Time-Varying Analysis of Investor Sentiment, Stock Market Liquidity, and Volatility: Based on Deep Learning BERT Model and TVP-VAR Model

Authors: Zhang Chenrui, Wu Xinyi, Deng Hailu, Zhang Huiwei, Zhang Chenrui

Date: 2022-05-10T09:51:09+00:00

Abstract

Based on comment data from Eastmoney.com' s Shenzhen Component Index stock forum between January 1, 2018 and December 31, 2019, this study extracts investor sentiment embedded therein using a deep learning BERT model, and employs a TVP-VAR model to investigate the time-varying dynamic relationships among investor sentiment, stock market liquidity, and volatility. The empirical results demonstrate that shocks from investor sentiment to stock market liquidity and volatility are more pronounced, while the reverse effects, although relatively weaker, exhibit more significant variation with market state changes. Moreover, short-term responses are more significant than medium- to long-term responses in all cases, and the impacts display asymmetry, with shocks being more intense during market downturns.

Full Text

A Time-Varying Study of Investor Sentiment, Stock Market Liquidity and Volatility—Based on Deep Learning BERT Model and TVP-VAR Model

ZHANG Chenrui, WU Xinyi, DENG Hailu, ZHANG Huiwei
School of Economics and Finance, South China University of Technology, Guangzhou 510006, China

Abstract

Based on comment data from the Shenzhen Stock Index BBS on Eastmoney.com between January 1, 2018 and December 31, 2019, this paper extracts investor sentiment using a deep learning BERT model and investigates the time-varying linkage relationships among investor sentiment, stock market liquidity, and volatility through a TVP-VAR model. Experimental results demonstrate that investor sentiment exerts a stronger impact on stock market liquidity and volatility, while the reverse effects, though relatively weaker, vary more significantly with market conditions. Furthermore, short-term responses are more pronounced than medium- and long-term responses in all cases, and the effects exhibit asymmetry, with shocks being more intense during market downturns.

Keywords: BERT model; TVP-VAR model; Shenzhen Stock Index; Investor sentiment; Stock market liquidity; Stock market volatility

This paper is one of the research results funded by the National College Students' Innovation and Entrepreneurship Training Program (Project No. 202110561076).

1. Introduction

The relative stability of stock markets forms the foundation for their development and growth, facilitating participation by domestic and foreign investors while helping to prevent financial crises. Abnormal market fluctuations erode investor confidence, distort price signals that serve as asset allocation benchmarks, and increase systemic risk across the financial system. Compared to mature markets in Europe and the United States, China' s stock market exhibits higher volatility. The dramatic boom-and-bust cycle of the Shanghai Composite Index between 2014 and 2015 prompted reflection among market participants and researchers. The underlying causes can be explained from two perspectives: behavioral finance, which emphasizes the important influence of investor behavior and sentiment, and fractal market theory, which highlights the role of market liquidity factors. Evidently, the dynamic evolution of investor sentiment and market liquidity has become a crucial determinant of market volatility. With the increasing internetization and democratization of stock trading, investigating the interactive mechanisms among investor sentiment, stock market liquidity, and volatility has assumed growing importance.

This paper makes two expected marginal contributions. First, we fine-tune BERT, one of the state-of-the-art deep learning models for text processing, to develop a more suitable and reliable BERT model for assessing sentiment in Chinese investor text, which we release as open source for future researchers to use and improve. Second, while existing literature has extensively examined the relationship between investor sentiment and either stock market liquidity or volatility individually, few studies have investigated the dynamic linkages among all three variables. This paper addresses this gap by employing a TVP-VAR model to enrich research in this domain.

2. Literature Review

With the internetization and democratization of stock market trading, examining how investor sentiment influences markets has become increasingly prominent in behavioral finance research. Our literature review focuses on two main areas: "measurement of investor sentiment" and "the linkage between investor sentiment and stock markets," which we further refine into specific research directions.

