Minds and Markets: The Relationship between Investor Sentiment and Stock Market Returns

Kaitlyn Pan

Received March 17, 2024
Accepted August 21, 2025
Electronic access October 15, 2025

This study examines a sector of behavioral finance, which is the intersection of economics and psychology, and aims to identify the relationship between investor sentiment and the stock market, addressing gaps in research across timeframes and sentiment indicators. The results contribute to the debate on whether markets are influenced by biases or if they incorporate all available information. This paper uses the Standard & Poor's (S&P) 500 stock data, Bullish Sentiment and Bearish Sentiment from the American Association for Individual Investors (AAII) Sentiment Survey, and the July 1987 to December 2024 Consumer Confidence Index (CCI) values from the Organization for Economic Co-operation and Development (OECD). These indicators were selected because the AAII captures direct investor expectations about market performance, while the CCI shows a broader view of market sentiment. The correlation and linear regression tests between Bullish Sentiment and Bearish Sentiment changes showed weak relationships with monthly S&P 500 Returns, with a correlation coefficient (r) of -0.22 and 0.21, respectively. The CCI change had the strongest predictive value, with a statistically significant relationship to S&P 500 returns in yearly regressions (p-value = 0.01, R^2 = 0.16). However, none of the correlation coefficients exceeded 0.4, indicating limited economic significance. This study implies that while sentiment may create price distortions temporarily, it is not a reliable long-term investment strategy. This research helps show how consumer confidence impacts market behavior and provides a framework for future research to explore alternate timeframes and sentiment indicators.

Keywords: behavioral economics, behavioral finance, stock returns, sentiment

Introduction

Problem Statement and Rationale

The stock market plays a crucial role in economic stability, wealth accumulation, and investment strategies. When the stock market performs badly, it affects the well-being of individuals, corporations, and the overall economy[1]. When biases such as overconfidence, anchoring, and confirmation bias form together to create an investor's sentiment, stock market participants can become irrational and collectively create market fluctuations and irrational returns[2]. This study explores how combined biases (which culminate in investor sentiment) affect market performance as opposed to isolating individual biases. In addition, while investor sentiment has been linked to market volatility, there is still a gap in how sentiment (measured through multiple indexes) related to market returns across multiple timeframes. There is limited on how longitudinal sentiment data quantifies the magnitude and direction of this relationship.

Significance and Purpose

This research paper addresses these gaps by analyzing how investor sentimentwhich can combine multiple biasesshapes stock market returns. By using two measures of investor sentiment and empirical analysis over an extended period, this research aims to provide a better understanding of whether psychological factors drive market returns. These findings will be valuable for individual investors, households, and policymakers who can develop strategies to navigate investment and the stock market.

Objectives

This study aims to examine the relationships between Bullish Sentiment, Bearish Sentiment, and CCI values with S&P returns. The guiding question is "How does investor sentiment influence stock returns?" My research is grounded in two contrasting theories, the Efficient Market Theory (EMH) and Behavioral Finance Theories. According to the EMH, stock prices reflect all available information, and investor sentiment should not predict or impact future market returns[3]. In contrast, behavioral finance argues that psychological biases can lead to deviations from market efficiency[4]. Based on these theories, this study has two testable hypotheses: 'There will be no statistically significant relationship between investor sentiment indicators (Bullish %, Bearish %, and CCI) and future S&P 500 returns (EMH consistent)' and 'Higher levels of bullish sentiment or

© The National High School Journal of Science 2025

consumer confidence will positively predict future S&P 500 returns, while higher bearish sentiment will negatively predict returns (Behavioral Finance consistent.)'

Theoretical Framework

The EMH and Behavioral Finance theories provide different perspectives on the role of sentiment and psychological biases in financial markets and help inform my analysis of sentiment-driven stock market movements[4,5].

The EMH, introduced by Fama, claims that all available information is reflected in stock prices and that investors act rationally when making investment decisions[3]. Under this idea, investor sentiment should have no effect on stock returns, as markets efficiently process information through mathematical formulas. Additionally, price fluctuations and market movements would be purely driven by rational decision-making, and emotions, heuristics, and biases would not have an impact.

In contrast, Behavioral Finance theories challenge the EMH by integrating psychological biases into financial decision-making and markets. Researchers such as Kahneman, Tversky, Shiller, and Thaler have demonstrated that investor actions often diverge from rational expectations[4–6]. Therefore, cognitive biases affect investor sentiment, which then can lead to anomalies such as herding behavior, speculative bubbles, and overreactions. From this perspective, investor sentiment is highly relevant to market movement.

This study operates within both frameworks. If investor sentiment exhibits a strong correlation with stock returns, it could provide evidence against the EMH and support the behavioral finance theory that emotions and biases influence market movements. Alternatively, if significant relationships aren't found, then it could reinforce the EMH in that market prices are governed by rational factors and mathematical models. This study aims to assess whether investor sentiment is meaningful in determining market performance or if it is simply a byproduct of market noise.

