5-2023

Exploring Social Media Sentiment Analysis in Relation to Common Stock

Landon Johnson

University of Arkansas, Fayetteville

Exploring Social Media Sentiment Analysis in Relation to Common Stock

by

Landon W. Johnson

Advisor: Sergio Santamaria

An Honors Thesis in partial fulfillment of the requirements for the degree Bachelor of Science in Business Administration in Finance

Sam M. Walton College of Business
University of Arkansas
Fayetteville, Arkansas

May 11, 2018

Table of Contents

Abstract

Many institutions, investors, and researchers have sought to gain insights into consumer attitudes and behaviors by leveraging social media. However, trading strategies utilizing social sentiment have been observed to often provide returns roughly equivalent to or less than those of standard market indexes. This observation raises several questions: does the efficacy of social sentiment analysis vary from firm to firm? Are price or volume directly correlated with social sentiment? Does social sentiment display any predictive ability? The following analysis attempted to answer these questions by comparing one week's worth of data from Twitter users to price and volume data from six publicly traded companies. Although more data is needed to validate this analysis' findings, there was evidence to suggest that social sentiment's efficacy varies significantly between firms, changes in prices and volume do not seem to be correlated to changes in social sentiment, and there does appear to be some patterns that could help predict trend reversals in price.

Introduction

Over the past two decades, social media has become an ever increasingly important part of our daily lives. According to a 2023 study, roughly 90% of Americans use some form of social media actively (Ruby, 2023). Similarly, investors, particularly younger investors, often use social media as a tool to research stocks, brainstorm potential trades, and learn popular trading strategies. In fact, over 50% of investors under the age of 35 turn to YouTube for investing advice (Nasdaq, 2021). However, many experts have criticized this behavior, arguing that over reliance on social media can lead to "herd mentalities" and other dangerous phenomena that end up destabilizing certain stocks as well as creating uncertainty in the broader market. Regardless of whether social media's overall impact on the market is beneficial or not, its use by both institutional and retail investors has already been firmly cemented, and it is likely that social media's importance concerning investing decisions will only continue to increase as time passes. With that being said, investors can gain insights from social media in two ways. First, and most commonly, an individual will directly consume news, opinions, and other relevant information directly from their own personalized "feed" or recommendations, and this experience may help to shape his or her own thoughts and opinions. In contrast, an individual may choose to sample general data gathered from a large number of social media users and analyze this data for underlying patterns and trends that may provide insights into market behaviors. Due to its more analytical nature, the focus of this paper's study will be on the latter method. More specifically, the following analysis will help explore the usefulness of social media analysis by attempting to answer several questions. First, does the effectiveness of social media sentiment analysis vary from firm to firm? Second, does social sentiment have any significant correlations with a stock's price or volume? Finally, does social sentiment analysis have any predictive power relating to price?

Literature Review

Social media sentiment, also known as social sentiment, is often analyzed by both academic and corporate researchers. According to Investopedia, a social sentiment indicator can be defined as an indicator that analyzes aggregated social media data to help businesses understand how they are performing in the eyes of consumers. Additionally, social indicators can be used by investors who wish to gain insights into a company's decisions, such as the successfulness of a new product or how the company's public image is faring in general. In a 2019 study, using posts from investor-oriented social media platform, StockTwits, researchers were able to demonstrate a significant correlation between stock returns and overall sentiment on the platform (Leung & Woon Sau, 2019). In a similar study, researchers analyzed Twitter users' sentiment around FOMC meetings and found correlations between sentiment before the meeting and returns immediately following (Azar et al., 2016). One study, which focused solely on the relationship between Twitter data and Tesla stock, found "clear correlation" between an increase in the number of tweets and an increase in Tesla's closing price (Bhadamkar & Bhattacharya, 2022). However, due to Tesla and Twitter's unique connection via Elon Musk, conclusions stemming from this study may not be as accurate when applied to other firms. Many studies concerning this topic, such as the analysis by Bhadamkar and Bhattacharya, opt to use machine learning algorithms or neural networks to find complex relationships between social sentiment and stock prices. Additionally, social media sentiment has been used as the basis for several stock market indices. For example, the S&P 500 Twitter Sentiment Index as well as the S&P 500 Twitter Sentiment Select Equal Weight Index follow the S&P 500 and track the top 200 and top 50 S&P 500 stocks, respectively, with the most positive sentiment from Twitter. As shown in Table 1, despite leveraging the opinions of Twitter users, both indexes have generally underperformed the base S&P 500 index. This may suggest that leveraging social media sentiment alone may not be adequate for consistently beating market returns by significant margins.