2.1 Investor Sentiment

Direct survey methods consume substantial time and resources while suffering from limited sample sizes and significant biases, making them less common in recent research. The market proxy indicator approach constructs sentiment proxies by selecting variables such as closed-end fund discounts and turnover rates, then applying principal component analysis. This method remains widely used due to its ease of implementation and significant research results, though it can only indirectly reflect investor sentiment. The dictionary method derives weighted sentiment indicators by matching sentiment words in financial texts against dictionaries. While this approach directly captures investor sentiment and has greatly facilitated financial text analysis research, its accuracy heavily depends on dictionary quality, and constructing dictionaries requires cumbersome manual work.

With advances in computing power and continuous updates to natural language processing algorithms, many scholars have begun employing models such as Support Vector Machines (SVM), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) to directly extract investor sentiment from financial texts using machine learning methods. The accuracy of such research continues to improve with model enhancements. Huicheng Liu utilized bidirectional LSTM to encode news texts and capture contextual sentiment signals, while Yao Qin developed a prediction method based on a dual-stage recurrent neural network time series model to study investor sentiment. Currently, machine learning represents the cutting-edge direction in investor sentiment measurement, prompting our adoption of this approach.

Sentiment classification is a subtask in natural language processing that essentially involves analyzing, processing, summarizing, and reasoning about subjective texts with emotional coloring. Since this study focuses on analyzing sentiment in stock BBS comments, sentiment classification facilitates reasonable monitoring of stock market public opinion.

With the development of computer hardware, BERT—a deeper and more complex deep learning model—was designed. Research shows BERT has been primarily applied in media and political commentary studies. For instance, Guoshuai Zhang et al. (2020) used BERT language models to analyze news from *The New York Times* to predict short-term U.S. policy changes, while Jiaqi Hou et al. (2020) proposed a new BERT-att model combining transformer bidirectional encoders, demonstrating its effectiveness for security risk assessment. However, few scholars have employed BERT in financial research, and even fewer have released their fine-tuned sentiment classification models as open-source tools. This study aims to fill this gap by fine-tuning BERT, one of the most advanced deep learning models for text processing, to create a more reliable model for judging sentiment in Chinese investor text.

2.2 Market Liquidity

Stock market liquidity refers to the ability to execute large volumes of trades quickly and at low cost without causing significant price changes. Liquidity encompasses several dimensions: immediacy of execution, low transaction costs, large trade quantities, and minimal price impact. Numerous liquidity indicators exist in current research, allowing researchers to construct different measures based on their specific needs. Yang Chaojun (2008) proposed new liquidity indicators building on the Amivest liquidity ratio and Hui-Heubel liquidity ratio concepts. Zeng Zhijian and Luo Changqing (2008) used turnover rate as a liquidity variable to study the linkage between two markets' liquidity using one year of Shanghai Stock Exchange data, finding no lead-lag relationship between monthly liquidity in stock and bond markets. Wang Yintian et al. (2010) investigated liquidity spillover effects between stock and bond markets, constructing price impact-based liquidity measures. Han Jinxiao et al. (2017) demonstrated that spread-based indicators from Crowin and Schultz (2012) effectively measure Chinese stock market liquidity. In stock markets, turnover rate remains the simplest liquidity indicator.

2.3 Market Volatility

Stock market volatility refers to price changes resulting from shifts in market expectations due to economic, policy, and market participant factors. Stock market fluctuations are influenced by fundamentals, funding conditions, investor sentiment, and other factors. Volatility is a crucial characteristic of financial markets, reflecting how financial asset returns change over time. Common volatility measurement methods in existing literature include variance or standard deviation of returns, conditional variance of returns, and price oscillation ranges. Engle (1982) distinguished between variance and conditional variance, establishing the ARCH model and pioneering new approaches to volatility research. Bollerslev (1986) subsequently proposed the GARCH model. Huang Hainan and Zhong Wei (2007) suggested using GARCH-type models to predict volatility in the Shanghai Composite Index returns. Dai Wen (2018) empirically analyzed ST stock return volatility using ARMA-GARCH model families. Existing literature on stock market volatility has revealed clustering, persistence, and asymmetry in security price fluctuations, confirming the applicability of ARCH-family models in modeling security price volatility. With various domestic and international models available for consideration, this study constructs stock market volatility indicators based on realized volatility from the CSMAR database.