Methodology Overview

This study uses quantitative analysis to examine the relationship between investor sentiment and stock returns by analyzing data from the American Association of Individual Investors (AAII) Sentiment Survey and the Consumer Confidence Index (CCI) from July 1987 to December 2024[7]. Correlation and Regression analysis are the main tests used. The AAII Sentiment Survey reflects the views of individual investors who are members of the AAII, which consists of older, wealthier people and is not representative of the broader investor population[7]. In addition, the CCI surveys a randomized range of households but doesn't just measure investor sentiment, which creates differences in analysis results[8]. Overall, AAII data reflects the market expectations of a more experienced investor population, while the CCI reflects broader economic sentiment. Therefore, the AAII sentiment might be more accurate while predicting stock market returns. Additionally, while there may be macroeconomic factors that could influence market returns, the time frame chosen was broad enough so that specific variables like interest rates or inflation could not influence the whole set of results.

Another potential constraint is the time periods of the data. This study looked at monthly and yearly data, but other time periods, such as weekly and quarterly data were not looked at. Despite these factors, this study provides a valuable foundation for further research into sentiment and market behavior. These datasets were chosen due to the large time frame of data collected, the reliability of their sources, and their established use. In the AAII Sentiment Survey, the variables analyzed were Bullish Sentiment, Bearish Sentiment, and S&P 500 Percentage Returns. To ensure the robustness of the findings, results were compared across multiple time horizons (monthly vs. yearly returns) and both the median splits and raw values were used for sentiment inputs. Tests were run on multiple variations of these variables, such as Above Median Bullish/Bearish Sentiment, Below Median Bullish/Bearish Sentiment, and Next Year Sentiment. Neutral Sentiment was not analyzed. This allowed the analysis to distinguish between periods of relatively high and low sentiment while minimizing the effect of outliers or irregular fluctuations in the raw percentages. The CCI data was also categorized into Above-Median, Below-Median, and Next-Year values for comparison. Statistical tests, including correlation and regression analysis, were used to assess these relationships. The study also uses a lag structure by comparing a year's variable results to the variable results of the year after to determine whether past sentiments influenced future market performance. This time frame allows sentiment effects to materialize while avoiding the dilution of influence that might occur with longer lag periods. Effect sizes from the data tests were interpreted in the context of real world market movements. While some relationships were modest, even small changes in sentiment can drive significant market movements. However, it is also important to consider the size of the correlation and p-value. The limitations in my survey include the differences between the AAII Sentiment survey and the CCII, which can influence the data tests differently.

Literature Review

The field of economics has been consistently evolving through the addition of new market theories and ideas. Traditional economic theories, such as the Expected Utility Theory, assume that investors make rational decisions to maximize their utility[9]. Similarly, the Efficient Market Hypothesis proposes that asset prices fully reflect all available information, implying that markets are rational and efficient[3].

© The National High School Journal of Science 2025

However, repeated financial crises, such as the dot-com bubble and the 2008 financial crisis, indicate that investors show irrationality and that traditional market theories can not fully explain changes in the market. These developments have led to the field of behavioral finance, which integrates psychology and economics and assumes that there are behavioral biases that prevent individuals from being completely rational. Kahneman and Tversky's paper on Prospect Theory laid the foundation for behavioral finance by showing how people weigh gains and losses differently, leading to risk-averse behavior when facing gains and risk-seeking behavior when facing losses[4]. Shiller introduced the concepts of investor overconfidence and herd mentality and concluded that these biases led to speculative bubbles[5]. His research provided evidence that the market was driven by psychological factors rather than efficient market movements.

Richard Thaler's seminal paper on mental accounting & consumer choice in 1985 highlighted how investors don't always calculate rationally, and treat money differently based on their own thoughts and biases[6]. It also introduced the Transactional Utility Theory, which states that people compare prices to a reference point and get satisfaction from the perceived quality of the deal that they get. Published in the same year, Shefrin and Statman's paper on the Disposition Effect showed that investors tend to hold on to losing stocks too long and sell their profiting stocks too quickly, adding to previous theories on loss aversion and risk-seeking behavior[10]. Robert Shiller's paper on excess volatility provided evidence that stock prices fluctuate more than what can be explained by rational behavior and traditional economic theories, suggesting that investor sentiment and behavioral biases play a role[11]. It challenged the Efficient Market Hypothesis by saying investor behavior and sentiment contribute to price movements. In De Bondt and Thaler's paper on the overreaction hypothesis, they show that markets tend to overreact to news due to collective sentiment, leading to herding[6]. De Bondt and Thaler stated that herding is a phenomenon where people make decisions based on other people's decisions rather than their own independent analysis, leading to collective behavior. Overall, there have been many pieces of literature that have solidified behavioral finance as an integral and solid factor in the stock market. However, literature is scarce on showing exactly which kinds of biases lead to different shifts in the market, such as specifically saying that overconfidence caused one instance of fluctuation, and herding behavior for another.