Table 1

S&P 500 Index compared to Social Sentiment Based Alternatives			
	S&P 500	Twitter Sentiment (Market Cap Weights)	Twitter Sentiment (Equal Weights)
1 Year Return	-8.39%	-10.23%	-11.75%
3 Year Return Annualized	15.51%	13.72%	14.20%
5 Year Return Annualized	9.53%	9.59%	10.03%
10 Year Return Annualized	10.21%	N/A	N/A

Table 1 was created using data provided by S&P Global Inc. All data in Table 1 is current as of April 9th, 2023. For each row, green coloration represents the highest return, yellow coloration represents the middle return, and orange coloration represents the lowest return of the three indexes charted. Due to the Twitter sentiment derived indexes' relatively recent creation, data concerning 10-year returns and longer is unavailable at the time of writing.

Data Collection and Review

In this analysis, six companies were chosen for study. Those companies were Tesla (TSLA), Apple (AAPL), Microsoft (MSFT), Johnson & Johnson (JNJ), Exxon (XOM), and JPMorgan Chase (JPM). These specific companies were chosen for several reasons. First, each of these companies has a relatively high market capitalization. Secondly, these companies cover a variety of sectors and represent a diversity of business models. Additionally, these companies each had a sufficient amount of data points (interest on Twitter) over the course of data collection. Many companies listed in the S&P 500 did not meet this requirement, and therefore, may have provided misleading results. Finally, several companies such as Amazon displayed significant amounts of apparent outliers in source Twitter data, and therefore, these companies were not chosen to be further analyzed. This analysis was primarily organized into the following two parts: data gathering and analysis.

For this analysis, all data was taken from Twitter using Twitter's official developer API from March 27, 2023, 7:00 a.m. Central Time to April 1, 2023, 12:00 a.m. Central Time. In this analysis, only one week of data was pulled from Twitter due to authorization restrictions imposed by Twitter. Similar to other studies and analytics providers, Twitter data was chosen due to Twitter's generous public API policy and a plethora of related documentation. For each company, relevant data was pulled from Twitter using a specific set of parameters. Social sentiment can be calculated in a multitude of ways, and the exact formula for how sentiment is calculated can often be a "black box". In this study, Twitter data for each company was collected by analyzing a count of tweets that occurred over the specified time period, contained the concerned company's symbol within the text of the tweet, and featured the words "buy" or "sell" within the text. With those data points, changes over time in the proportion of "buy" and "sell" remarks as well as the raw number of mentions of "buy" or "sell" were used to infer social sentiment or momentum on a certain stock.

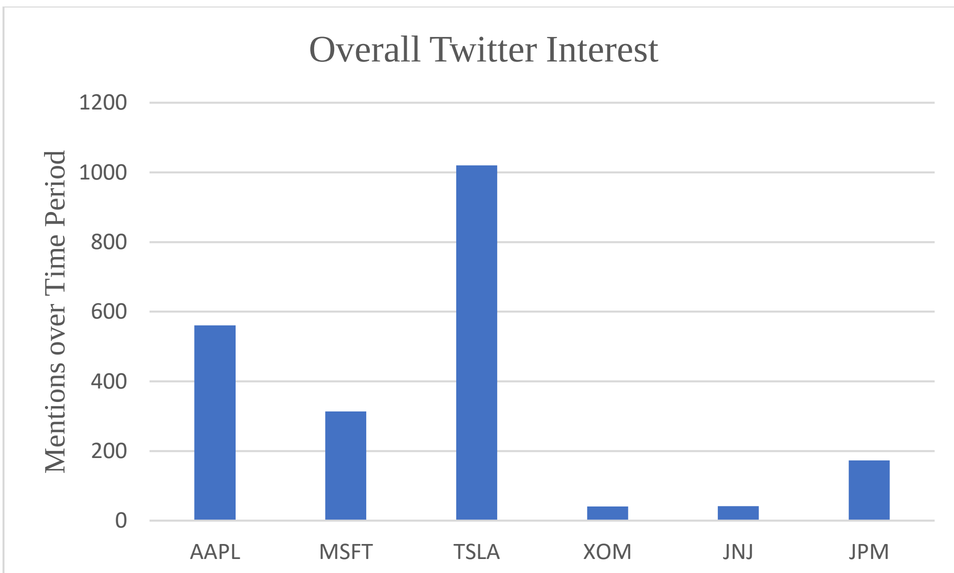
During the second part of this analysis, trends observed in the Twitter derived data were compared to independent data taken from a Bloomberg Terminal. More specifically, Twitter sentiment data trends were compared to the associated stock's price and volume during trading hours.