2.4 Research on Investor Sentiment, Market Liquidity, and Market Volatility

Regarding the impact of investor sentiment on liquidity, Chen et al. (2009) studied the connections among investor sentiment, market liquidity, and trading behavior, finding that when investor sentiment turns pessimistic, more investors choose to sell, reducing net buying volume and consequently lowering market liquidity, with individual investors reacting more strongly to sentiment changes than institutional investors. Li Chunhong and Peng Guangyu (2011) used Granger causality tests to demonstrate that investor sentiment Granger-causes stock liquidity, indirectly affecting market capitalization rates and economic growth levels. Wang Danfeng and Liang Dan (2012) found that changes in investor sentiment increase market liquidity and significantly affect liquidity premiums in expected stock returns. Liu Xiaoxing et al. (2016) discovered that investor sentiment positively influences market liquidity, though this effect weakens as investors' cognitive ability regarding market information improves. Research also reveals differential impacts of positive versus negative sentiment on stock market liquidity. Yang Xiaolan et al. (2016) found that the positive effect of local attention on trading volume is stronger under positive sentiment than under negative sentiment. Shi Guangping et al. (2016) used a TVP-SV-SVAR model to analyze how investor sentiment and market liquidity affect stock market bubbles, finding that optimistic sentiment during bull markets has a more significant impact than pessimistic sentiment during bear markets. These studies examine investor sentiment's effect on market liquidity from various perspectives, revealing differentiated impacts across sentiment types and market conditions.

Concerning investor sentiment's impact on stock market volatility, Lee et al. (2002) found that sentiment significantly affects market volatility, with positive sentiment reducing volatility and negative sentiment increasing it. Zhang Zongxin and Wang Hailiang (2013) constructed an investor sentiment index using principal component analysis to analyze its relationship with stock market volatility, discovering that sentiment has a significant positive effect on market returns and volatility—higher sentiment levels lead to greater return fluctuations, larger deviations from intrinsic value, and increased market volatility. Hu Changsheng and Chi Yangchun (2013) found that both rational and irrational investors contribute to market volatility, with sentiment causing markets to deviate from rational frameworks and generating abnormal fluctuations. Lu Jianqing and Chen Mingzhu (2013) proposed the "expected sentiment spillover effect" hypothesis, suggesting that micro-level investor psychological expectations trigger herd behavior, inducing macro-level "stock market resonance effects" that amplify volatility. Ba Shusong (2016) noted that margin trading amplifies investor sentiment and exacerbates market volatility. These studies confirm from various angles that investor sentiment constitutes a systematic risk factor and represents an important driver of stock market return volatility.

3. Data Processing and Model Construction

3.1 Data Selection

To ensure our research object provides a comprehensive overview of the entire A-share market and accurately reflects actual investor sentiment conditions, we select Eastmoney.com stock BBS as our text data source and use realized volatility ( ) and turnover rate ( ) as variables to study the impact of investor sentiment.

3.1.1 Text Data Selection We crawled BBS data from Eastmoney.com between January 1, 2018 and December 31, 2019, comparing data across multiple individual stocks and index BBS platforms before ultimately selecting the Shenzhen Stock Index BBS. Compared to the Shanghai Composite Index and other index BBS data, the Shenzhen Stock Index BBS offers appropriate data volume, high user engagement, and strong representativeness. Moreover, the Shenzhen Stock Index comprises 500 high-quality listed companies of certain scale on the Shenzhen Stock Exchange, covering various industries, thus effectively reflecting overall stock market conditions. Following standard text analysis procedures, we removed duplicate data, tables, HTML tags, and other non-text elements.

