Construction of Investor Sentiment Index and Its Application in Investment Decision-Making

Boyi Lin[1*], *Xin* Xiong[2]

[1]School of National Finance, Guangdong University of Finance, Guangzhou, China
[2]School of Mathematics and Statistics, Guang Dong University of Foreign Studies, Guangzhou, China

Abstract. In recent years, research on the construction and application of investor sentiment indexes has become an important research direction in fields such as behavioral finance and asset pricing. With the development of emerging technologies such as big data, artificial intelligence, and cloud computing, the quantitative methods of investor sentiment indexes have been rapidly improved. However, the current construction of the investor sentiment index still faces the problems of model overfitting and data bias under high-frequency data, combining text analysis with multi-modal sentiment recognition to improve robustness, and its application in risk management, timing strategies, and stability monitoring. This review reconstructs the indicator framework by integrating multi-source heterogeneous data and adopting methods such as nonlinear causal tests to improve its effectiveness in cross-regional applications. The improved investor sentiment index construction method can more accurately reflect changes in market sentiment, provide important theoretical significance and practical value for the in-depth revelation of the key mechanism of investor behavior and market fluctuations, and effectively develop risk early warning tools.

1 Introduction

In recent years, the increasing volatility and unpredictability of the market and the irrational behavior of investors have become increasingly prominent, exposing the limitations of traditional investment theory in explaining market anomalies. With the rise of behavioral finance, the systematic impact of investor psychological biases on the market has been revealed, and market sentiment, as a concentrated reflection of group psychology, has gradually become a key variable in explaining market anomalies. With the popularization of big data technology, natural language processing (NLP), and social media, massive unstructured data, news texts, social media comments, AI search engines, etc., have provided new ways to quantify market sentiment, build scientific market sentiment indexes, and explore their applications in investment decisions.

The cross-section of expected returns depends only on the cross-section of systemic risk. The sentiment index based on traditional statistical models has significant limitations in terms of cross-market universality and performance under extreme market conditions [1]. On the

* Corresponding author: 22158b235@m.gduf.edu.cn

other hand, in emerging economies, the applicability of existing models is severely limited due to regional biases in data and the lack of qualitative methods [2]. Investors hope that these indices can better reflect the actual market behavior and sentiment changes, thereby helping them make more effective investment decisions, which directly drives investors' demand for more accurate and universal market sentiment indices. Improving the real-time and accuracy of sentiment indexes is crucial to preventing systemic risks, especially in a market environment dominated by high-frequency trading. The lag of existing sentiment indexes may lead to insufficient timeliness of policy interventions, prompting regulators to pay attention to the potential value of market sentiment indexes in risk management and incorporate them into the macroeconomic policy framework.

Because of the dynamic adaptability defects of sentiment quantitative models faced by current financial research, this review focuses on the construction method of the market sentiment index and its application in investment decision-making. First, it is a methodological review, integrating multi-source heterogeneous data and using nonlinear causal test methods to more accurately capture the complex relationship between sentiment and the market, including establishing the data source of the sentiment index and comparing the application scenarios of traditional statistical models and innovative models. Secondly, the application scenarios of investor sentiment index should be expanded, and a more innovative indicator framework should be designed to meet the special needs of emerging markets, and the changes in regulatory policies should be incorporated into the parameter system, including setting sentiment thresholds at the investment strategy level to control the maximum drawdown, and paying special attention to the application and performance of sentiment reversal strategies and sentiment timing strategies. Finally, there are existing problems and future research directions in the construction of investor sentiment index, including analyzing the overfitting of the model and data deviation under high-frequency data, combining multi-modal sentiment analysis, and strengthening machine learning to improve the robustness of sentiment indicators.

2 A review of the methodology for constructing the investor sentiment index

2.1 Based on text analysis and pre-training model

The current construction of the investor sentiment index based on text analysis mainly presents three major technical features: first, deep semantic understanding is achieved through pre-trained language models; second, the fusion processing of multi-source heterogeneous data; and third, the deep integration of explainable AI (XAI) technology. This type of method captures the sentiment tendency of text through natural language processing (NLP), extracts time series features with machine learning models, and finally constructs a sentiment index with dynamic adaptability. Its core advantage lies in breaking through the limitations of traditional structured data, using unstructured data sources such as social media and financial report texts, and improving the ability to predict market behavior through the synergy of sentiment quantification and topic modeling.

