

Industry-Level Sentiment and Stock Returns

Thi Thuy Trang Truong*

January 11, 2026

Abstract

This paper studies how industry-level investor sentiment affects industry returns. Industries with high investor sentiment outperform those with low investor sentiment by 8.48% per year. This outperformance is not explained by existing predictors of industry returns, is most pronounced when institutional ownership and analyst coverage are low, persists for up to two quarters, and generalizes to alternative measures of industry-level sentiment. Across industries, customer industry sentiment predicts supplier industry sentiment and, ultimately, customer industry returns. Taken together, the results suggest that industry-level investor sentiment is incorporated gradually into asset prices, in contrast to the negative relation between market-level sentiment and aggregate market returns documented in prior literature.

Keywords: Industry-Level Sentiment, Return Predictability, Sentiment Spillovers, Customer-Supplier Relationships

EFM Classification Codes: 320, 330, 370

*ESCP Business School, 6-8 avenue de la Porte de Champerret, 75017 Paris, France. Email: thi_thuy_trang.truong@edu.escp.eu.

I am profoundly thankful to my thesis advisor, Paul Karehnke, whose guidance and support have been invaluable to the development of this work. I sincerely appreciate the valuable feedback received from Linda Allen, Martijn Boons, Feng Dong, Fredj Jawadi (discussant), Hiep Manh Nguyen, Stanislava Nikolova, Lei Zhao, and seminar participants at ESCP Internal Finance Seminar, Nova SBE PhD Research Group, Foreign Trade University (HCMC), the 41st International Conference of the French Finance Association (AFFI) PhD Workshop, the FMA 2025 European Doctoral Consortium, and the EFMA 2025 Doctoral Seminar. Part of this research was conducted while I was visiting Nova School of Business and Economics, whose hospitality is gratefully acknowledged.

1. Introduction

Understanding the drivers of industry returns is essential for investors and corporations, as industries capture not only systematic risks but also present unique sector-specific opportunities. Existing research demonstrates that industry returns exhibit persistence and predict broader market movements (Hong, Torous, and Valkanov, 2007; Moskowitz and Grinblatt, 1999). Industries also serve as fundamental building blocks for diversification, portfolio construction (Chan, Karceski, and Lakonishok, 1999), and cost of capital estimation (Fama and French, 1997). Prior research identifies several determinants of industry returns, including expected inflation (Boudoukh, Richardson, and Whitelaw, 1994), forecast demand changes (DellaVigna and Pollet, 2007), industry concentration (Bustamante and Donangelo, 2017; Corhay, Kung, and Schmid, 2020; Hou and Robinson, 2006), and intra-industry value spreads (Bustamante, 2015). However, investor sentiment at the industry level receives little attention, and its role in explaining variations in industry returns remains unexplored. Given its well-documented role in explaining stock-level returns and the distinctive economic characteristics of industries, sentiment represents a plausible and important factor in explaining variations in industry returns.

In this paper, I investigate whether industry-level investor sentiment can predict industry returns and examine the sentiment spillover effects among customer-supplier industries. Studying industry-level sentiment is particularly important due to its potential differences from market-wide sentiment. At the aggregate market level, high investor sentiment typically leads to lower future returns, as overly optimistic investors drive prices above fundamental values, followed by subsequent corrections (Baker and Wurgler, 2006; Huang, Jiang, Tu, and Zhou, 2015). However, sentiment at the industry level may behave differently, potentially reflecting early signs of sector-specific growth that take time to manifest in broader market indicators. This distinction is critical because industry-specific sentiment can capture local-

ized economic conditions, supply chain dynamics, or technological advancements that are not fully reflected in aggregate market sentiment. For example, Bustamante (2015) and Bustamante and Donangelo (2017) emphasize that firms' returns depend not only on their own actions but also on the behavior of their industry peers. Menzly and Ozbas (2010) further show that returns in supplier and customer industries exhibit strong cross-predictability. These findings suggest that industry-level sentiment may influence stock returns through different mechanisms than market-level sentiment.