As stock market capitalization constantly changes, so do constituent stocks. Table 1 shows partial constituent stocks of the Shenzhen Stock Index as of January 1, 2018. However, constituent stocks changed minimally during our study period, allowing us to ignore the impact of such changes. Additionally, our daily BBS crawling captured information including timestamps, comment content, read counts, and comment numbers, as illustrated in Table 2.

Table 3 presents descriptive statistics showing that the average daily comment title length is 62.7 characters with a right-skewed distribution. Unlike traditional (relatively short) comments, some longer posts in BBS data are typically copied and pasted from other sources such as news reports and analytical briefings. To mitigate potential outlier effects, we retained only comments with fewer than 150 Chinese characters. Furthermore, considering the varying daily comment volumes and the influence of different read counts, we selected the top 50 daily comments by read count for our analysis.

3.1.2 Stock Trading Data Selection This study uses the Shenzhen Stock Index as the research object, analyzing its daily trading data including total market capitalization, tradable market capitalization, total shares outstanding, tradable shares, turnover rate, price-earnings ratio, and price-to-book ratio. We obtained this data through the open-source Python API Tushare, which returns results as Pandas.DataFrame objects. Table 4 shows that both market volatility and turnover rate average approximately 1.2%, with right-skewed distributions.

Market volatility is particularly pronounced, with a range of approximately 10% between maximum and minimum values.

3.2 Model Construction

3.2.1 Deep Learning Model Design Based on Financial Corpus In the deep learning era, NLP pretraining widely employs word embeddings, where trained words are converted into word vectors as neural network input layers. Model performance largely depends on training set size, with larger corpora producing better word vectors. However, word vector training ignores contextual meaning, assigning identical vectors to polysemous words. To address this limitation, Devlin et al. (2019) proposed the pretrained language model BERT in 2018, which achieved state-of-the-art results on 11 NLP tasks, representing a major advance in the field.

As shown in Figure 1, the BERT model structure takes input characters 1, 2,··· , , which pass through a bidirectional Transformer feature extractor to obtain text features and output corresponding vectors 1, 2,···, . Figure 2 illustrates BERT' s training approach. As a model where all layers can incorporate contextual semantics, BERT' s input comprises three vectors: token embeddings, segment embeddings, and position embeddings. BERT employs the Masked Language Model (MLM), which masks certain characters (similar to fill-in-the-blank tasks) and predicts the masked tokens through iterative training to achieve contextual understanding.

BERT is fundamentally a two-stage NLP model. The first stage, Pre-training, trains a language model using existing unlabeled corpora. To improve prediction accuracy, we adopt the Chinese-BERT-wwm pretrained model released by the Harbin Institute of Technology and iFLYTEK Joint Laboratory, which uses whole word masking. After pretraining, BERT can be applied to financial entity sentiment recognition tasks as output for subsequent networks. Given the unique characteristics of financial texts, we perform second-stage Fine-tuning on a local GPU using sentiment-labeled financial texts to train a more accurate domain-specific model.

Model Training: We select the chinese_{roberta}*{wwm}*{large}_{ext}L-24_H-1024_A-16 pretrained model (24-layer, 1024-hidden, 16-heads) released by the HIT-iFLYTEK Joint Laboratory, which employs 24 Transformer layers with hidden dimension 1024, 16 attention heads, and total model size of 330MB. Training uses batch size 16, learning rate 2e-5, and maximum sequence length 128. We employ three-class sentiment labeling: 0 for negative, 1 for neutral, and 2 for positive. From our crawled data spanning January 1, 2018 to December 30, 2019, we randomly selected 4,000 comments for manual sentiment labeling, splitting them into training and test sets at an 8:2 ratio. Table 5 shows sample processed data including sentiment labels and probability scores.

