Prioritizing the Net Sentiment Score: A Banking Industry Case Study

Por qué el índice de sentimiento neto debería ser una prioridad: un estudio de caso de la industria bancaria

José Guadalupe Mendoza-Macías

Quantum Analytics, Sinaloa, Mexico

Román Alejandro Mendoza-Urdiales

Quantum Analytics, Sinaloa, Mexico

272

Abstract

This study analyzes the impact of social media comments on the stock performance of banks registered on the United States stock market. We use artificial intelligence to monitor and extract comments in real time, together with natural language processing, to identify the sentiment of each comment. Comments were classified as positive or negative and were added by the hour for each bank during the observed period. Our results showed that both positive and negative comments have a significant effect on the stock performance, with negative comments having a more pronounced, asymmetrical impact. This study contributes to understanding how social media interactions influence the market value of companies, highlighting the importance of monitoring and managing the online perception of the companies listed on the stock market.

Keywords: *stock performance, net sentiment score (NSS), robust linear model regression, brand value impact.*
JEL Classification: *G41.*

Resumen

El artículo analiza el impacto de los comentarios en redes sociales sobre el rendimiento de las acciones de los bancos en el mercado de valores de EE.UU. Se empleó inteligencia artificial para monitorear y extraer comentarios en tiempo real, y se utilizó el procesamiento de lenguaje natural para calcular el sentimiento de cada comentario. Los comentarios se clasificaron como positivos o negativos y se agregaron, por hora, para cada banco durante el período observado. Los resultados mostraron que tanto los comentarios positivos como los negativos tienen un efecto significativo en el rendimiento de las acciones, con un impacto asimétrico más pronunciado en el caso de los comentarios negativos. Este estudio contribuye a la comprensión de cómo la interacción en redes sociales puede influir en el valor de mercado de las empresas y destaca la importancia para las compañías que cotizan en bolsa de monitorear y gestionar la percepción en línea.

Palabras clave: *rendimiento accionario, índice de sentimiento neto (ISN), regresión de modelo lineal robusto, impacto en el valor de la marca.*
Clasificación JEL: *G41.*

1. Introduction

The practice of manipulating information for personal financial gain predates academic literature. During the Napoleonic wars, a man claiming to be Colonel du Bourg stated that Napoleon I was murdered in the battle against the Bourbons (Evans,1859). Within a very short period, this claim greatly benefited the securities price backed by the British government. It was later revealed that the news was false, and a subsequent investigation by the Committee of the Stock Exchange concluded that Lord Cochrane had orchestrated the manipulation, selling over £1 million in assets during the frenzy.

The above example is but one of the multiple documented instances of the stock market being impacted significantly and swiftly through news manipulation, with the information source seldom verified. This is primarily due to market players seeking premiums, and being exposed to (and accepting) the risk. In modern times, more contemporary methods are employed to induce similar market effects in order to garner substantial yields within a brief period.

Social media serves as a common medium through which influence on publicly traded companies can be exerted. These platforms, either consciously or unconsciously, have virtually eliminated the gap between consumers and firms. In recent years, consumers have discovered that they wield significant power (at their fingertips) (Plourde, 2023). Situations arise where individuals, feeling unheard by the brands managed by firms, utilize the power of social media to demand attention not only from the brands themselves, but also from other customers, regulators, government agencies, and so forth. This phenomenon has informally and positively established social media as the "official" channel for communication between brands and end users, removing any intermediaries.

Brands have noticed this and are leveraging social media to successfully promote their products, as well as monitoring their reputation on the platforms. Reichheld (2011) describes the introduction of the term "net promoter score (NPS)" by the management consulting company, Bain and Company, in 2003. This metric was designed to measure customer loyalty towards a brand, company, product, or service, based on the probability of recommending the product to peers. More recently, Van Velthoven (2014) linked NPS to net sentiment score (NSS), contrasting the already calculated NPS of Vistaprint (an online provider of printed and promotional material

and marketing to small businesses and consumers) with comments mentioning the company on X (formerly Twitter). The study revealed that there is indeed a positive moderated correlation between NPS and NSS.

While firms and decision-makers generally agree on the direct impact of social media on sales (Donthu *et al*., 2021), the indirect effects on brand value and the workings of social media posts/trends-sales-brand popularity/value mechanism require further exploration. This study proposes a model that integrates economic, financial, linguistic, and psychological factors to capture the impact of social media comments on a firm's total value through stock prices. Additionally, we investigate how quickly this mechanism affects stock prices and company value.

In the era of digital communication, the influence of social media on public perception and behavior has become a significant area of interest. This study delves into the impact of social media interactions on the stock prices of publicly traded banks, a topic that has seen increasing scholarly attention in recent years. Notably, works such as Bollen *et al*. (2011), who investigated the relationship between the Twitter mood and the stock market, and Tetlock (2007), who examined the impact of media sentiment on securities markets, provide a foundational understanding of the nexus between public sentiment expressed online and market outcomes. Building upon this literature, our research aims to quantify the specific effects of social media sentiment—both positive and negative—on banking stocks, and thereby contribute to a more nuanced understanding of social media as a potent market mover.