Results and Interpretation

In this study, data from Twitter was gathered and compiled into several forms: number of "buy" mentions every hour by company, number of "sell" mentions per hour by company, net interest per hour by company (calculated by "buy" mentions minus "sell" mentions), total interest per hour by company (calculated by "buy" mentions plus "sell" mentions), and overall twitter interest (the sum of all mention counts over the study time period). Chart 1 and Chart 2 were created using data collected from Twitter.

Chart 1 helps to visualize the relative popularity of this study's companies on Twitter during the study time period.

Chart 1



Relatively speaking, there was a significant difference between how often each of these companies was tweeted about. This could be partially explained by retail investors' greater desire to speculate and share their opinions about "interesting" stocks such as Tesla and Apple compared to what could be perceived as more "boring" companies such as Johnson & Johnson.

Chart 2

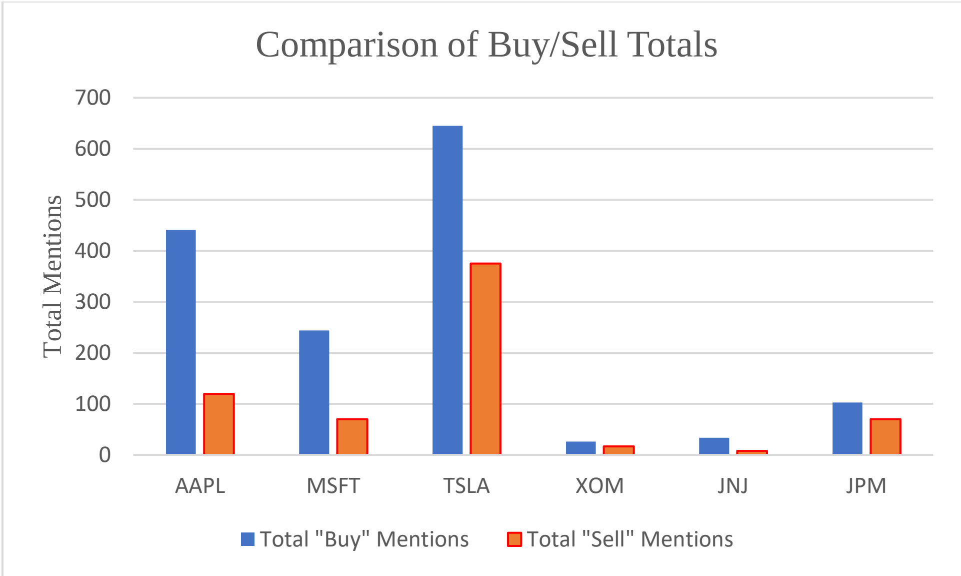


Chart 2 expands upon Chart 1 and provides a more detailed look at how mentions were broken down by "buy" and "sell" mentions. Generally, buy mentions can be considered a sign of positive sentiment, while sell mentions can be considered a sign of negative sentiment. According to this logic, the overall sentiment for every stock considered was positive for the time period.

Table 2

Comparison of Returns of Period			
TSLA	**Starting Price**	**End Price**	**Period Return**
AAPL	159.91	164.96	3.2%
MSFT	279.86	289.27	3.4%
TSLA	195.04	207.46	6.4%
XOM	104.13	109.66	5.3%
JNJ	153.46	154.93	1.0%
JPM	126.81	130.28	2.7%

Table 2 was compiled using data from Bloomberg. Inferring overall sentiment during the study to be positive was supported by the fact that returns for every company studied were positive over the same period. Additionally, Tesla, which received the most interest over the study period, also had the highest return of the period. However, looking at other companies, this trend does not seem to hold up. Following Tesla, Exxon had the second highest return; however, Exxon's overall interest level was significantly below Apple and Microsoft's interest levels despite having considerably greater price appreciation during the period. This could be one sign to help indicate that social sentiment analysis tools may not yield the same results when applied to fundamentally different companies. Specifically, social sentiment analysis may lead to subpar results when applied uniformly across companies. For example, analyzing social sentiment for companies such as Amazon and Google may require a significantly different approach and interpretation compared to lesser talked about, quieter companies such as, for example, Texas

Instruments. To further investigate the differences between how social sentiment reflects the returns of different companies, correlations between price, volume, “buy” mentions, “sell” mentions, net Twitter interest, and total Twitter interest were compared in Tables 3 and 4. Tables 3 and 4 were created using data collected from Twitter and data downloaded from Bloomberg.