3.2.2 Investor Sentiment Indicator Based on Stock Reviews Following Antweiler and Frank (2004), we construct a bullishness indicator based on BBS post classification:

= (−) / (+ +) (1)

In Equation (1), represents the weighted sum of message type { , , } over period ( ), where denotes positive sentiment, negative sentiment, and neutral sentiment. The indicator is a dummy variable equal to 1 if a message belongs to type and 0 otherwise. When all weights equal 1, this represents the total message count of type during period ( ). The sentiment indicator ranges between -1 and 1, expressing the relative bullishness of investors independent of total post volume. As shown in Figure 3 and Table 6, negative sentiment predominates, particularly concentrated in 2018, consistent with that year' s downward market trend.

4. Empirical Analysis

4.1 Stationarity Test

We conduct stationarity tests on three indicators: investor sentiment ( ), realized volatility ( ), and turnover rate ( ). Results show that at the 1% significance level, investor sentiment ( ) and realized volatility ( ) are integrated of order zero, while the turnover rate ( ) is integrated of order one.

4.2 TVP-VAR Model Specification

The TVP-VAR model assumes time-varying coefficients and error term variances that follow first-order random walk processes. This time-varying nature enables the model to better capture potential nonlinear, gradual, or structural changes in the dynamic relationships between investor sentiment and stock market fluctuations. However, time variation also makes traditional likelihood-based estimation unreliable. Nakajima (2011) proposed using Markov Chain Monte Carlo (MCMC) algorithms to overcome these limitations. The approach involves first specifying prior distributions for time-varying parameters based on experience, then iteratively drawing random samples of state vectors through MCMC. When certain conditions are met, the Markov chain' s state transition matrix converges to a stable probability distribution, allowing selection of post-convergence samples from the stationary distribution to effectively simulate the conditional posterior distributions of time-varying parameters and enable statistical inference.

Based on comprehensive evaluation of six criteria (LogL, LR, FPE, AIC, SC, HQ), we determine the optimal lag order for the VAR model to be 3. Following Nakajima (2011) and Zheng Tingguo et al. (2018), we specify the following time-varying parameter vector autoregression (TVP-VAR) model for investor sentiment ( ), realized volatility ( ), and the first-differenced turnover rate series ( ):

$$= 1 \quad -1 + \quad 2 \quad -2 + \quad 3 \quad -3 + \quad , \qquad (, \ ) \ (2)$$

where is a 3 × 1 column vector (, ,) , and (= 1,2,3) are 3 × 3 coefficient matrices. The random disturbance term is a 3 × 1 vector reflecting structural shocks, with being a lower triangular matrix with diagonal elements of 1 that captures simultaneous relationships among variables arising from structural shocks identified through recursive identification. is a diagonal matrix with elements (1 , 2 , 3) representing the standard deviations of random disturbances.

Transforming Equation (2) and adding an intercept term yields the final specification:

$$= \quad + \ 1 \quad -1 + \ 2 \quad -2 + \ 3 \quad -3 + \quad , \qquad (, \quad -1 \qquad (\quad -1))$$
(3)

Let denote the vector stacking the elements of (= 1,2,3) by rows, denote the vector stacking the lower triangular elements of , and = (1 , 2 , 3) where = for = 1,2,3. Assuming all parameters in Equation (3) follow first-order random walk processes with error terms distributed as:

$$+1 = \quad + \quad , \quad +1 = \quad + \quad , \quad +1 = \quad + \qquad (4)$$

where (,), (,), and (,). Following Zheng Tingguo (2018), we specify initial states as:

$$0 \quad (\ 0, \ 0), \ 0 \quad (\ 0, \ 0), \ 0 \quad (\ 0, \ 0) \ (5)$$

with 0 = 0 = 0 = and 0 = 0 = 0 = × .

4.3 Granger Causality Test

We conduct Granger causality tests on investor sentiment ( ), realized volatility ( ), and the first-differenced turnover rate series ( ) using 3 lags, consistent with our TVP-VAR model. Results indicate that at the 10% significance level, we reject the null hypothesis that investor sentiment ( ) does not Granger-cause changes in turnover rate ( ). At the 5% level, we reject the null that turnover rate ( ) does not Granger-cause investor sentiment ( ) and that investor sentiment ( ) does not Granger-cause realized volatility ( ). At the 1% level, we reject the null that realized volatility ( ) does not Granger-cause turnover rate ( ). Thus, lagged changes in investor sentiment effectively explain fluctuations in turnover rate and volatility; lagged changes in turnover rate effectively explain changes in investor sentiment; and lagged changes in realized volatility effectively explain turnover rate fluctuations.