Specifically, we seek to answer the following questions:

1) Is it possible to calculate the impact of social media interactions at a company level even when companies have a limited social media presence?

2) Is there an asymmetric effect between negative and positive interactions?

3) Is it possible to separate positive and negative sentiment for each company using robust linear model regression?

This paper is structured as follows: The "related work" section outlines similar work. The "problem definition" and "method" sections explain the proposed methodology. In "results and discussion" we show the experiment results and evaluation of the proposed method. Finally, in the "conclusions", some final assessments are made and direction for future research is provided.

2. Related Work

An interesting proposal focuses on measuring the impact of customer satisfaction on brand value. Colicev and O'Connor (2020) explore how social media data affects brand value by monitoring customer satisfaction, corporate reputation, word-of-mouth, and awareness using Partial Least Squares Path Modelling (PLS-PM). The study explores official content and user content to estimate latent variable scores, and concludes that social media marketing has a positive impact on customer satisfaction and that brand value is a precursor of sales and shareholder value.

Mendoza-Urdiales *et al*. (2021) present a methodology that categorizes and measures the impact of social media comments on the daily closing performance of 23 companies over a ten-year period. Their study obtained a success rate of over 80% in confirming the influence of social media on some of the world's largest publicly traded companies. Mendoza-Urdiales *et al*. (2022) propose a method to categorize positive and negative comments and their impact on daily performance by creating negative and positive variables. Using two methods, EGARCH and Transfer Entropy, the results indicated an asymmetric impact of negative comments compared to positive ones, persisting for 2-3 days after the comments had been posted. Núñez-Mora and Mendoza-Urdiales (2023) propose a big data approach in which they extracted all comments on social media that mentioned the 2557 most important publicly traded companies in the U.S. equity market. They categorized the comments as positive or negative, and were able to capture the asymmetric effect for the 508 largest publicly traded companies in the United States, with less than 30% missing data in the social media comments observations. Furthermore, the results show that the signal from social media impacts the stock market in under an hour.

Kirtac and Germano (2024) propose leveraging the ability of large language models (LLM) over traditional tools, like the Loughran-McDonald dictionary, to predict stock returns through sentiment analysis of financial news. Utilizing over 965,000 U.S. financial news articles from 2010 to 2023, they applied regression analysis and various performance metrics to evaluate the predictive accuracy of LLMs. The Open Pre-trained Transformer language model (OPT) showed exceptional performance with high accuracy and a notable Sharpe ratio in trading strategies. The study emphasizes the significant potential of LLMs in enhancing financial market prediction and investment strategy formulation, and advocates for their integration into financial analysis to improve market prediction and decision-making processes.

3. Problem Definition

The sample of banks included in our study were selected from the largest publicly traded companies in the United States. Specifically, we focused on companies that represent 99% of the total market capitalization, arranged in descending order of size, and totaling approximately 2557 entities. Within this universe, we targeted the 167 firms classified as banks under the industry RIC label. This selection criteria ensured that our analysis encompassed those banking institutions operating within the stock market that collectively represent 99% of the total market valuation on the U.S. stock exchange.

This methodical selection of banks allowed for the comprehensive examination of the impact of social media sentiment on stock prices across a significant portion of the banking sector. By focusing on such a substantial share of the market, the study aimed to develop insights that are both statistically robust and broadly applicable to major financial institutions within the United States. In Table 1, a sample of ten of the 167 banks is presented, with a brief description of the universe used in the study (see Table 1). Both global and local brands can be observed.

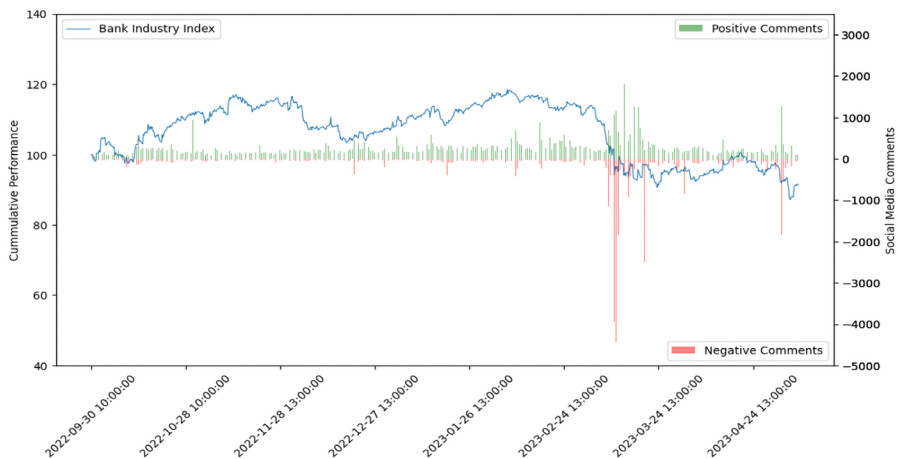**Table 1**. Sample of International and Local U.S. Banks