Table 3

Comparison of Correlations between Social Sentiment and Stock Price

Company	Price & "Buy"	Price & "Sell"	Price & Net Interest	Price & Total Interest
TSLA	0.19	-0.17	0.29	0.04
MSFT	0.14	-0.26	0.22	0.05
JPM	0.30	-0.01	0.20	0.16
AAPL	-0.07	0.12	-0.09	-0.04
XOM	-0.22	-0.13	-0.12	-0.11
JNJ	-0.24	0.02	-0.26	-0.18

Table 3 details correlations between the price of a stock and the corresponding number of “buy” or “sell” mentions that occurred during the same hourly period, as well as statistics derived from these mentions (net interest and total interest). Correlations can range from 1, a perfect positive relationship, to -1, a perfect negative relationship. The correlations observed in this set of data ranged from very weak (+/- 0-0.19) to weak (+/- 0.2-0.39). Additionally, correlations tended to grow stronger after using “buy” and “sell” mentions to calculate net interest. This may indicate that net interest is a better indicator of sentiment than simpler interpretations. While weak correlations are not very impressive by themselves, this could indicate that social media sentiment could be a relevant variable while building a more complete model to analyze stock prices. If used in conjunction with other indicators, the model’s overall R-value could be significantly higher.

Table 4

Comparison of Correlations between Social Sentiment and Volume

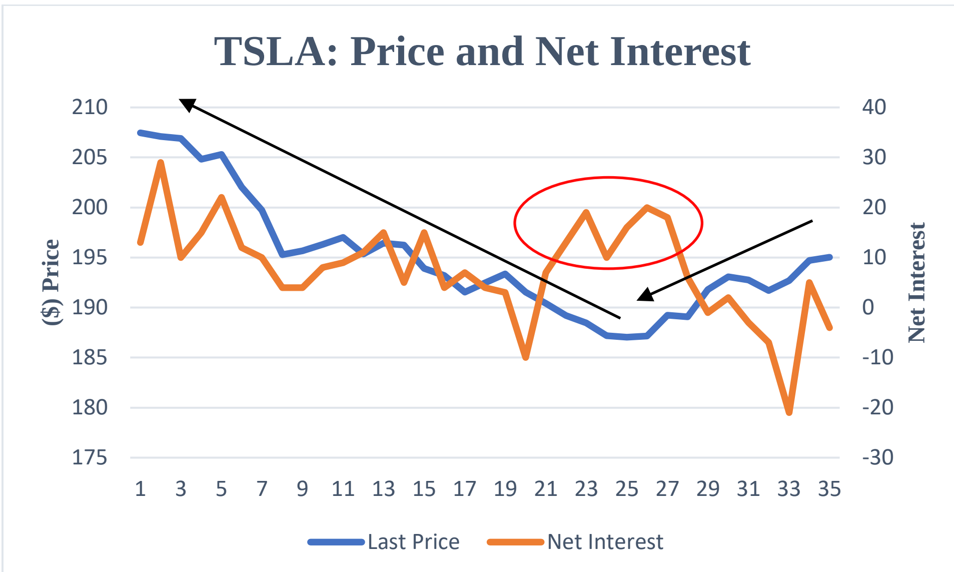
Company	Volume & Net Interest	Volume & Total Interest
TSLA	-0.06	0.11
MSFT	-0.35	-0.28
JPM	-0.31	0.30
AAPL	-0.57	-0.04
XOM	0.06	0.04
JNJ	0.06	0.34

Interestingly, when comparing the volume of a stock to net interest, either very weak correlations or weak to moderate negative correlations were found. Furthermore, companies with more interest on Twitter seemed to display negative correlations with net interest; however, companies with less Twitter interest (XOM and JNJ) displayed slightly positive correlations. In

Table 4, the correlation between Apple's volume and net interest was highlighted in red due possessing a significant negative correlation.