Investor Sentiment, a key part of behavioral finance, reflects the overall mood or attitude of investors toward the stock market[12]. It is a collective attitude towards the market or economy based on various cognitive biases that investors might hold. Sentiment is different than mood or attitude in that while sentiment is the collective feeling of a group of people, mood or attitude is typically more temporary and personal. Sentiment is typically categorized as Bullish (optimistic), Bearish (pessimistic), or Neutral and can have strong impacts on stock movements and volatility. Multiple factors, like weather and mood, can impact investor sentiment, which consequently affects market behavior[13]. Additionally, influxes of positive sentiment will drive up prices and contribute to different market anomalies[14]. Fisher and Statman studied the sentiment of three different groups of investors, small (individual), medium (newsletter writers), and large investors (Wall Street Strategists), to show that the sentiments of different groups of people were often different, especially between small and large investors[15]. Individual investor sentiment was found to be a contrary indicator for S&P performance, with a negative relationship and high statistical significance. Brown and Cliff argued that high investor sentiment drives stock prices above fundamental values, leading to low returns as the assets return to their rightful values, suggesting a negative sentiment-return relation[16]. Their idea was similarly proven by Schmeling across a study of 18 different countries while analyzing a wide range of stocks, including value stocks, growth stocks, and small stocks[17].

Overall, existing sentiment literature mainly shows a negative relationship between positive sentiment and stock returns, suggesting that the existence in sentiment drives up stock prices to irrational values. However, there is less research done on the magnitude of these effects, and the results are often impacted highly by the context of the study done, such as the type of investor, geographical aspects, and timeframe.

Methods

Research Design

My research study is observational, and I conducted analysis on two data sets: The American Association for Individual Investors (AAII) Sentiment Survey and the Consumer Confidence Index (CCI). My study was based on time series analyses, since data was collected on the AAII Sentiment survey and the CCI over a period of time. data sets were used for a more holistic view, and to compare differences in results.

0.1 Participants or Sample

The dataset includes 37 years (1987-2024) of yearly data and 449 months of monthly data. The AAII Sentiment Survey is an indicator of individual investor sentiments towards the future stock situation. To take the survey, you must be an AAII member, which costs $74 for a one-year membership and $390 for a lifetime membership[18]. According to AAII, the typical AAII member is a male in his mid-60s with a bachelor's or graduate degree, has a median portfolio size of more than $1 million, and describes themselves as having a moderate level of investment knowledge and engaging primarily in fundamental analysis[7]. The second data set that was analyzed was the CCI, which is a standardized confidence indicator used to forecast future

developments of the consumption and savings of households. The index is randomly sent to 5,000 households and is based on answers regarding household unemployment, sentiment about the general economic situation, expected financial situation, and capability for savings [12]. The differences in these two indexes dilute the data results, as they have different survey audiences and different questions asked.

Data Collection

The data used was collected through surveys. The AAII Survey asks individual investors on question: "Do you feel the direction of the stock market over the next six months will be up (Bullish/optimistic), no change (neutral), or down (Bearish/pessimistic)?" Data collection started in 1987 and is asked weekly from Thursday at 12:01 a.m. to Wednesday at 11:59 pm. AAII has approximately 170,000 members, and a response rate of around 300 members on average every week [19]. The data is added every week to an Excel spreadsheet that is publicly available. The CCI is also conducted through an online survey. The questions that are asked are split into two indexes. The Present Situation Index makes up the first two questions and the Expectations Index makes up the last three questions. The questions assess:

1 Respondents' appraisal of current business conditions

2 Respondents' appraisal of current employment conditions

3 Respondents' expectations regarding business conditions in six months

4 Respondents' expectations regarding employment conditions in six months

5 Respondents' expectations regarding their total family income in six months

The five survey questions each have three response options: positive, negative, or neutral. The CCI has approximately 3000 responses weekly and is conducted on a monthly basis by the conference board. The CCI is sent out to 5,000 random households each month, receiving around 3,000 responses on average [20].

Variables and Measurement

The independent variables I looked at in the AAII Investor Sentiment data set were the percentage of investors that were Bullish and the percentage of investors that were Bearish. The independent variable for the CCI survey was the CCI value. Bullish sentiment, Bearish Sentiment, Change in CCI, and Change in S&P are reflected in Figure 2. The dependent variable for both data sets was the S&P 500's percentage return. AAII is calculated by finding the percentage of responses that were Bearish, Bullish, and neutral. The CCI is calculated by separating each of its questions into a portion called the "relative value." Each question's positive responses are divided by the sum of its positive and negative responses. The relative value for each question is then compared with the relative value of 1985 since the CCI is benchmarked against 1985. From the comparison, we get an index value, and then the index values for all five of the questions are averaged together to form the CCI.