Bank Name	Reuters Instrument Code	Description
JPMorgan Chase & Co.	JPM.N	One of the largest and most influential global financial institutions, known for its extensive banking operations.
Bank of America Corp.	BAC.N	A major player in the global finance industry, serving millions of consumers and businesses worldwide.
Citigroup Inc.	C.N	A leading global bank with approximately 200 million customer accounts, operating in more than 160 countries.
Wells Fargo & Co.	WFC.N	One of the largest banks in the USA, known for its commercial and consumer banking services.
The PNC Financial Services Group, Inc.	PNC.N	Noted for its wealth management, asset management, and corporate banking services.
U.S. Bancorp	USB.N	The parent company of U.S. Bank, ranked as one of the largest commercial banks in the United States.

Bank Name	Reuters Instrument Code	Description
Goldman Sachs Group Inc.	GSBC.OQ	Primarily known for investment banking, but also a significant player in other banking sectors.
Morgan Stanley	MSBI.OQ	Recognized for its investment banking, wealth management, and asset management services.
Citizens Financial Group, Inc.	CFG.N	One of the oldest and largest financial services firms in the United States.
KeyCorp	KEY.N	The parent company of KeyBank, a regional bank headquartered in Cleveland, Ohio.

Source: Prepared by the authors.

The market cap weighted cumulative performance of the industry was calculated and paired with the aggregated classified comments grouped in positive/negative, as shown in Figure 1 (see Figure 1). It shows a correlation in the short-term variation between type of comments and performance. A higher number of positive comments than negative comments can be observed, except for the month of March 2023, when the regional bank bubble occurred and Silicon Valley Bank liquidity issues arose. Additionally, during the same period, the performance dropped drastically.

Figure 1. Market Cap-Weighted Performance of 167 Banks Operating in the U.S. Stock Market, Paired with Real-time Comments Grouped According to Sentiment (September 30, 2022 - May 5, 2023)

Source: Quantum Analytics.

Figure 2 presents the total volume of comments per hour per company during the analyzed period (see Figure 2). On average, there are 500 comments per hour, with peaks reaching up to 2500 comments and above. This indicates continuous interactions between banks and third parties. Social media users were categorized into two groups.

3.1 Direct Relationship

This refers to users with direct interaction with the company, either as current or potential customers, or through a direct interest in the company's products or services.
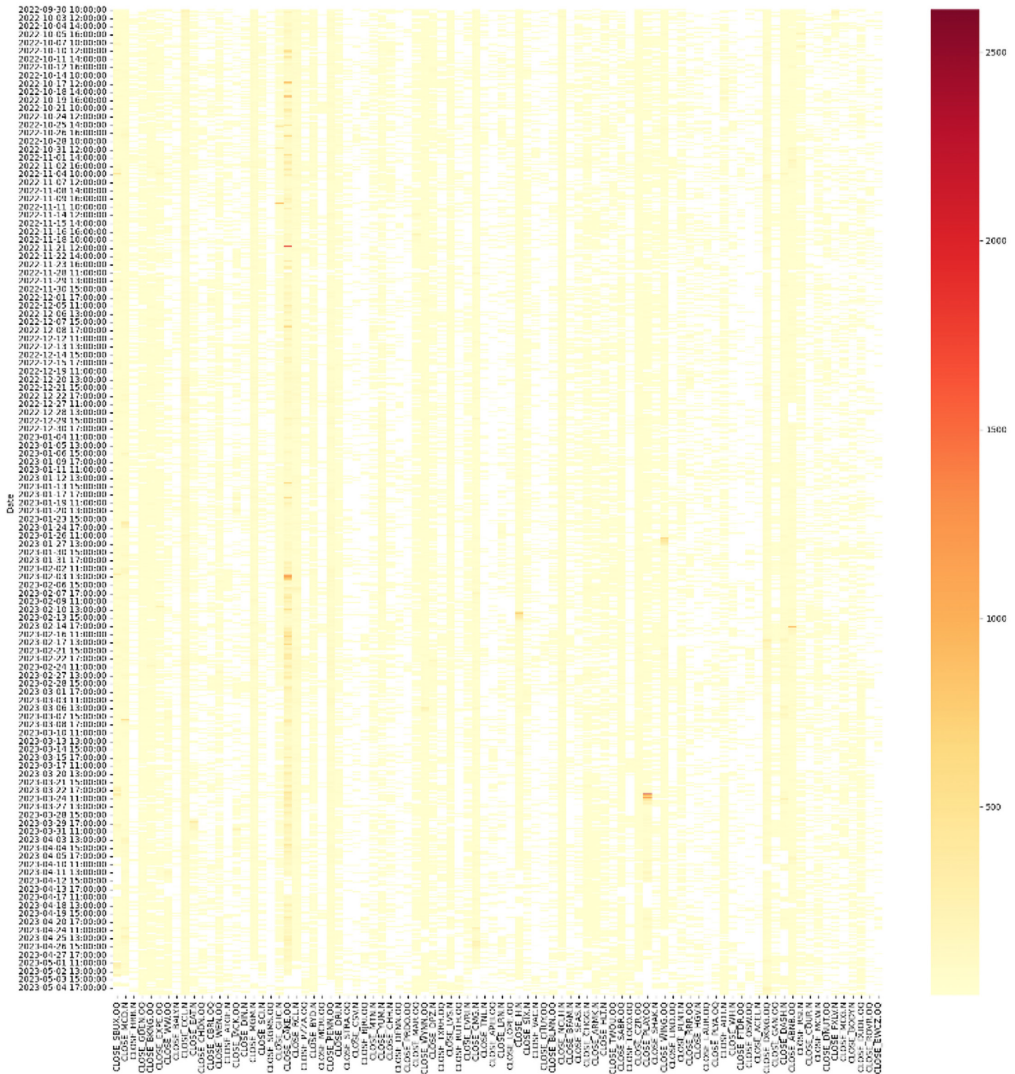
1) *Customers:* They have purchased or are directly interested in purchasing the company's products or services.