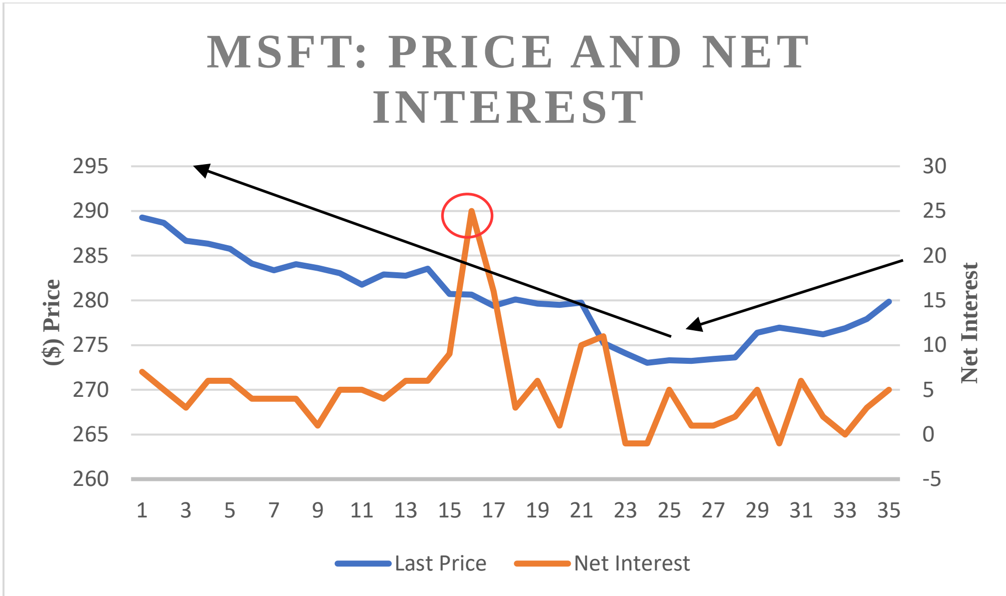
Charts 3-9 were created using data pulled from Twitter and pricing data downloaded from Bloomberg. These charts display visualizations of the relationship between price and net interest. Additionally, it should be noted that the horizontal axis, numbered from 1 through 35, indicates the remaining number of open market hours in the study period. Therefore, data points above the label "1" represent data from the final hour of the study period. Similarly, data points above the label "35" represent data from the first hour of the study period. Effectively, this means that **Charts 3-9 should be read from right to left to better reflect the flow of time**.

Chart 3
Tesla: Price vs. Net Interest



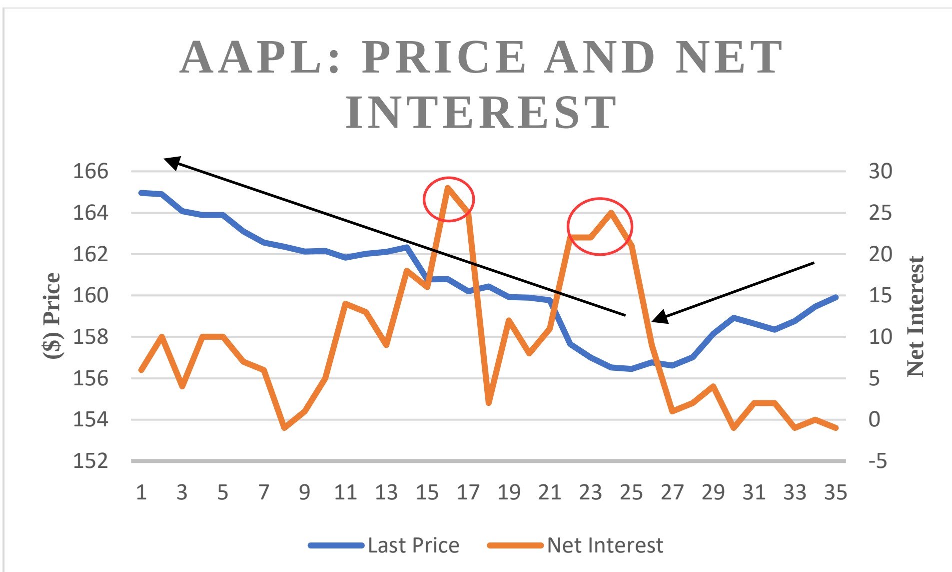
In Chart 3, a spike in net interest immediately precedes a visible reversal in price trend. Additionally, this spike displays a double top pattern.

Chart 6
Microsoft: Price vs. Net Interest



In Chart 4, a clear spike in net interest occurs shortly after a reversal in the trend of price.