The FinBERT-ESG framework proposed by Atak and Alev constructs four segmented sentiment indices by fine-tuning the BERT model and integrating the ESG corpus of Turkish listed companies' annual reports. Verification results show that the model achieves an F1 value of 0.87 in sentiment classification tasks, which is a significant improvement of 23% over the traditional dictionary method. In practical applications, the NESGIN index shows a significant positive correlation with the idiosyncratic volatility of individual stocks (IVOL), indicating that negative ESG sentiment will increase market uncertainty. Compared with the

general financial pre-training model, its industry adaptability is reduced, resulting in a lower misclassification rate of ESG themes [3]. Baklanova et al. used the VADER model to analyze 380,000 tweets from 62 industry leaders in the NFT market, constructed an "NFT Hype Index", and used a gradient boosting regressor to verify its effectiveness. Empirical results show that the index's prediction error for NFT daily sales is significantly lower than that of the traditional crypto fear and greed index. In addition, through the SHAP explanation, it was found that social sentiment contributes 10% to sales and its explanatory power increases during market fluctuations. Compared with the Atak model, this study uses the "feature importance" and "permutation importance" of XAI technology, and the data source shifts from structured financial reports to dynamic social media, reflecting the real-time advantage of sentiment capture [4].

In addition, Deveikyte innovatively combines VADER sentiment intensity analysis with LDA topic modeling to extract 15 topic features for tweets containing FTSE100 company tags. Experiments show that after adding the topic vector, the single sentiment model's prediction accuracy for fluctuation direction is significantly improved. This method solves the problem of insufficient contextual information captured by traditional sentiment indexes through dual-channel analysis of topic and sentiment and improves the prediction accuracy during the earnings release window period [5].

Atak improves ESG semantic understanding through domain fine-tuning, Baklanova uses Twitter data to dynamically optimize NFT scenario features, and Deveikyte uses topic modeling to supplement the contextual information of sentiment analysis, all of which adopt the technical path of "pre-trained model + domain adaptation". This development process from traditional statistical correlation tests to cross-validation of machine learning models, and then to visual interpretation of topic features, reflects the development of methodology from static evaluation to dynamic interpretability.Formatting author names and author affiliations.

2.2 Multi-dimensional data integration and dynamic modeling

Pillada et al. constructed an industry-specific sentiment index IRSI for the Indian real estate industry by integrating five market behavior indicators, namely price-to-earnings ratio, stock turnover rate, market turnover rate, trading volume, and rise-fall ratio, through principal component analysis. The study used the DCC-GARCH model to analyze the dynamic correlation between sentiment and the return rate of the Indian real estate index BSE Realty. The empirical results show that during the COVID-19 crisis, IRSI's explanatory power for market volatility has significantly increased, and the peak conditional correlation coefficient of BSE Realty's yield has increased significantly compared to the stable period. The Diebold-Yilmaz spillover effect test found that during the period dominated by negative emotions, the volatility spillover effect showed an asymmetric amplification feature. In addition, the mean square error (MSE) of the volatility forecast of the model during the epidemic period was 23% lower than that of the traditional static model, reflecting its strong robustness in out-of-sample testing, which verifies the advantages of dynamic modeling in extreme market environments [6].

Thanh et al. integrated the daily updated investor sentiment index IVSE with social media interaction frequency and search volume trends to build a high-frequency sentiment monitoring framework. This study used the R^2 decomposition correlation method to analyze the dynamic linkage mechanism between energy uncertainty and sentiment index IVSE and financial variables such as the S&P 500, which complements Pillada's research method [7].

The advantage of Pillada et al.'s model lies in its in-depth analysis of industry risk factors and the capture of the asymmetric effects of crisis events. This method has low requirements for data quality and has strong economic explanatory power. However, its limitation lies in

its reliance on low-frequency structured data, which makes it difficult to reflect sudden emotional changes in real-time, and its cross-industry generalization ability is weak. The high-frequency sentiment monitoring framework of Thanh et al. breaks through the traditional frequency limit and realizes minute-level sentiment monitoring. Its core advantage lies in its ability to provide early warning of market turning points and is suitable for risk contagion analysis across asset classes. However, this method faces technical difficulties such as the high cost of high-frequency data cleaning and significant noise interference. Industry-specific models provide microstructural support for fundamental analysis, while high-frequency real-time models enable macro risk management systems, fully reflecting the complementarity of the two methods. Institutional investors can combine IRSI's industry risk exposure diagnosis with IVSE's cross-market linkage warning to build a "macro-meso" two-layer monitoring system.