2) *Prospective Customers:* Individuals interested in the company's offerings who may become customers in the future.

3) *Fans or Followers:* Users who actively engage with the company's content due to a genuine interest in the brand.

4) *Brand Advocates:* Satisfied customers who actively promote the company and its products or services.

5) *Employees or Team Members:* People working for the company who engage with the profile to support and promote the brand.

3.2 Indirect Relationship

These users may not have a direct commercial interest in the company but interact with the profile for various reasons.

1) *Trolls:* Users who engage with the profile to provoke or disrupt without a direct interest in the company's offerings.

2) *Competitors:* Other companies in the same industry that may monitor the profile but do not directly engage in business with the company.

3) *Influencers:* Individuals with a significant following who can impact the company's reputation, although their engagement might not be directly tied to a commercial interest in the company.

Figure 2. Total Number of Comments for Each Analyzed Company (September 30, 2022 - May 5, 2023)

Source: Quantum Analytics.

4) *Critics:* Users who may voice concerns or criticism without being direct customers or having a direct commercial interest in the company.

5) *Neutral Observers:* Users who follow the profile out of curiosity or for informational purposes without actively engaging nor with a direct interest in the company's products or services.

Although several recent studies have focused on measuring the impact of social media comments on stock price performance (Li & Yang, 2024; Chung & Chang, 2024), it is still necessary to explore the impact of social media comments on individual company performance. Additionally, dividing the signal into negative and positive, aimed at capturing the asymmetric effect, is still a pending question to be solved in academia. The problem of missing data for social media has been addressed previously and is presented as a limitation (Núñez-Mora & Mendoza-Urdiales, 2023). In this study, we address this problem by running a robust linear regression model which avoids the manipulation of data and allows the use of raw data without imputation methods.

Figure 2 shows that certain companies maintain a consistent presence on social media, suggesting they are mainstream. In contrast, other banking institutions have minimal social media presence, posing a challenge from an individual analysis perspective (see Figure 2). This issue will be addressed through a combination of approaches in the following section.

4. Methodology

This study analyzes the industry level effect of social media comments on companies' performance. While the sentiment calculation and method to process and classify each comment as positive or negative is carried out individually for each company, the statistical modeling is a global approach in a single modeling using a robust linear regression.

When analyzing financial data, it is common to encounter anomalies and outliers, such as those observed during a market contagion or, as in this case, the regional banking crisis of 2023. These irregularities can frequently impact the variance of data, consequently affecting the accuracy of estimates and inferences derived from traditional panel data regression models. The implementation of robust statistical

methods, especially robust linear regression (RLM), offers distinct advantages over conventional panel regression techniques in managing such data.

In both Huber (1973) and Huber (1981), the resilience of statistical estimators to deviations from typical model assumptions is highlighted, particularly in the presence of outliers and heteroscedasticity. These methods utilize alternative loss functions, such as the Huber loss, which are less sensitive to outliers than the squared error loss used in an ordinary least square (OLS) regression. This adaptability allows for more accurate estimations even when the data strays significantly from standard assumptions like normality and homoscedasticity.

Additionally, a robust regression is adept at managing the unique variations across different panel units by diminishing the impact of outliers, a frequent occurrence in economic and financial datasets. As highlighted by Croux and Rousseeuw (1992), these methodologies not only bolster the reliability of statistical inferences but also enhance computational efficiency. This dual benefit is essential for extensive panel data analyses common in contemporary research settings, providing a robust foundation for statistical modeling that effectively addresses the complexities and specificities of real-world data.

Therefore, robust linear models enhance traditional panel data analysis frameworks by offering robustness against data anomalies and flexibility in accommodating non-standard data distributions, thereby becoming vital tools in the statistical analysis of panel data across various fields.