4.4 Experimental Results Analysis

We set MCMC algorithm iterations to 11,000, discarding the first 1,000 "burn-in" samples and using the remaining 10,000 stationary samples for estimation. TVP-VAR model estimation results based on MCMC are presented in Table 6. All estimated parameters fall within their 95% confidence intervals, and all Geweke diagnostic values remain below the 5% significance level critical value, indicating that all parameter Markov chains satisfy convergence requirements. These results confirm the validity of our parameter estimates.

4.4.1 Impulse Response Analysis at Different Lags Following Yang Wenqi et al. (2020) and Ren Yongping et al. (2020), we select lag periods of 1 day, 7 days (one week), and 14 days (two weeks) to examine short-term, medium-term, and long-term impacts of unit shocks.

Investor Sentiment and Market Liquidity: Throughout the study period, impulse responses of market liquidity to investor sentiment remain consistently positive with notable time-varying effects. From June 2018 (near t=100) to July 2019 (near t=360), the impact of lag-1 sentiment shocks on market liquidity increased gradually. When investors confirmed the market had entered a bear phase (near t=20), uncertainty about future trends and duration, combined with risk aversion preferences, likely prompted frequent trading and market overreaction. This effect persisted until investors confirmed the transition to a bull market (near t=246). Additionally, lag-1 (short-term) effects are significantly stronger than lag-7 and lag-14 (medium- and long-term) effects, reflecting stock market information efficiency consistent with Yang Wenqi (2020). Conversely, while market liquidity' s impact on investor sentiment is relatively smaller, its time-varying effect is more pronounced (more sensitive to market state changes).

Investor Sentiment and Market Volatility: As shown in Figure 5, lag-1 investor sentiment exhibits consistently negative impulse responses to market volatility throughout the study period, with stronger shocks during bear markets and minimal medium- and long-term effects.

4.4.2 Impulse Response Analysis at Different Time Points Our study period covers the late stage of a slow bull market (January 2, 2018 to January 29, 2018), a bear market (January 30, 2018 to January 4, 2019), and the early-to-middle stage of a new bull market (January 5, 2018 to December 31, 2019). We select three representative time points—January 15, 2018 (t=10), July 2, 2018 (t=120), and June 27, 2019 (t=360)—to examine differential impacts across bull market late stage, bear market, and bull market early-to-middle stage.

Investor Sentiment and Market Liquidity: At all three time points, investor sentiment shocks produce consistently positive impulse responses in market liquidity, with short-term effects stronger than long-term effects. Lag-3 sentiment effects on market liquidity diminish significantly. Bull market sentiment impacts on liquidity are slightly stronger. These results support the view that sentiment effects on stock markets are asymmetric, with optimistic sentiment more likely to push stock prices upward away from fundamental values, leading to frequent trading and price bubbles. Conversely, market liquidity' s impulse responses to investor sentiment are positive within lag-2 periods before declining to negative values and gradually weakening toward zero. The likely transmission mechanism involves herd behavior: when investors observe frequent trading and rising turnover rates, their sentiment initially becomes optimistic, prompting them to follow the trend. Individual irrational decisions can evolve into convergent market-wide sentiment and actions, causing frequent trading, asset price deviations, and market volatility. However, after investors process information and make rational judgments, the impact of liquidity changes weakens. Notably, bear market liquidity changes exert significantly stronger impacts on investor sentiment than bull market changes.

Investor Sentiment and Market Volatility: Results indicate that investor sentiment affects market volatility more strongly during bear markets than bull markets, with influence magnitude experiencing repeated fluctuations in the first five days before gradually decaying. This pattern likely relates to investors' information processing capacity and information dissemination efficiency.