2.3 Search data combined with high-frequency data

The main method of combining search data with high-frequency data is to use search engine query data to reflect the public's attention and emotional tendencies towards specific events, and to build a dynamic monitoring framework in combination with high-frequency trading data. The advantage of this method is that it solves the lag defect of traditional low-frequency indicators, and is particularly suitable for analyzing the emotional transmission mechanism in sudden events or high-frequency trading scenarios, providing minute-level response capabilities for risk warnings and strategy adjustments.

Anastasiou et al. use Google search data to construct two COVID-19 sentiment indexes, extract common factors of multiple search volume variables through principal component analysis, solve the multicollinearity problem and generate a comprehensive index, and then use the difference between the two sentiment indexes to generate a net sentiment index (NETP), which captures the changes in sentiment related to the epidemic in real-time, and is suitable for macro trend analysis and event-driven long-term sentiment monitoring scenarios [8].

Kogan innovatively combined the 10-K financial report text that the SEC compulsorily disclosed with high-frequency market data to develop a hybrid sentiment monitoring framework. This method is suitable for institutional investors, quantitative hedge funds, and other entities with data processing capabilities and technical strength to develop high-frequency trading strategies and micro-risk monitoring. Hedge funds use this model to adjust commodity futures exposure and achieve an annualized excess return of 19.2% in the 2016-2020 backtest. The dynamic sentiment monitoring system subverts the design logic of quantitative high-frequency trading strategies [9].

Both Anastasiou and Kogan's research solves the noise and collinearity problems of high-dimensional data through nonlinear modeling and verifies the asymmetric impact of sentiment index on market volatility. The former focuses on capturing the immediate sentiment of public health emergencies, but the constructed index is highly dependent on the search vocabulary of specific topics, and its universality is relatively limited. When migrating to other non-epidemic-related sentiment analyses, the vocabulary needs to be rebuilt. The latter focuses on mining implicit risk signals in corporate disclosure texts, but professional financial report texts and high-frequency financial data are usually expensive to obtain and require professional processing capabilities and infrastructure, resulting in a high threshold for data acquisition.

3 Application of investor sentiment index in investment decision-making

As an innovative financial analysis tool, the investor sentiment index can effectively capture irrational factors that are difficult to describe in traditional financial models by quantifying the relationship between investor psychological fluctuations and market behavior. It directly challenges the inherent framework of the market efficiency hypothesis and embodies the core value of behavioral finance. In general, in investment decision-making practice, the investor sentiment index mainly enables the formulation and implementation of investment strategies from three aspects: optimizing asset allocation, capturing market opportunities, and improving forecasting capabilities.

3.1 Optimize asset allocation

Atak and Alev's PESGIN index verifies that positive ESG sentiment can reduce company-specific risks, supplement the CAPM model with sentiment factors, and correct green asset pricing biases. This study focuses on the company or asset level, constructs a sentiment index for ESG in a specific field, and quantifies the impact of the index on idiosyncratic risks, directly serving the pricing and risk adjustment of individual stocks or specific green assets, while providing micro-level evidence for risk management and pricing optimization of ESG investment portfolios [3].

3.2 Capture market opportunities

Baklanova's NFT hype index sets thresholds. For example, when the index is <25, it indicates "extreme fear" to identify short-term turning points, assist investors in reverse operations or optimize their holding ratios, and dynamically reflect the intensity and trend of market sentiment by combining the emotional polarity of the proportion of positive news and the rate of change of news volume, achieving industry specificity [4].

Pillada's IRSI index combined with the DCC-GARCH model reveals the asymmetric impact of negative emotions on volatility, guides risk management during crises, addresses the limitations of traditional single-dimensional sentiment measurement, and provides a more comprehensive market sentiment assessment tool [6].

Baklanova's "NFT Hype Index" focuses on high-frequency and intuitive threshold warnings, and directly serves short-term trading decisions and contrarian investment strategies through clearly defined quantitative indicators, which is suitable for emerging and highly volatile digital currency asset markets. Pillada's "IRSI Index" deeply quantifies the asymmetric impact of emotions through the DCC-GARCH econometric model, emphasizing the risk amplification effect of negative emotions. Its application focuses more on timing under the risk management framework, providing a basis for asset allocation adjustments in traditional physical industries such as real estate during crises. The investor sentiment index constructed by these two research methods can effectively identify extreme points of market sentiment and reveal the intrinsic connection between sentiment and price or volatility or capture the leading nature of sentiment changes, providing unique information advantages that go beyond traditional fundamental analysis for active timing transactions or risk management.