The framework used includes extracting all the public interactions of the social media platform X (formerly Twitter) in real time with the 167 banks that operated in the U.S. stock market from September 30, 2022, to May 5, 2023. The comments were analyzed using natural language processing algorithms that individually assigned each comment a negative or positive grade [-1,1] in which -1 is a fully negative classification and 1 is a fully positive classification. The process included cleaning the text of spelling errors, removing stop words, and retaining words that gave meaning to each comment. In this way, the individual sentiment for each comment was calculated for each company. If a comment mentioned more than one monitored bank, it was used to construct the sentiment for both banks.

Subsequently, the aggregated sentiment and the number of comments classified as either positive or negative were added up in hourly frequencies for each bank during the observed period. This approach aimed to observe how the number of positive and negative comments influenced the NSS of the firms.

The explanatory variables constructed for the model are sentiment (net promoter score), negative interactions, and positive interactions. The sentiment variable represents the hourly aggregated individual sentiment for each bank, while the positive and negative variables represent the total number of positive and negative mentions for each bank during the same hour according to the previously mentioned classification method. This implies that for each observation and each bank, there should be a value for sentiment, positive interactions, and negative interactions. The dependent variable is the standardized hourly returns for the banks. Finally, these variables are aggregated into time series for each bank and used in the RLM regression. The model aims to explain the normalized return of each bank's stock prices ('ZRET') using the sentiment ('SENT') and presence of positive ('POS') and negative ('NEG') comments. The following equations were constructed:

1. The equation 1 defines the sentiment score for company *A* at time *t*, which is calculated as the sum of individual post sentiments on social media.

2. The equation 2 represents a time series of sentiment scores, capturing the evolution of social sentiment over time, which allows for dynamic analysis of how sentiment impacts stock prices.

3. The equation 3 calculates the return for company *A* at time *t* as the percentage change in price from the previous time period. It is a fundamental metric in financial analysis, used to assess investment performance. The returns are then standardized in the equation 4, and the time series was constructed with the standardized results in equation 5.

4. The equation 6 calculates the total positive sentiment for company *A* at time *t*, summing all positive post sentiments. It reflects the aggregate positive public perception at a given time. Conversely, the equation 7 calculates the total negative sentiment.

5. The positive variable (POS) for time *t* was created by aggregating all comments with positive sentiment (P >0) for time *t* (equation 8). Conversely, the negative variable (NEG) was created by aggregating all comments with negative sentiment (P<0) (equation 9).

6. Finally, the equation for the robust linear model regression is presented (equation 10).

$$SENT_{A,t} = \sum_{n=1}^{i} P_{n,t} \tag{1}$$

$$SENT_A = [SENT_{A,t}, SENT_{A,t+1}, SENT_{A,t+2}, \dots] \tag{2}$$

$$RET_{A,t} = \frac{PRICE_{A,t}}{PRICE_{A,t-1}} - 1 \tag{3}$$

$$ZRET_{A,t} = \frac{RET_{A,t} - \mu RET_A}{\sigma RET_A} \tag{4}$$

$$ZRET_A = [ZRET_{A,t}, ZRET_{A,t+1}, ZRET_{A,t+2}, \dots] \tag{5}$$

$$POS_{A,t} = \sum_{i=1}^{n} P_{i,t} \; \forall Pi > 0 \tag{6}$$

$$NEG_{A,t} = \sum_{i=1}^{n} P_{i,t} \; \forall Pi < 0 \tag{7}$$

$$POS_A = [POS_{A,t}, POS_{A,t+1}, POS_{A,t+2}, \dots] \tag{8}$$

$$NEG_A = [NEG_{A,t}, NEG_{A,t+1}, NEG_{A,t+2}, \dots] \tag{9}$$

$$ZRET_{i,t} = \alpha_{i,t} + \beta_1 SENT_{i,t} + \beta_2 POS_{i,t} + \beta_3 NEG_{i,t} + \epsilon_{i,t} \tag{10}$$

This methodology allowed for a comprehensive analysis of the impact of social media sentiment on stock returns across a large number of banks.

5. Results

The robust linear model results with lag = 0 are shown in Table 2, where all exogenous variables (NSS, negative comments, and positive comments) yielded statistically significant coefficients (p-value < 0.05) (see Table 2). Additionally, an asymmetric effect between the coefficients of the negative and positive comments is evident, indicating that negative comments have a larger negative impact than the positive impact of positive comments. This is concordant with the asymmetric effect hypothesis.