To address the research question, I first construct the sentiment index for each industry using five sentiment proxies from Baker and Wurgler (2006): the number of initial public offerings (IPOs), first-day returns of IPOs, the equity share in new issues, the dividend premium, and the closed-end fund discount. In their paper, Baker and Wurgler (2006) apply principal component analysis (PCA), where the sentiment index is derived as the first principal component of the correlation matrix of these proxies. While PCA captures the common variation in sentiment, Huang et al. (2015) argue that it may incorporate substantial approximation errors and variations in the proxies unrelated to true investor sentiment, since these errors are part of the common variation extracted by PCA. To address this issue, Huang et al. (2015) use partial least squares (PLS), a technique originally developed by Wold (1966, 1975) and later applied in finance by Kelly and Pruitt (2013, 2015). PLS extracts the common component from the proxies while explicitly maximizing its covariance with future stock returns, thereby reducing the influence of irrelevant approximation errors and enhancing the predictive power of the index. Because of these advantages, I follow Huang et al. (2015) and use PLS to construct the sentiment index for each industry.

The indices are generated using an out-of-sample approach with an expanding window, ensuring that only information available up to time t is incorporated. This methodology addresses concerns about data snooping and ensures the indices can be implemented in real time. The sample consists of all common stocks (share codes 10 and 11) listed on the New

York Stock Exchange (NYSE), NASDAQ, and NYSE American (formerly AMEX) from 1993 to 2024, excluding stocks priced below five dollars. Industries are classified according to the Global Industry Classification Standard (GICS), which comprises 11 industries.

The empirical results demonstrate that industries with high investor sentiment outperform those with low sentiment by 8.48% per year, a pattern that persists for up to two quarters. This relationship remains statistically and economically significant even after controlling for industry momentum, industry concentration, institutional ownership, and analyst coverage. The effect is particularly pronounced in industries with low institutional ownership or analyst coverage, where information asymmetry allows sentiment-driven mispricing to persist. The findings are robust to alternative measures of industry-level sentiment. Moreover, I also document that customer industry sentiment predicts supplier industry sentiment and, in some cases, supplier industry returns. This evidence highlights the importance of sentiment spillovers across economically linked sectors. Collectively, these findings suggest that industry-level sentiment is incorporated gradually into asset prices, unlike the negative relationship between market-level sentiment and aggregate market returns found in prior literature. The gradual diffusion of sentiment at the industry level aligns with the theoretical models of limited attention and slow information processing proposed by Hong et al. (2007). Additionally, the spillover effects between customer and supplier industries emphasize the role of economic linkages in shaping return predictability.

This paper makes three key contributions to the literature on behavioral finance, asset pricing, and corporate finance. First, it is the first study to introduce an industry-level sentiment measure. Since industries exhibit distinct economic characteristics, their sentiment dynamics often differ from aggregate market-wide measures, making industry-specific sentiment analysis essential. Second, this paper investigates the unexplored relationship between industry-level sentiment and industry returns, providing empirical evidence that industry-level sentiment significantly explains cross-sectional variation in returns. This finding high-

lights how mispricing and behavioral biases shape investment decisions at the industry level, offering new insights into the drivers of sector-specific performance. Third, by exploring the dynamic transmission of sentiment between economically linked supplier and customer industries, this study documents a predictable lead-lag relationship in sentiment spillovers. Understanding how sentiment propagates through supplier-customer networks helps firms anticipate demand shocks and make more informed capital investment decisions.