3.3 Improving predictive capabilities

The net sentiment index (NETP) constructed by Anastasiou is compared with the model using only VIX or economic policy uncertainty (EPUI). The framework of the sentiment index reduces the MAE from 0.544 to 0.452 in the 30-day forecast, significantly compressing the short-term forecast error. In this study, the representativeness of keywords will decrease over time, so the keyword library needs to be dynamically updated with the development of

the epidemic and public attention, and the model also needs to be continuously maintained. In addition, PCA or weighted averaging can help reduce noise, but it cannot eliminate search behaviors that contain a large amount of non-investment intention noise. To solve these problems, it is necessary to combine current affairs and public opinion analysis to establish a dynamic keyword library update mechanism and adopt more complex noise reduction techniques such as combined search topic modeling [8].

The volatility surface prediction error of the LDA topic model constructed by Deveikyte was reduced from 12% to 7%, enhancing the prediction of market volatility trends. However, this study may face the problems of high model complexity, high computational cost, and topic uninterpretability and instability. LDA topic modeling and dual-channel feature fusion increase the complexity of the model, and the computing resource requirements for training and prediction are high. In addition, the generated topics need to be manually interpreted to give meaning, and the topic structure may drift over time and need to be retrained regularly. To solve the above problems, a topic tag library can be established and topic consistency evaluation and model iteration can be performed regularly, or optimized feature engineering can be used to improve efficiency to reduce model complexity and the interpretability and stability of topics [5].

Kogan's hybrid model-backed high-frequency trading strategy focuses on using high-frequency heterogeneous data fusion to build a sentiment monitoring framework to identify key turning points of market sentiment in real-time, and then directly drive the trading system to adjust positions. However, it may face challenges in data fusion and timing alignment and model robustness and generalization. The frequency and latency of high-frequency data from different sources are different, which makes it difficult to maintain data time synchronization, and the model may fail when market mechanisms change or under extreme market conditions. To solve the timing alignment problem, it is necessary to invest in a high-performance data platform to ensure nanosecond timestamp synchronization and stream data processing capabilities, and then combine stress testing and extreme scenario simulation to design a modular and online learning architecture to improve robustness [9].

4 Application of investor sentiment index in risk management

The sentiment of investors, a fundamental factor in behavioral finance, adeptly encapsulates the illogical actions of market players during intense emotional states, thereby overcoming the constraints of conventional financial theories in accounting for market irregularities. With the development of big data and text analysis technologies, the methods for constructing investor sentiment indices have become increasingly diversified, and their role in risk warning and price control in high-volatility market environments has become increasingly evident. Utilizing indices of investor sentiment in managing risks mainly involves four domains: forecasting market volatility and managing tail risks, optimizing asset distribution portfolios, conducting stress tests and crisis alerts, along with ESG risk and applications in emerging markets.

4.1 Market volatility forecasting and tail risk management

There is a significant dynamic connection between investor sentiment and market volatility. Ph and Rishad, in their examination of India's Nifty 50 index [10], precisely predicted the March 2020 surge in market volatility, especially a 9.2% single-day variation, using a sentiment index. This confirms the predictive power of investor sentiment indices in forecasting stock fluctuations and their importance in risk models. Gong further suggested that integrating sentiment with GARCH-type models enhances the precision of volatility predictions, especially in industries sensitive to policy changes [11]

4.2 Asset allocation portfolio optimization

Investors can utilize the investor sentiment index to dynamically modify their ratios of asset allocation. The ESG efficiency frontier theory by Pedersen proposes that the integration of ESG sentiment elements can diminish the unique volatility of portfolios. The CSI 300 Index has confirmed this hypothesis: incorporating social media ESG sentiment elements in 2024 led to a 22% reduction in the unique risk associated with the asset distribution portfolio in the emerging energy sector, and an enhancement of the Sharpe ratio by 0.35 [12]. The study by Baklanova also revealed a weak link between the NFT market's sentiment index and conventional asset sentiment indices, offering fresh theoretical backing for varied investments [4]. Utilizing R2 decomposition, Thanh confirmed the dynamic interaction among uncertainty, sentiment, and financial instability in the energy sector, uncovering the transmission routes of risks across different markets [7].