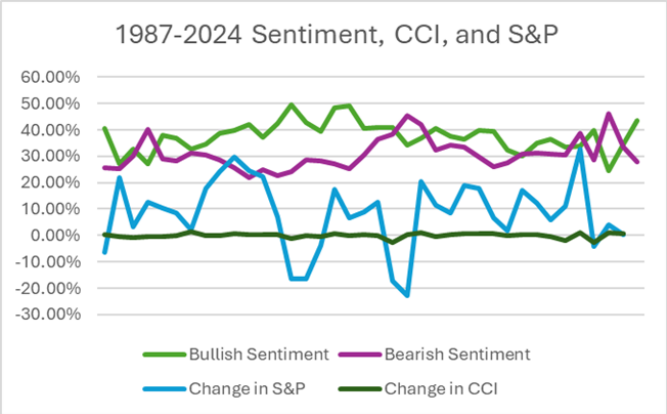

Fig. 1 1987-2024 Bullish Sentiment, Bearish Sentiment, S&P percentage change, and CCI percentage change.

Procedure

Both data sets were analyzed using Excel in a monthly data version and a yearly data version between July 1987 and December 2024. In the AAII data set, I analyzed monthly and yearly data of the independent and dependent variables to test for correlations and statistical significance between them. The following is the procedure for the AAII Sentiment Survey:

1 A Pivot table was used on the Weekly data to find monthly averages for Bullish sentiment, Bearish sentiment, and S&P Returns.

2 The change in S&P in percentage value was calculated through the formula (next year price-current year price)/current year price.

3 In order to test a lag structure, a column was created for "Next year's S&P Above Median Return".

4 The Median was calculated for Bullish sentiment, Bearish sentiment, and S&P percentage return.

5 Columns in binary numbers for Above Median Bearish years, Above Median Bullish Years, Above Median S&P Percentage Return, Below Median Percentage Median, and Below-Median S&P percentage return were created

© The National High School Journal of Science 2025

through the function =IF('year 1 value>median for column', 1, 0) and =IF('year 1 value'<' median for column', 1, 0).

6 Columns for "Percent Change for Bullish Sentiment," "Percent Change for Bearish Sentiment," and "Percent Change for Next Year's S&P Return" were also created.

7 A correlation test using the Excel function =CORREL was performed. Then, a t-value equation "$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$" was used, with r being the correlation value and n being the number of pairs used. Finally, a p-value test was conducted using the function =TDIST(ABS('t-value', 'n-2, 2).
a. This process was repeated for above median Bullish sentiment and above median S&P percentage return, above median Bullish sentiment and next year above median &P percentage return, above median Bearish sentiment and next year above b. Median S&P percentage, above median Bearish and above median S&P percentage return, below median Bullish sentiment and below median S&P percentage, S&P percentage change and change in Bullish sentiment, and S&P percentage change and change in Bearish sentiment.

8 Linear regression using the function "=LINEST" was to find the slope and intercept for the pairs Bullish sentiment and S&P monthly percent change, Bearish sentiment and S&P monthly percent change, change in Bearish sentiment and S&P monthly percentage change, percent change in Bullish sentiment and S&P percentage change, S&P percent change on both Bullish and Bearish sentiment, and S&P percentage change on both change in Bullish and change in Bearish sentiment.

9 Lastly, the r^2 value was found using the "=RSQ" value.

A similar procedure with the same functions was followed for the CCI, with only the independent variable changing.

1 The median was found for the CCI, and the S&P percentage change.

2 Columns for above median CCI, above Median S&P percentage change, next year above median S&P percentage change, below median CCI, below median S&P percentage change, next year below median S&P change, and Change in CCI were created.

3 The "=CORREL" Excel function was used to find the correlation number between above median CCI and above median S&P percentage change, above median CCI and next year above median S&P percentage change, and below-median CCI and next year below median S&P percentage change using the same equation format as the AAII sentiment survey. The t-value and p-value, linear regression values, and r^2 values were also found using the same equations and format.

Data Analysis

Data analysis was conducted using statistical tests to determine the relationship between investor sentiment, consumer confidence, and stock market performance.