Table 2. Results of Robust Linear Model Regression

```
              Robust linear Model Regression Results
================================================================================
Dep. Variable:                 Ret   No. Observations:            64661
Model:                         RLM   Df Residuals:                64657
Method:                       IRLS   Df Model:                        3
Norm:                       HuberT
Scale Est.:                    mad
Cov Type:                       H1
Date:               Mon, 13 May 2024
Time:                       10:11:25
No. Iterations:                 36
================================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
--------------------------------------------------------------------------------
const          0.0110      0.003      4.083      0.000       0.006       0.016
Sentiment     -0.0021      0.000     -4.705      0.000      -0.003      -0.001
Vader_Neg     -0.0009      0.000     -2.945      0.003      -0.001      -0.000
Vader_Pos      0.0008      0.000      2.941      0.003       0.000       0.001
================================================================================
```
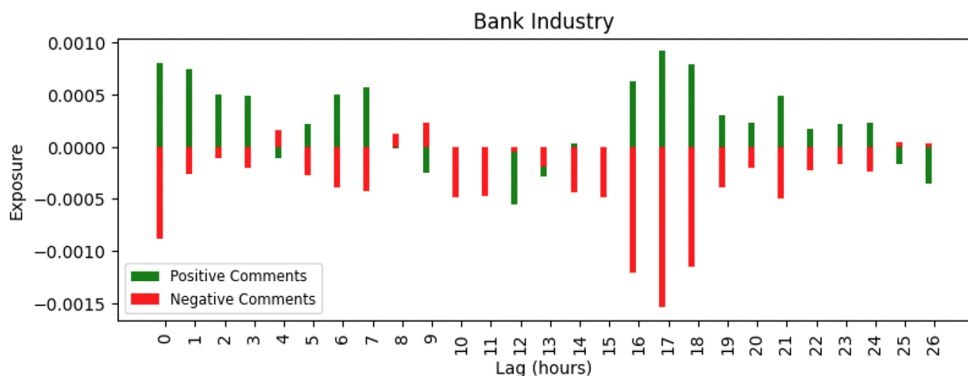
Source: Prepared by the authors.

The statistically significant coefficients suggest that social media sentiment does indeed influence stock prices. The asymmetric effect observed, where negative comments have a larger impact than positive comments, supports the notion that investors and market participants may react more strongly to negative news or sentiment. This finding could have implications for firms in managing their social media presence and monitoring the sentiment of online discussions about their company. It highlights the importance of not only promoting positive sentiment but also actively managing and responding to negative comments to mitigate potential negative impacts on stock performance.

The analysis was repeated 24 times, each time incorporating a 1-hour lag to the exogenous variables to assess the evolution of the impact of positive and negative interactions on performance. The results, illustrated in Figure 3, show that performance exposure is maintained over time, that the signal intensity varies, and that the opposing effect between negative and positive comments is clearly sustained (see Figure 3). These results suggest that social media interactions not only have an almost immediate effect on stock performance, but that their impact

can also be persistent over time with varying levels of intensity. Specifically, during the first eight hours after the comments are posted, their impact increases, and thus, the exposure of companies' performance to comment sentiment increases. Subsequently, the effect is dramatically reduced but continues to remain present for up to the following 24-26 hours. This finding suggests that monitoring and managing social media sentiment should be an ongoing effort rather than a one-time action, as the effects on stock performance are enduring and can vary in intensity over time.
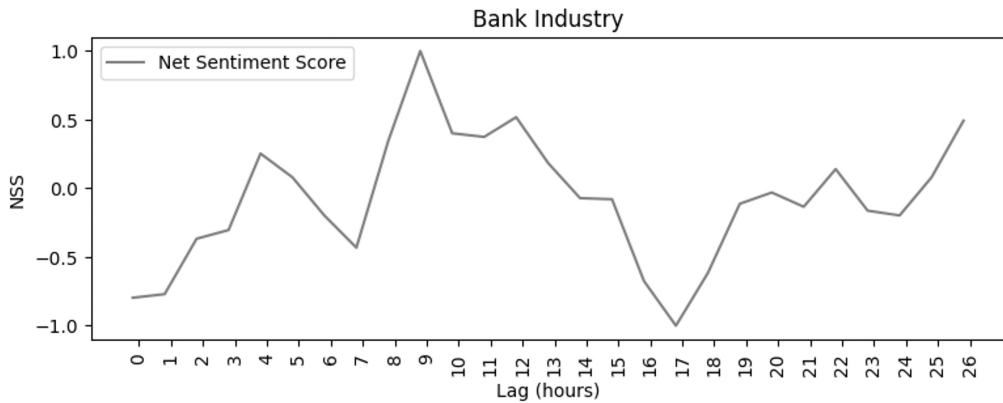
Figure 3. Influence of Positive and Negative Comments on Stock Performance Over Time

Source**:** Prepared by the authors.

The developed methodology enabled the calculation of the effect of the sentiment variable (SENT) in real time, which is widely known as the NSS. The results in Figure 4 intuitively demonstrate how NSS is aggregated using positive and negative comments, and how its lower scores are present when the coefficients of the negative comments increase negatively (see Figure 4). This highlights the importance of closely monitoring social media sentiment, particularly negative comments, as they can have a substantial impact on the overall sentiment score. This demonstration can aid in developing strategies to manage and mitigate negative sentiment to maintain a positive NSS.

Figure 4. Overall Net Sentiment Score Over Time

Source: Prepared by the authors.