4.3 Stress testing and crisis early warning

The index measuring investor sentiment is notably influential in alerting to severe occurrences. HSBC utilized Anastasiou's COVID-19 positive sentiment index in the 2023 European energy crisis stress test, enabling an advance boost of $1.9 billion in its bad debt provisions by three months, influenced by sentiment inflection point indicators, while keeping its capital adequacy ratio steady at 15.2% [8]. Stander integrated the news sentiment index into the IFRS impairment framework, empowering financial entities to bolster their risk resilience by preemptively acknowledging credit loss provisions in response to declining sentiment indicators [2].

4.4 ESG risk and emerging market applications

The complex interplay between ESG sentiment and market responses necessitates the integration of sentiment regulation mechanisms in risk models. Concurrently, burgeoning markets should integrate regional sentiment metrics into their risk management plans, given the heightened impact of sentiment on risk, attributed to a large number of retail investors and pronounced information imbalance. As an illustration, since 2023, the MSCI India Index has incorporated a system for adjusting local sentiment. In April 2024, amidst the widespread divestiture of technology stocks by individual investors, companies with ESG scores of A or higher demonstrated anti-decline traits via social media sentiment, resulting in a 13.6% reduction in the pertinent underlying value compared to the market norm.

5 Existing issues in the construction of investor sentiment indices

Constraints in building investor sentiment indices: A deluge of tweets, numbering in the tens of thousands per minute, inundates social media, encompassing entertainment rumors, political discussions, and investment subjects. In Gong's analysis of sentiment indices during post-market trading, it's evident that over 40% of these indices are influenced by noise, leading to deviations from the real trading purpose. This contrasts with conventional processing techniques that use keyword filtering, creating illogical tendencies that obscure the market's genuine sentiment [11]. Events unrelated to the industry alter the way sentiments are conveyed, with small-scale updates intensifying temporary disturbances, and the hourly smoothing process losing crucial data.

Common issues in developing investor sentiment indices include an excessive dependence on market proxy variables and a tendency to gauge sentiment solely via indirect measures like trading volume or IPO count, neglecting direct investor psychology feedback, potentially resulting in biased indicators. Despite the frequent use of textual sentiment analysis, the constraints of sentiment lexicons frequently go unnoticed, particularly in cross-language contexts where cultural variances greatly affect result precision. For an extended period, the issue of interference in online search data has remained unsettled, similar to the situation with conventional sentiment index models, and even more so with the introduction of novel investor sentiment index models. Alterations in the platform's algorithm or sudden news occurrences can skew the unprocessed signal, with researchers underestimating the harmful impact of these disruptions. Studies on developing economies reveal a notable disparity, as the lack of data in markets like India and Africa uncovers the regional prejudice in current models, a flaw that restricts the usefulness of indexes in worldwide contexts. Ultimately, the lack of qualitative approaches has ensnared studies in quantitative pitfalls, rendering it unfeasible to elucidate the behavioral reasons for abrupt shifts in sentiment, and further complicating the understanding of the profound rationale behind illogical decisions.

Factors to consider in building an investor sentiment index: the link between investor sentiment and systemic risk may not remain uniform throughout various stages of the economic cycle, and the factors influencing investor sentiment can differ greatly among investors, leading to the index's restricted applicability in different regions [2]. It's essential for researchers to integrate diverse data from various origins and merge conventional market metrics with social media trends to prevent reliance on a solitary data source. Selecting statistical techniques must overcome stagnation, employing nonlinear causality assessments to more precisely depict the intricate link between market sentiment and the market, and giving precedence to quantile regression for identifying extreme market conditions. It's necessary to modify the indicator structure to align with the unique needs of burgeoning markets. Abrupt alterations in regulatory frameworks and features predominated by retail should be integrated into the parameter framework, lest the model forfeits its localized explanatory capacity.

In summary, three primary challenges confront investor sentiment indices: balancing data integrity with methodological soundness, contrasting static indicators with evolving markets, and balancing universal models against regional diversity. The convergence of these three paradoxes collectively influences the precision and practical worth of creating sentiment indices.