Correlation Analysis, T-Value, and P-Value

To measure the strength and direction of relationships between sentiment (AAII Bullish and Bearish percentages, CCI values) and S&P 500 percentage returns, the correlation coefficient (r) was calculated using Excel's "=CORREL" function. To determine statistical significance, a t-value was found using the formula "$t = \frac{r(n-2)}{1-r^2}$", with n being the number of pairs examined and r being the correlation coefficient. Then, the p-value was determined using the "=TDIST" function to evaluate the probability of obtaining the found correlation. Correlation values were interpreted based on the guidelines that $| r | > 0.5$ indicated a statistically significant correlation, $0.3 \leq | r | < 0.50$ indicated a moderate relationship, and $| r | < 0.3$ | indicated a weak relationship. p-values were interpreted through a p-value less than 0.05 being strongly statistically significant, a p-value between 0.05 and 0.1 being moderately significant, and a p-value more than 0.1 being weak and not statistically significant.

Regression Analysis

Linear regression tests were used in addition to correlation tests to quantify the effect of sentiment on returns, in addition to the relationship and strength of the relationships already found. This study used the "=LINEST" function to perform Linear regression in Excel to analyze the relationship between sentiment indicators and S&P 500 returns. The regression output included the slope and intercept values, which were then used to find the extent to which investor sentiment and consumer confidence influenced stock markets. The coefficient of determination (r^2) was then calculated using the Excel function "=RSQ" to measure the power of sentiment indicators on stock returns.

Lag Structure Analysis

In addition to conducting current year and current year's statistics, a lagged analysis was conducted by using next year's S&P 500 percentage return as a dependent variable. This enabled us to see if the previous year's sentiment or CCI would influence next year's statistics.

Ethical Considerations

My study doesn't have any ethical concerns since the surveys are done with the consent of the responders, and the data is made publicly available.

Data Quality

Both the AAII Sentiment Survey and the CCI are widely used for economic and financial research. Both datasets contain several decades worth of data and use fixed question formats. However, there is a difference in the outcomes of the AAII Sentiment Survey and the CCI since there is a difference in the surveys themselves. The AAII asks one question, while the CCI asks 5 and is much broader. Additionally, the AAII only asks about investor sentiment, while the CCI asks about other information, such as expectations of employment, family income, and future business conditions.

Results

The results from the data tests have been compiled into four tables below. Tables, as opposed to other visualizations such as scatter plots or time series plots, were used because the data consists of many variable combinations with single summary values. Visual trends wouldn't be as meaningful, so tables are the most effective way to present the results clearly. The first table contains statistics that give context to the overall numbers in the data tests. The next three tables summarize the relationships between investor sentiment, the CCI, and S&P 500 returns across both yearly and monthly timeframes.

The average AAII Bullish sentiment was 37.49% on a yearly basis and 37.74% on a monthly basis, while Bearish sentiment averaged 30.92% and 30.93%, respectively. The CCI average stayed the same across both timeframes at approximately 100.19. The S&P 500 average return over the sample period was 8.65% annually and 0.18% monthly, with median returns slightly higher than the mean in both cases. The number of years in which the S&P 500 return was above the mean return was 19, while it was below the mean in 18 years (see Table 1). The differences between monthly and yearly data could support reasoning for why monthly and yearly data results differed.

Table 2 presents the correlation results between AAII Sentiment and S&P 500 returns at both the yearly and monthly levels. Apart from S&P Returns and Bullish % Change and S&P Returns and Bearish % Change, the p-values of the yearly tests are either weakly significant or not significant at all since they are more than 0.05 (p-value¡0.01 = strongly statistically significant, 0.05>p-value>0.01 = statistically significant, p-value> 0.1 = not statistically significant). A low p-value suggests that the observed pattern is unlikely to be due to chance. These cutoffs were applied consistently across all correlation and regression tests to evaluate statistical significance. At the yearly level, S&P and Bullish % change show a slightly strong positive correlation, meaning that as Bullish Sentiment increases, so do S&P Returns. S&P Returns and Bearish sentiment also show a slightly strong negative correlation, and both correlations are statistically significant since they have a p-value of less than 0.05. For the monthly data, the lag structure did not seem to create a stronger statistical significance between Above Median Bullish + Next Year Above Median S&P and above Median Bearish + Next Year Above Median S&P, meaning that this year's sentiment doesn't have a strong relationship with next year's S&P Returns. On the monthly level, above Median Bearish and Next Year Below Median S&P, S&P and Bullish % Change, and S&P + Bearish % Change had the strongest correlation levels (0.24, 0.22, and 0.21, respectively). While the correlations are modest in size (below 0.5), their statistical significance indicates that the relationships are unlikely to be due to chance.

Table 1 AAII Sentiment Survey Data Analysis Results.