6. Discussion

This study analyzed the impact of social media sentiment on stock performance, revealing that both positive and negative comments have a significant effect. Evans (1859) created a precedent regarding the influence of news on the stock market during the commercial crisis of 1857-1858, suggesting that information flow can impact market dynamics. However, the role of social media in shaping investor sentiment is a more recent phenomenon. Plourde (2023) discusses how consumers use social media as a platform for voicing their opinions about companies, which can influence brand perception and, indirectly, stock performance. Figure 1 captures the relationship of positive and negative comments in the performance of the banking industry. This is aligned with the findings of Reichheld (2011), who emphasized the importance of customer-driven strategies in enhancing company value.

Colicev and O'Connor (2020) argue that it is possible to differentiate the type of person making a comment depending on their relationship with the brand and calculate a direct influence. Nevertheless, our study results were conclusive without taking these distinctions into account. This suggests that while the source of a comment may influence its impact, the overall sentiment expressed in social media interactions remains a significant driver of brand value and stock performance.

287

Furthermore, Van Velthoven (2014) demonstrated a correlation between social media sentiment and net promoter score, indicating that positive online discussions can enhance brand loyalty and potentially affect stock prices. This is supported by the systematic review of Donthu *et al*. (2021), that mapped the electronic word-of-mouth research landscape and showed the growing significance of online sentiment in business outcomes. Figures 3 and 4 present the variation of the negative and positive comments impact through time and how this influences the NSC. Colicev and O'Connor (2020) found a positive impact of social media marketing on customer satisfaction, further emphasizing the link between online sentiment and company performance.

Mendoza-Urdiales *et al.* (2021) and Núñez-Mora and Mendoza-Urdiales (2023) provided empirical evidence of the direct relationship between social media comments and stock prices in individual firm analysis. Along with Kirtac and Germano (2024), this highlights the potential of leveraging large language models for sentiment trading. Furthermore, Li and Yang (2024) and Chung and Chang (2024) explored the broader implications of investor sentiment on market volatility and stock prices, respectively. Figure 2 presents the results of the combination of several approaches, in which the individual monitoring of several firms through a RLM captures the asymmetric effect by leveraging a robust natural language processing model.

7. Conclusion

The study reveals that firms are highly responsive to social media comments, with an immediate impact that persists over time and varies in intensity. It also shows that the NSS is more affected by negative comments than positive ones, suggesting a crucial insight for decision-makers monitoring this metric.

The study employed a bottom-up methodology to analyze the impact of social media on the stock performance of 167 publicly traded banks, revealing that both positive and negative comments have asymmetric effects on intraday data, with persistence in the signal post-commentary. The analysis also accounted for variability in the data, based on the popularity of each bank. Furthermore, the consistency of these findings across various types of studies, including different frequencies and modeling approaches (industry and firm-level), underscores the importance of natural language processing over the choice of statistical models in measuring the impact of social media interactions on stock performance.

This study provides a comprehensive understanding of social media sentiment's impact on publicly traded banks' stock performance. By employing a methodology that integrates natural language processing and RLM regression, the research demonstrates the significant and persistent influence of both positive and negative comments on stock prices. The findings highlight the importance of continuous monitoring and management of social media sentiment for firms, as it can substantially affect their stock performance and overall market perception.

Furthermore, the study reveals the asymmetric effect of positive and negative comments, with negative sentiment having a more pronounced impact on stock prices. This underscores the need for firms to develop strategies not only to promote positive sentiment but also to effectively address and mitigate negative comments to maintain a positive NSS.

The impact of social media comments and news on stock market performance has been extensively explored in recent decades. The 2017 Nobel Prize in Economics recipient, Richard H. Thaler, focused much of his research on how rationality influences decision-making, examining individual choices rather than their collective effect. These cognitive limitations can significantly impact financial markets. Thaler's work on behavioral economics (Thaler, 2017) highlights the role of psychological factors in economic decisions, which can be seen in how investors react to social media sentiments.

While this study observes how social media sentiment influences stock prices, it is important to acknowledge certain limitations that could affect the breadth and applicability of the findings. Firstly, the period under study may not fully capture long-term trends or the impact of extraordinary market events (seven months, from October 2022 to April 2024). Additionally, the reliance on social media data, which can be inherently biased, introduces potential variability in the sentiment. The information used in the study, extracted from public social media, may not represent all factors influencing stock market behavior, such as undisclosed financial information, which can also play a critical role. Acknowledging these constraints is crucial for a balanced interpretation of the results and for future research.

In conclusion, the intersection of behavioral economics and social media sentiment analysis presents a fertile ground for further research. Understanding the psychological underpinnings of investor reactions to online information can provide valuable insights into market behavior and inform strategies for managing the impact of social media on stock performance.

Further work could focus on grouping comments according to user type to analyze whether a varying level of influence among those groups exists. It can also analyze the asymmetric effect of negative and positive comments over time across the broader financial market. Additionally, further investigation into the mechanisms behind the asymmetric effect of positive and negative sentiment could provide deeper insights into investors' psychology and market dynamics.