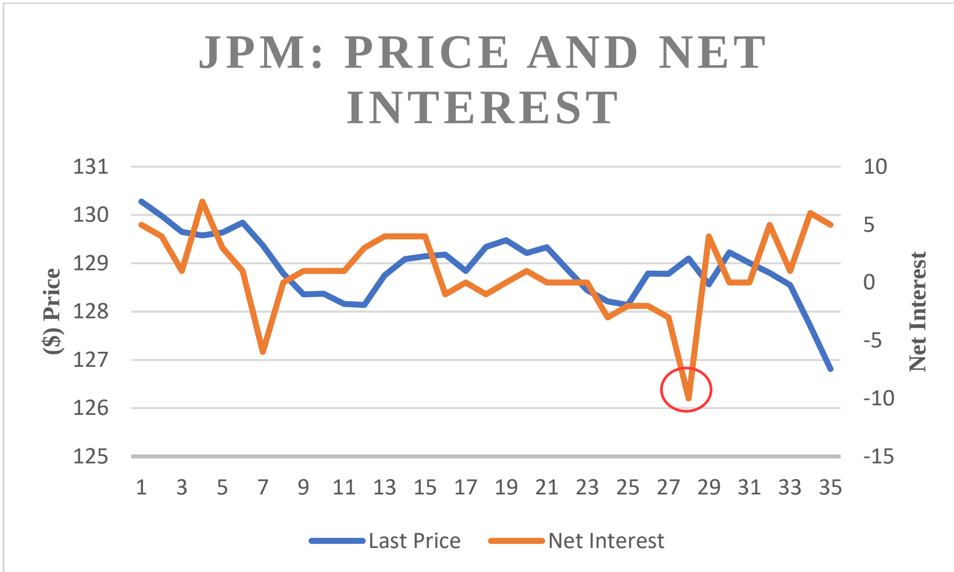
Chart 5
Apple: Price vs. Net Interest



In Chart 5, one spike in net interest coincides with a reversal in trend. While another occurs as the trend continues.

Chart 6

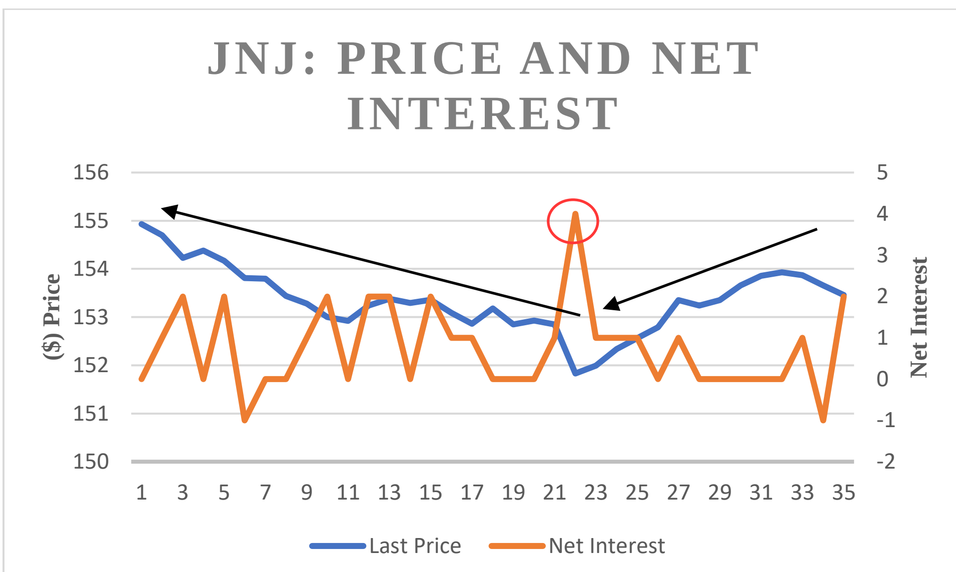
JP Morgan: Price vs. Net Interest



In Chart 6, a significant dip in net interest is followed by a sideways movement in price.