5. Conclusions and Recommendations

This study investigates investor sentiment' s impact on stock markets using deep learning methods. Based on data from Eastmoney.com and the Shenzhen Stock Index, we construct investor sentiment using the BERT model and incorporate market liquidity and volatility indicators. Experimental results show high correlations, confirming that investor sentiment significantly influences stock markets. Remaining deviations may stem from imperfect data crawling, high noise levels in BBS data, or minor model specification biases. However, overall results support our hypotheses and demonstrate strong model robustness.

Our conclusions are: (1) The BERT model effectively extracts investor sentiment from stock BBS comments; (2) The TVP-VAR model successfully captures relationships among investor sentiment, liquidity, and volatility.

Policy recommendations include: (1) Chinese stock market investor sentiment, particularly individual investor sentiment, significantly impacts markets and is easily influenced by social platforms like stock BBS. This necessitates further improvement of market institutions. As markets mature and institutional investors increase, stock price sensitivity to sentiment will decrease, reducing market risk. (2) While developing institutional investors as a key strategic priority for China' s securities market, equally important is improving their quality and rationalizing their investment behavior. Only by solving this problem can we enhance individual investor quality and improve market efficiency.

This study has limitations that warrant further research. Future work could examine the impact of national policies, market events, and unexpected incidents, incorporate additional explanatory variables to reduce noise, and test longer time series or smaller samples to further refine the model and parameters.

References

[1] Huicheng Liu. Leveraging Financial News for Stock Trend Prediction with Attention-Based Recurrent Neural Network.

[2] Yao Qin et al. A Dual-Stage Attention-Based Recurrent Neural Network for Time Series Prediction.

[3] Guoshuai Zhang. Learning to Predict U.S. Policy Change Using New York Times Corpus with Pre-Trained Language Model.

[4] Jiaqi Hou. A Neural Relation Extraction Model for Distant Supervision in Counter-Terrorism Scenario.

[5] Liao Shiguang, Yang Chaojun. Research on Liquidity Value in China' s Stock Market—Empirical Evidence from A- and B-Share Dual Listings[J]. South China Journal of Economics, 2008(12):64-78.

[6] Zeng Zhijian, Luo Changqing. Empirical Research on Liquidity Linkage Between Stock and Bond Markets[J]. Finance and Economics Theory and Practice, 2008(04):45-49.

[7] Wang Yintian, Wen Zhiying. Research on Liquidity Spillover Effects Between Stock and Bond Markets[J]. Journal of Financial Research, 2010(03):155-166.

[8] Han Jinxiao, Wu Weixing. Liquidity and Asset Pricing—Empirical Research Based on a New Spread-Based Liquidity Measure[J]. Shanghai Finance, 2017(12):51-57.

[9] Engle, R. Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation[J]. Econometrica, 1982, (50):987-1007.

[10] Bollerslev, T. Generalized Autoregressive Conditional Heteroskedasticity[J]. Journal of Econometrics, 1986, (31):307-327.

[11] Dai Wen. Empirical Analysis of ST Stock Return Volatility—Based on ARMA-GARCH Model Family[J]. Contemporary Accounting, 2018(10):9-10.

[12] Huang Hainan, Zhong Wei. Evaluation of GARCH Model Volatility Forecasting[J]. Chinese Journal of Management Science, 2007(06):13-19.

[13] Chen W, Chiu J, Chung H. Investor Trading Behavior, Market Liquidity, and the Role of Investor Sentiment[C]. 2009 FMA Annual Meeting. 2009.

[14] Li Chunhong, Peng Guangyu. Investor Sentiment, Stock Market Liquidity, and Economic Growth[J]. Technology Economics, 2011, 30(05):105-108+123.

[15] Wang Danfeng, Liang Dan. Stock Market Liquidity from an Investor Sentiment Perspective—Research on B-Share Opening to Domestic Residents[J]. Journal of Applied Statistics and Management, 2012, 31(02):363-373.