References

Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science*, 2(1), 1-8. https://doi.org/10.1016/j.jocs.2010.12.007

Chung, M.H., & Chang, Y.K. (2024). Financial reporting complexity, investor sentiment, and stock prices. *Finance Research Letters*, 62(Part A), article 105026. https://doi.org/10.35692/07183992.13.1.8

Colicev, A., & O'Connor, P. (2020). How Social Media Impacts Brand Value: The Mediating Role of Customer Satisfaction. *Multidisciplinary Business Review*, 12(1), 15-26. https://doi.org/10.35692/07183992.13.1.8

Croux, C., & Rousseeuw, P. J. (1992). Time-efficient algorithms for two highly robust estimators of scale. Computational Statistics, Physica-Verlag HD.

Donthu, N., Kumar, S., Pandey, N., Pandey, N., & Mishra, A. (2021). Mapping the Electronic Word-of-Mouth (eWOM) Research: A Systematic Review and Bibliometric Analysis. *Journal of Business Research*, 135, 758-773 https://doi.org/10.1016/j.jbusres.2021.07.015

Evans, D. M. (1859). *The history of the commercial crisis, 1857-58, and the stock exchange panic of 1859*. Groombridge and Sons.

Huber, P. J. (1973). The 1972 Wald Memorial Lectures: Robust regression: Asymptotics, conjectures, and Monte Carlo. *The Annals of Statistics*, 1(5), 799-821. https://www.jstor.org/stable/2958283

Huber, P. J. (1981). *Robust statistics*. John Wiley & Sons.

Kirtac, K., & Germano, G. (2024). Sentiment Trading with Large Language Models. *Finance Research Letters*, 62(Part B), article 105227. https://doi.org/10.1016/j.frl.2024.105227

Li, J., & Yang, J. (2024). Financial shocks, investor sentiment, and heterogeneous firms' output volatility: Evidence from credit asset securitization markets. *Finance Research Letters*, 60, article 104860. https://doi.org/10.1016/j.frl.2023.104860

Mendoza-Urdiales, R. A., García-Medina, A., & Núñez-Mora, J. A. (2021). Measuring information flux between social media and stock prices with Transfer Entropy. *PLOS ONE*, 16(9), article 0257686. https://doi.org/10.1371/journal.pone.0257686

Mendoza-Urdiales, R. A., Núñez-Mora, J. A., Santillán-Salgado, R. J., & Valencia-Herrera, H. (2022). Twitter Sentiment Analysis and Influence on Stock Performance Using Transfer Entropy and EGARCH Methods. *Entropy*, 24(7), 874. https://doi.org/10.3390/e24070874

Núñez-Mora, J. A., & Mendoza-Urdiales, R. A. (2023). Social sentiment and impact in US equity market: an automated approach. *Social Network Analysis and Mining*, 13, 111. https://doi.org/10.1007/s13278-023-01116-6

Plourde, A. (2023). Criticizing companies via social media and digital platforms: an alternative forum of justice for consumers? *Option consommateurs*. CID: 20.500.12592/m8xnp1. https://canadacommons.ca/artifacts/8246598/criticizing-companies-via-social-media-and-digital-platforms/9163537/

Reichheld, F. (2011). *The Ultimate Question 2.0 (revised and expanded edition): How Net Promoter Companies Thrive in a Customer-Driven World*. Harvard Business Review Press.

Tetlock, P. (2007). Giving content to Investor Sentiment: The role of Media in the Stock Market. *The Journal of Finance*, 62(3), 1139-1168. https://doi.org/10.1111/j.1540-6261.2007.01232.x

Thaler, Richard H. (2017). The Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2017. Nobel Prize 2017.https://www.nobelprize.org/prizes/economic-sciences/2017/thaler/facts/

Van Velthoven, S.T.M. (2014). *Sentiment analysis on social media to predict Net Promoter Score* (master's thesis), Eindhoven University of Technology Research. https://pure.tue.nl/ws/portalfiles/portal/46988096/783003-1.pdf

▬▬ **About the authors**

José Mendoza is a distinguished professional in the technology and financial sectors, with over three decades of experience in leadership roles and systems development. He demonstrates exceptional skill in integrating technological solutions into complex business operations. Throughout his career, José has led a variety of innovative projects, excelling in the implementation of biometric systems and customer identification platforms that have transformed banking and retail operations in various countries. His commitment to continuous technological advancement and his ability to adapt new technologies to business demands highlight his strategic vision and executive competence. José contributes not only his vast technical knowledge but also his practical experience in process optimization and the application of strategic technologies, which is essential for companies in the technology and financial sectors aspiring to lead in a highly competitive market.

mendoza@quantumanalytics.mx
https://orcid.org/0009-0006-7804-4849

Román Mendoza is a leader in the artificial intelligence industry within consulting, working on developing solutions, strengthening capabilities, and generating knowledge in the field. His work focuses on delivering cutting-edge projects that integrate big data analytics and natural language processing to drive data-driven decision-making. With a solid track record of over 12 years in consulting projects, Román has specialized in implementing advanced technologies for operations research.

roman@quantumanalytics.mx
https://orcid.org/0000-0003-2888-156X