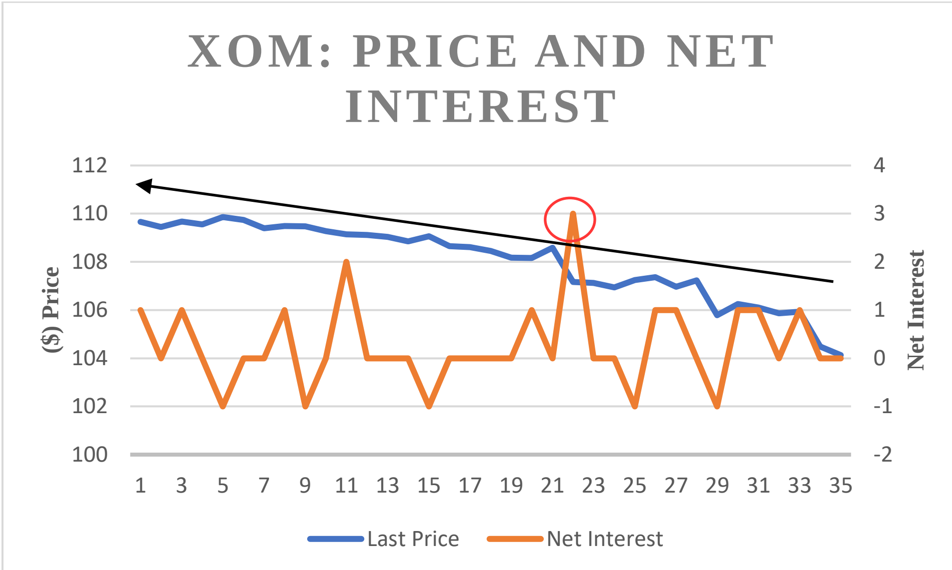
Chart 7

Johnson & Johnson: Price vs. Net Interest



In Chart 7, a spike in net interest coincides with a reversal in price trend.

Chart 8
Exxon: Price vs. Net Interest



In Chart 8, after a spike in net interest, price continues to climb in a stable pattern.

Chart 9
JP Morgan: Price vs. "Buy" Mentions

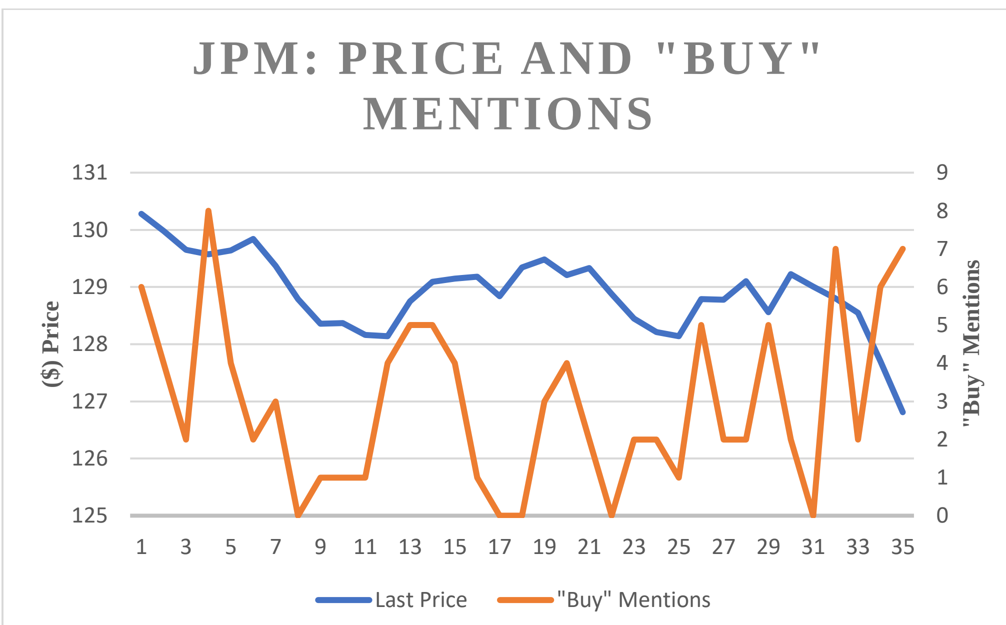


Chart 9 helps visualize a moderate inverse relationship between "buy" mentions and price. This inverse relationship may indicate that certain forms of social sentiment may actually be contrarian indicators.

As displayed in Charts 3,4,5,7, and 8, despite overall weak or negative correlations between price and net interest, there does seem to at least be a visual connection between peaks in net interest and a reversal in price trend to the positive direction, or if already moving in a positive trend, further support for that movement. Troughs in net interest occurred much less often and resulting trends were nonconclusive. Therefore, this study's findings could be interpreted to support the idea that social sentiment could be an indicator that is especially useful for spotting reversals in stock prices as well as providing evidence for a continuation in trend. Due to the limited scope of this study, implementing trading rules and assessing a trading strategy utilizing these spikes was difficult. However, from a technical analysis perspective, these initial findings are interesting and could warrant further attention in a study using more data points.

Conclusion

In summary, this analysis sought to answer several questions. First, is social media sentiment analysis an appropriate tool for analyzing most stocks? Second, does social sentiment have any significant correlations with a stock's price or volume? Finally, does social sentiment analysis have any predictive power relating to price? Because this analysis only included data from one week of trading, no definitive answers could be made to the previous questions. However, insights and patterns from this analysis could be used to help formulate questions and hypotheses that could be validated or rejected in additional studies with access to more data.