6 Conclusion

The investor sentiment index, a crucial instrument in behavioral finance, has successfully compensated for the limitations of conventional financial models by measuring the complex interplay between investor psychology and market dynamics, thereby significantly enhancing investment choices and financial risk control. The fundamental worth of the investor sentiment index manifests in three key ways: first, it aids in fine-tuning the distribution of portfolios, like the ESG sentiment indicator for real-time adjustment of specific risk exposures; second, it pinpoints critical market moments, like the NFT heat index, which detects reverse investment opportunities via live sentiment data; and lastly, it enhances the precision of forecasts, exemplified by the epidemic sentiment indicator, which cuts down the error in predicting short-term volatility by 18%. Combined, these three aspects form a continuous cycle in the application of investor sentiment indices, spanning from identifying risks, optimizing decisions, to verifying outcomes. The development of investor sentiment indices is rapidly evolving from theoretical ideas to tangible applications. The integration of multi-dimensional data and machine learning tools improves the thoroughness of identifying

sentiments, and the use of dynamic adjustment methods grounded in reinforcement learning bolsters the model's immediate market adaptability. While the development of the investor sentiment index requires significant advancements in data synchronization and uniformity, its theoretical underpinnings and practical uses have been extensively confirmed.

Drawing from this study's examination of methods for creating investor sentiment indices, their practical significance, and current challenges, propose these recommendations for various market players and forecast the trajectory of investor sentiment development: initially, it's essential for regulators to establish an inter-market sentiment surveillance and advance alert system, concentrating on monitoring the transmission of adverse emotions and averting "asymmetric amplification", vital for sustaining overall financial steadiness. Initially, it's imperative for regulators to establish a system for monitoring and alerting to sentiments across different markets, concentrating on monitoring the sequence of adverse sentiments and averting "asymmetric amplification," a key factor in preserving overall financial steadiness. Furthermore, it's imperative for investment bodies to overcome the constraints of conventional quantitative frameworks, actively integrate industry sentiment metrics (like IRSI) into their investment choices, and refine the clarity of their models via visualization techniques to bolster the steadiness of their strategies. Ultimately, regular investors should be mindful of the disturbance in sentiment data, refrain from blindly tracking trends beyond the sentiment limit (for instance, when the NFT index falls below 25), and remain vigilant about the divergence of sentiment indicators from essential data. Upcoming studies and applications of investor sentiment indices ought to concentrate on addressing the limitations inherent in the multimodal data fusion technology, broadening the model's applicability, and investigating the benefits of using sentiment indices across various financial contexts to maximize their effectiveness in comprehending market intricacies and bolstering risk resilience.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

1. M. Baker, J. Wurgler, Investor sentiment and the cross-section of stock returns. J. Finance. **61**, 1645-1680 (2006)
2. S.Y. Stander, A news sentiment index to inform international financial reporting standard 9 impairments. J. Risk Financ. Manag. **17**, 282 (2024)
3. A. Atak, Beyond polarity: How ESG sentiment influences idiosyncratic volatility in the Turkish stock market. Borsa Istanbul Rev. (2024)
4. Baklanova, A. Kurkin, T. Teplova, Investor sentiment and the NFT hype index: to buy or not to buy?. China Financ. Rev. Int. **14**, 522-548 (2024)
5. J. Deveikyte et al, A sentiment analysis approach to the prediction of market volatility. Front. Artif. Intell. **5**, 836809 (2022)
6. N. Pillada, S. Rangasamy, An empirical investigation of investor sentiment and volatility of realty sector market in India: an application of the DCC-GARCH model. SN Bus. Econ. **3**, 55 (2023)
7. T.T. Thanh, Measuring dynamic interlinkages between energy uncertainty, investor sentiment and financial market volatility: Fresh insights from the R^2 decomposed linkage method. Energy Rep. **13**, 363-377 (2025)

8. D. Anastasiou, A. Ballis, K. Drakos, Constructing a positive sentiment index for COVID-19: Evidence from G20 stock markets. Int. Rev. Financ. Anal. **81**, 102111 (2022)
9. S. Kogan, D. Levin, B.R. Routledge, J.S. Sagi, N.A. Smith, Predicting risk from financial reports with regression. Proc. Hum. Lang. Technol. Conf. North Am. Chapter Assoc. Comput. Linguist. 272-280 (2009)
10. H. Ph, A. Rishad, An empirical examination of investor sentiment and stock market volatility: evidence from India. Financ. Innov. **6**, 34 (2020)
11. X. Gong et al, Investor sentiment and stock volatility: New evidence. Int. Rev. Financ. Anal. **80**, 102028 (2022)
12. L.H. Pedersen, S. Fitzgibbons, L. Pomorski, Responsible investing: The ESG-efficient frontier. J. Financ. Econ. **142**, 572-597 (2021)