[16] Liu Xiaoxing, Zhang Xu, Gu Xiaoxian, Yao Dengbao. How Does Investor Behavior Affect Stock Market Liquidity?—Analysis Based on Investor Sentiment, Information Cognition, and Short-Selling Constraints[J]. Journal of Management Sciences in China, 2016, 19(10):87-100.

[17] Yang Xiaolan, Shen Hanbin, Zhu Yu. Local Bias, Investor Sentiment, and Stock Returns: Empirical Evidence from Online Forums[J]. Journal of Financial Research, 2016(12):143-158.

[18] Shi Guangping, Liu Xiaoxing, Wei Yuesong. Investor Sentiment, Market Liquidity, and Stock Market Bubbles—Analysis Based on TVP-SV-SVAR Model[J]. Financial Economics Research, 2016, 31(03):107-117.

[19] Zhang Zongxin, Wang Hailiang. Investor Sentiment, Subjective Belief Adjustment, and Market Volatility[J]. Journal of Financial Research, 2013, (04):142-155.

[20] Hu Changsheng, Chi Yangchun. Investor Sentiment, Asset Valuation, and Stock Market Volatility[J]. Journal of Financial Research, 2013, (10):181-193.

[21] Lu Jianqing, Chen Mingzhu. Analysis of the Internal Mechanism of Investor Sentiment' s Impact on Stock Market Volatility[J]. Shandong Social Sciences, 2015(S2):122-123.

[22] Ba Shusong, Zhu Hong. Margin Trading, Investor Sentiment, and Market Volatility[J]. International Finance Research, 2016(08):82-96.

[23] Tong Yuansong. Research on the Relationship Among Investor Sentiment, Stock Market Liquidity, and Volatility[J]. Technology Economics and Management Research, 2021(02):76-82.

[24] He Jibao, Lu Zhi. What Influences Investors' Decisions to Sell or Continue Holding?[J]. Journal of Management Sciences in China, 2006(06):53-63.

[25] Pang B, Lee L. A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts. In: Scott D, ed. Proc. of the ACL 2004. Morristown: ACL, 2004. 271-278.

[26] Wang Jingyi, Huang Yiping. Characterizing Fintech Media Sentiment and Its Impact on Online Lending Markets[J]. China Economic Quarterly, 2018, 17(04):1623-1650.

[27] Al-Nasseri A, Ali F M. What Does Investors' Online Divergence of Opinion Tell Us About Stock Returns and Trading Volume?[J]. Journal of Business Research, 2018, 86(1).

[28] Pawar K, Jalem R S, Tiwari V. Stock Market Price Prediction Using LSTM RNN[M]. Emerging Trends in Expert Applications and Security. Singapore: Springer, 2019:493-503.

[29] Devlin J, Chang M W, Lee K, et al. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding[C]//Proceedings of the 2019

Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg, PA: Association for Computational Linguistics, 2019: 4171-4186.

[30] Antweiler, W., Frank, M.Z. Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards[J]. The Journal of Finance, 2004, 59(3):1259-1294.

[31] Nakajima, J. Time-varying Parameter VAR Model with Stochastic Volatility: An Overview of Methodology and Empirical Applications[J]. Monetary and Economic Studies, 2011(29): 107-142.

[32] Zheng Tingguo, Liu Tangyong. Analysis of Stock Market Volatility Spillover Effects and Their Influencing Factors[J]. China Economic Quarterly, 2018, 17(02):669-692.

[33] Yang Wenqi, Li Kehan. How Do Investor Sentiment and Market Liquidity Affect Stock Market Stability?—Time-Varying Research on the STAR Market Using TVP-VAR Model from a Stock Price Crash Perspective[J]. Research in Institutional Economics, 2020(04):73-93.

[34] Ren Yongping, Li Wei. Economic Policy Uncertainty, Investor Sentiment, and Stock Price Synchronicity—Time-Varying Parameters Based on TVP-VAR Model[J]. Journal of Shanghai University (Natural Science Edition), 2020, 26(05):769-781.

Note: Figure translations are in progress. See original paper for figures.