In regard to the first question, several interesting insights were made. Generally, index funds that employ social sentiment as the primary tool for portfolio allocation do not outperform the market consistently. One possible explanation for this may be that social media analysis is a tool that is best used to compare time-series data from a single company instead of trying to compare different companies to each other. For example, the S&P 500 Twitter Sentiment Index tracks the top 200 S&P 500 members with the most "positive" sentiment on Twitter. Presumably, 200 companies are tracked in order to narrow down the S&P 500 while still providing adequate diversification; however, due to this relatively large number of firms, generalized social sentiment tools may not work for analyzing every firm. This is supported by the fact that total volume of tweets concerning certain companies varies wildly from firm to firm, and a handful of firms tend to consistently dominate discussions in social media spaces.

In regard to the second question, this analysis did not consistently find strong or even moderate correlations between positive Twitter sentiment and stock price increases. In fact, in half of the studied firms, the opposite was found; increases in positive sentiment accompanied declines in price. This could indicate that, in many cases, social sentiment actually possesses more value as a contrary indicator as opposed to a predictive indicator. This provides further support for the idea laid out in the previous paragraph. With mixed correlations between sentiment and price, it is logical that indexes utilizing similarly derived social sentiment tools would exhibit mixed returns. When analyzing the relationship between volume and social sentiment, a relationship ranging from weakly inverse to no meaningful relationship was observed.

In regard to the final question, somewhat consistent social sentiment trends were found that could be useful in predicting future price movements. In five of the six charts comparing price and net interest, visually sharp increases in net interest preceded significant accompanying moves in price. In Charts 3,4,5,7, and 8, these sharp peaks in net interest were highlighted using a red circle, and the accompanying price movement was modeled using an arrow, which emphasized the general trend. Due to an insufficient amount of data over a longer time period, it is difficult to claim that this observed pattern was both significant and consistent. Despite this,

this pattern could still be of interest to traders. For example, in a declining market, a spike in social media interest could be analyzed in conjunction with other indicators to predict a reversal in a trend, and additional peaks in social sentiment after this point could be viewed as support for this trend.

Overall, conclusions drawn from this analysis support the idea that social sentiment may have value as an indicator in technical analysis as a tool that could be used in conjunction with others to help predict reversals and gauge strength of trends. However, leveraging social sentiment in isolation, such as in the S&P 500 Twitter Sentiment Index, may not improve returns over long periods.

Works Cited

"6 in 10 Young Investors Use Social Media as a Source of Investment Information. Should You?" *Nasdaq.com*, 2021, [www.nasdaq.com/articles/6-in-10-young-investors-use-social-media-as-a-source-of-investment-information.-should-you. Accessed 4 Mar. 2023](www.nasdaq.com/articles/6-in-10-young-investors-use-social-media-as-a-source-of-investment-information.-should-you).

Azar, Pablo, and Andrew W. Lo. "The Wisdom of Twitter Crowds: Predicting Stock Market Reactions to FOMC Meetings via Twitter Feeds." *Papers.ssrn.com*, 11 Mar. 2016, papers.ssrn.com/sol3/papers.cfm?abstract_id=2756815. Accessed 19 Feb. 2023.

Leung, Woon Sau, et al. "Social-Media Sentiment, Portfolio Complexity, and Stock Returns." *Papers.ssrn.com*, 24 Nov. 2019, papers.ssrn.com/sol3/papers.cfm?abstract_id=3492722. Accessed 19 Feb. 2023.

Ruby, Daniel. "Social Media Users — How Many People Use Social Media in 2022." *Demandsage,* 4 Jan. 2023, www.demandsage.com/social-media-users/. Accessed 5 Mar. 2023.

"Social Sentiment Indicator." *Investopedia*, [www.investopedia.com/terms/s/social-sentiment-indicator.asp. Accessed 4 Mar. 2023](www.investopedia.com/terms/s/social-sentiment-indicator.asp).

Bhadamkar, Bhattacharya. "Tesla Inc. Stock Prediction using Sentiment Analysis" *ro.uow.edu.au*, 2022, ro.uow.edu.au/cgi/viewcontent.cgi?article=2283&context=aabfj Accessed 8 May 2023.

"S&P 500 Twitter Sentiment Index." *S&P Dow Jones Indices*, www.spglobal.com/spdji/en/indices/strategy/sp-500-twitter-sentiment-index/#overview. Accessed 8 May 2023.

"S&P 500 Twitter Sentiment Select Equal Weight Index." *S&P Dow Jones Indices*, www.spglobal.com/spdji/en/indices/strategy/sp-500-twitter-sentiment-select-equal-weight-index/#overview. Accessed 8 May 2023.