Summary Statistics	Yearly	Monthly
Years of Sample	1987-2024	7/1987-12/2024
AAII Average Bullish Sentiment	37.49%	37.74%
AAII Median Bullish Sentiment	37.50%	37.48%
AAII Average Bearish Sentiment	30.92%	30.93%
AAII Number of Observations	37	449
CCI Average	100.19	100.18
CCI Average Change	-0.03%	-0.01%
CCI Number of Observations	37	449
S&P Average Return over sample period	8.65%	0.18%
S&P Median Return over sample period	9.57%	0.30%
S&P Number of Yearly Returns> Mean Return	19	244
S&P Number of Yearly Returns ¡ Mean Return	18	205

The tests run for the CCI data were shorter because there was only one variable (CCI value) to make combinations rather than two (Bullish Sentiment and Bearish Sentiment). The results for the CCI tests were all statistically insignificant, with p-values far exceeding 0.05 (the threshold for statistical significance) or even 0.1 (the threshold for weak statistical significance). The highest correlations were seen between below median CCI and next year below median S&P, and above median CCI and next year above median S&P, which were the combinations where lag structure was used. However, without statistical significance, the observed

© The National High School Journal of Science 2025

Table 2 AAII Sentiment Survey Data Analysis Results. In the table, above is abbreviated as Abv, below is abbreviated as Bel, median is abbreviated as Mdn, year is abbreviated as Yr, and the S&P refers to the percentage change in the S&P 500.

Variable Combination	Monthly Correlation	Monthly t-value	Monthly p-value	Yearly Correlation	Yearly t-value	Yearly p-value
Abv Mdn Bullish + Abv Mdn S&P	0.18	3.92	0	-0.11	-0.66	0.51
Abv Mdn Bullish + Next Yr Abv Mdn S&P	0.05	1.15	0.25	-0.3	-1.85	0.07
Abv Mdn Bearish + Next Yr Abv Mdn S&P	0.02	0.5	0.62	0.24	1.48	0.15
Abv Mdn Bearish + Next Yr Bel Mdn S&P	-0.24	-5.33	0	-0.15	-0.88	0.39
Abv Mdn Bearish +Abv Med S&P	-0.12	-2.65	0.01	0.17	1	0.32
Bel Mdn Bearish+ Bel Mdn S&P	0.13	2.75	0.01	0.03	0.17	0.86
Bel Mdn Bullish + Bel Mdn S&P	-0.19	-4.02	0	-0.08	-0.47	0.64
S&P and Bullish % Change	-0.22	-4.73	0	0.37	2.32	0.03
S&P + Bearish % Change	0.21	4.52	0	-0.35	-2.2	0.03

Table 3 CCI Data Analysis Results. In the table, above is abbreviated as Abv, below is abbreviated as Bel, median is abbreviated as Mdn, year is abbreviated as Yr, and the S&P refers to the percentage change in the S&P 500.

Variables	Yearly Correlation	Yearly t-value	Yearly p-Value	Monthly Correlation	Monthly t-Value	Monthly p-Value
Bel Mdn CCI and Next Yr Bel Mdn S&P	0.16	0.93	0.36	-0.03	-0.66	0.51
Abv Mdn CCI and Next Yr Abv Mdn S&P	0.11	0.62	0.54	-0.03	-0.66	0.51
Bel Mdn CCI and Bel Mdn S&P	0.05	0.31	0.76	0.02	0.38	0.71
CORREL Abv Med CCI and Abv Med S&P %	0	0	1	0.02	0.38	0.71

correlations are likely due to random market fluctuations. There were no significant correlations (correlation > 0.5) for any of the variable combinations on both the yearly and monthly level.

The regression analysis assesses the relationship between Bearish and Bullish sentiment and S&P returns on both a monthly and yearly basis. Table 4 shows relatively high regression values compared to the weak correlation values in Table 3. This is because regression can detect predictive relationships between variables that correlation might miss, especially when the independent variable has a meaningful linear effect on the dependent variable but is embedded in noisy data. As for results, linear regression for monthly S&P Returns on CCI Change showed strong statistical significance with a high T-statistic of 8.44, a low P-value of 0.00, and an R-squared of 0.1359, meaning that my CCI change explains 13.6% of the variation in S&P returns. Although the standard error was higher than 0.30, it doesn't necessarily weaken the relationship since the t-statistic and p-value are still strong. It does suggest some caution as there might be more variability and fluctuations. Similarly, linear regression for yearly S&P returns on CCI change also had the strongest correlation, with strong t-statistics, p-values, and r-squared values. Again, the standard error is also higher. For yearly data, the relationships have weaker correlation and less statistical significance. Bullish change and Bearish change variables are statistically significant (P-values = 0.03 and 0.04, respectively), but their R-squared values are lower at approximately 0.1. On the other hand, Bullish and Bearish sentiments did not have strong p-values (0.26 and 0.43, respectively). The adjusted R^2 values across all regression models are relatively low, ranging from 0.01 to 0.14. This indicates that while some variables (such as CCI change) are statisticaly significant predictors of S&P 500 returns, they only explain a small portion of the total variance in returns. The low effect sizes suggest that while lagged sentiment measures may show some association with future returns, they are not strong or reliable predictors. Therefore, their predictive value should be taken with caution. This study did not include residual analysis, multicollinearity, or verify regression assumptions. Additionally, given the many variable combinations that were used in the tests (above median, below median, percent change, lagged), a summary table

Table 4 Linear Regressions on both AAII Variables and CCI variables. In the linear regressions, Sentiment and CCI values were the independent variables, and S&P Returns were the dependent variables. In the table, above is abbreviated as Abv, below is abbreviated as Bel, median is abbreviated as Mdn, year is abbreviated as Yr, and the S&P refers to the percentage change in the S&P 500.

