Clustering, Prediction, and Financial Performance Implications. *Business Strategy & Development*, *8*(3). https://doi.org/10.1002/bsd2.70150

Taheripour, E., Sadjadi, S. J., & Amiri, B. (2025). A multi-criteria approach to ESG-based portfolio optimization incorporating historical performance, forward-looking insights, and credibilistic CVaR: a case study on the DJIA. *Scientific Reports*, *15*(1), 39088–39088. https://doi.org/10.1038/s41598-025-24242-x

Tavasoli, A., Sharbaf, M., & Madani, S. (2025). Responsible Innovation: A Strategic Framework for Financial LLM Integration. In *ArXiv.org*. https://doi.org/10.48550/arxiv.2504.02165

Xiao, Y., Wu, D., Li, Z., & Xu, H. (2023). On the prediction of stock price crash risk using textual sentiment of management statement. *China Finance Review International*, *14*(2), 310–331. https://doi.org/10.1108/cfri-12-2022-0250