Variables	Coefficient	Standard Error	T-statistic	p-value	Adjusted R Squared	Observations
Monthly S&P Returns on Bullish Change	-0.01	0	-4.73	0	0.05	449
Monthly S&P Returns on Bearish Change	0.01	0	4.52	0	0.04	449
Monthly S&P Returns on Bullish Sentiment	0.01	0.01	2.41	0.02	0.01	449
Monthly S&P Returns on Bearish Sentiment	-0.02	0.01	-3.42	0	0.02	449
Monthly S&P Returns on CCI Change	2.34	0.28	8.44	0	0.14	449
Monthly S&P Returns on CCI Index	0	0	0.68	0.5	0	449
Yearly S&P Returns on Bullish Change	0.28	0	2.29	0.03	0.11	36
Yearly S&P Returns on Bearish Change	-0.26	0.12	-2.17	0.04	0.1	36
Yearly S&P Returns on Bullish Sentiment	-0.45	0.4	-1.14	0.26	0	36
Yearly S&P Returns on Bearish Sentiment	-0.3	0.37	-0.79	0.43	0	36
Yearly S&P Returns on CCI Change	6.23	2.25	2.77	0.01	0.16	36
Yearly S&P Returns on CCI Index	0	0.02	-0.01	0.99	-0.03	36

may understate or overstate the importance of a certain variable combinations if certain variable combinations were included or excluded. Instead, the Conclusion states the key findings and distinctions between monthly and yearly outcomes.

Conclusion

Restatement of Key Findings

This study analyzed the relationship between investor sentiment, as measured by the AAII Sentiment Survey and the Consumer Confidence Index, and stock market returns, represented by S&P 500 percentage changes. The correlations between Bullish Sentiment, Bearish sentiment, and CCI values with S&P returns present weak to moderate correlations with either high statistical significance or no substantial correlation. The strongest predictor of S&P 500 returns, according to linear regression tests, was the CCI change, with both monthly and yearly regression results showing significant relationships. However, CCI change and S&P returns are weak to moderate since their correlation values are less than 0.5.

Implications and Significance

These findings help contribute to the fields of both market efficiency and behavioral finance. My findings support the EMH

© The National High School Journal of Science 2025

more than they support behavioral finance theories since they indicate that long-term sentiment does not have an impact on stock market returns. However, there are weak but statistically significant correlations to behavioral impacts in the short term. Compared to previous studies, these results both build upon and challenge existing findings. For instance, Brown and Cliff found that high sentiment drives returns below fundamental value, while Fisher and Statman observed a negative relationship between individual sentiment and S&P 500 performance.17, 16, My results also differ from Schmeling, who found an international pattern of sentiment predicting stock returns.18 My data, which focused on US sentiment, showed that the same strength of predictability was not observed. This suggests that geographical, dataset, and investor populations can all contribute to variability in observed effects, and can explain why this study did not predict market One possible reason why CCI change showed slightly more predictive power is that consumer confidence may incorporate broader macroeconomic signals, such as expectations around employment or inflation, which can more directly affect stock market activity. However, these influences may be short-lived or easily overtaken by more fundamental economic drivers.

Connection to Objectives

My research objectives were to determine the relationships between Bullish sentiment, Bearish sentiment, CCI values, and S&P 500 returns. The study found some evidence for correlations, such as between Above median Bullish sentiment and the next year's S&P returns. However, none of the correlation values made it past 0.3, implying that all the variables were weakly or not at all correlated with each other. Additionally, many of my results were not statistically significant.

Recommendations

Since short-term results seemed to provide more significant results, time periods such as weekly and quarterly data could be assessed. Other behavioral indicators can also be studied more in depth, such as media sentiment or other sentiment surveys. Additionally, this study only examined investor sentiment, which can consist of multiple behavioral biases. The ways in which specific biases (mental accounting, herding, overconfidence, etc.) affect the stock market can be investigated more. Although previous research has explored this area, isolating a single bias as the cause of a specific market fluctuation remains challenging. Investors should also avoid overreliance on aggregate sentiment indexes when making long-term portfolio decisions, as their predictive power is weak. However, short-term traders, market analysts, and financial advisors could use sentiment trackers as short-term tools to understand how crowd behavior influences returns temporarily. My study provides a foundation for future work by providing evidence for the correlation and significance of the relationship between investor sentiment and stock returns.

Limitations

The acknowledged limitations help explain the lack of strength of the observed relationships. For example, the AAII Sentiment primarily reflects the views of individual investors, and as shown in the demographics of the survey, primarily men with a median portfolio size of more than $1 million dollars. The sentiment of individual investors will vary from the sentiment of other groups of people, such as institutional investors. The CCI, while it captures a more general view of consumer sentiment, also asks about other things that are not limited to investor sentiment. This dilutes the data results and potentially weakens them, as the AAII Survey is solely targeted to individual investors, and the CCI is targeted toward the whole economy. Another limitation in my study is the reliance on bivariate analysis, where relationships were only tested between two variables at a time. While the correlations and regressions show statistical relationships, they do not account for other macroeconomic factors that may be influencing the variables. This could weaken the link between sentiment and returns by omitting key explanatory variables. Additionally, given the bivariate focus and the use of p-values and r-squared values to assess significance and relationship strength, confidence intervals were not included, although they can offer insight into estimate precision. Future research can address this limitation by incorporating multivariate regression models. There are also several volatile periods in the market that could have diluted the results, as shown in Figure 1. Finally, although other statistical methods such as GARCH models or vector autoregression could have been used, the methods chosen are appropriate for identifying broader sentiment-return relationships over time, and maintain interpretability and conciseness.

Fig. 2 A chart of volatile movements in the US stock market from Manchester Capital Markets [1]

Closing thoughts

Investor Sentiment, behavioral biases, and consumer bias play an undeniable role in shaping market behavior. However, in the data sets analyzed, their effects appear to be more pronounced short term and weak. This study highlights the relationship between psychological biases and sentiment while analyzing stock market returns. As financial markets evolve, integrating behavioral insights with traditional economic theories will be crucial in understanding a more nuanced approach to understanding and predicting market movements.

Acknowledgments

Thank you to Professor Joshua Pierce for supporting my research process.

References

1 Manchester Capital Management, *Market Resilience*, https://mcmllc.com/market-resilience/, 2021.

2 D. Chatterjee and K. Nayyar, *SSRN*, 2024, **2**, year.

3 E. Fama, *The Journal of Finance*, 1970, **25**, 383–417.

4 D. Kahneman and A. Tversky, *Econometrica*, 1979, **47**, 263–291.

5 R. Shiller, *Irrational Exuberance*, Princeton University Press, 2000.

6 R. Thaler, *Marketing Science*, 1985, **4**, 199–214.

7 American Association of Individual Investors, *AAII Sentiment Survey*, https://www.aaii.com/sentimentsurvey/sent_results, 2025.

8 Organization for Economic Co-operation and Development (OECD), *Consumer Confidence Index (CCI)*, https://www.oecd.org/en/data/indicators/consumer-confidence-index-cci.html?oecdcontrol-b2a0dbca4d-var3=1974-06&oecdcontrol-b2a0dbca4d-var4=2024-12, 2025.

9 J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, 1944.

10 H. Shefrin and M. Statman, *Journal of Finance*, 1985, **40**, 777–790.

11 R. Shiller, *American Economic Review*, 1981, **71**, 421–436.

12 Investopedia, *Consumer Confidence Index (CCI)*, https://www.investopedia.com/terms/c/cci.asp, 2025.

13 M. Schneider, *Journal of Behavioral Finance*, 2014, **15**, 120–135.

14 A. Gormus and F. Gunes, *Applied Econometrics and International Development*, 2010, **10**, 103–114.

15 K. Fisher and M. Statman, *Financial Analysts Journal*, 2000, **56**, 16–23.

16 G. W. Brown and M. T. Cliff, *The Journal of Business*, 2005, **78**, 405–440.

17 M. Schmeling, *Journal of Empirical Finance*, 2009, **16**, 394–408.

18 American Association for Individual Investors, *AAII Membership Details*, https://invest.aaii.com/shop/products/#:~:text=Annual%20fee%20%2429.,Includes%203%20FREE%20Investor%20Guides!&text=Year%20AAII%20Membership-,Enjoy%20your%2030%2Dday%20%241%20trial,4%2DYear%20fee%20%2499, n.d.

19 American Association of Individual Investors, *Analyzing the AAII Sentiment Survey Without Hindsight*, https://www.aaii.com/journal/article/analyzing-the-aaii-sentiment-survey-without-hindsight?printerfriendly=true, n.d.

20 Investopedia, *Understanding the Consumer Confidence Index*, https://www.investopedia.com/insights/understanding-consumer-confidence-index/, 2023.

© The National High School Journal of Science 2025