Related Literature This paper is primarily related to the literature on return predictability and sentiment. While extensive research has explored factors explaining stock-level returns (Chen and Velikov, 2023; Feng, Giglio, and Xiu, 2020; Green, Hand, and Zhang, 2017; Harvey, Liu, and Zhu, 2016; Hou, Xue, and Zhang, 2020; Kelly, Malamud, and Pedersen, 2023), the predictability of industry returns remains relatively understudied despite evidence that industry-level factors generate substantial profits (Moskowitz and Grinblatt, 1999) and that industry portfolios can predict broader market movements (Hong et al., 2007). Existing studies on industry return predictors have focused on fundamental factors such as expected inflation (Boudoukh et al., 1994), forecast demand changes (DellaVigna and Pollet, 2007), industry concentration (Bustamante and Donangelo, 2017; Corhay et al., 2020; Hou and Robinson, 2006), and intra-industry value spreads (Bustamante, 2015). Notably, customer-supplier relationships between industries exhibit strong cross-sectional return predictability (Cohen and Frazzini, 2008; Menzly and Ozbas, 2010; Sharifkhani and Simutin, 2021), suggesting that economic linkages between industries create important return dynamics that extend beyond firm-specific characteristics. Recent studies have also incorporated industry returns into asset pricing models to better account for idiosyncratic risks and improve predictive power (Daniel, Mota, Rottke, and Santos, 2020; Hou, Xue, and Zhang, 2015).

At the market level, investor sentiment has been documented as a key determinant of cross-sectional returns. Baker and Wurgler (2006) define investor sentiment as the propensity

to speculate, driven by emotions such as optimism or fear. Investor sentiment is distinct from investor belief, which is based on rational expectations of fundamentals. Baker and Wurgler (2006) find that low sentiment predicts higher subsequent returns for speculative stocks. This finding has been corroborated internationally (Baker, Wurgler, and Yuan, 2012) and extended to show sentiment’s predictive power for market returns (Huang et al., 2015) and its negative impact on short-leg performance in long-short strategies (Stambaugh, Yu, and Yuan, 2012). Da, Engelberg, and Gao (2015) further demonstrate that sentiment predicts short-term return reversals and volatility spikes. While these studies establish sentiment’s importance at the market level, its role in explaining industry returns remains unexplored.

Another strand of the literature focuses on measuring sentiment through textual analysis (Garcia, Hu, and Rohrer, 2023; Jiang, Lee, Martin, and Zhou, 2019; Loughran and McDonald, 2011; Soo, 2018; Tetlock, 2007), developing methodologies to quantify market mood and investor expectations. However, these sentiment measures have primarily been applied at the market or firm level, leaving industry-level sentiment dynamics unexamined. This paper fills the gap in the literature by investigating how industry-level sentiment and its transmission across economically linked sectors contribute to return predictability.

The remainder of the paper proceeds as follows. Section 2 describes the methodology and data. Section 3 analyzes the return predictability of industry-level sentiment. Section 4 examines sentiment transmission between supplier and customer industries. Section 5 presents robustness checks using alternative industry-level sentiment measures. Section 6 concludes. The Appendix shows the graphs comparing the sentiment index of each industry with the market-level sentiment index.

2. Methodology and Data

2.1. Methodology

To construct the sentiment indices at the industry level, I use five sentiment proxies from Baker and Wurgler (2006): number of initial public offerings ($NIPO$), first-day returns of IPOs ($RIPO$), equity share in new issues (S), dividend premium (P^{D-ND}), and closed-end-fund discount ($CEFD$). $NIPO$ is the total number of initial public offerings over the previous four quarters. $RIPO$ is the $NIPO$ -weighted average of quarterly first-day returns of IPOs over the previous four quarters. P^{D-ND} is the natural logarithm of the difference between the value-weighted average market-to-book ratios of dividend payers and non-payers. The indexes use the $t - 4$ value. S is the total volume of equity issues over the prior four quarters divided by the total volume of equity and debt issues over the prior four quarters. The closed-end fund discount ($CEFD$) is the average difference between the net asset values (NAV) of closed-end stock fund shares and their market prices. To eliminate the effect of recessions, all variables are orthogonalized by regressing each of the five raw proxies on a dummy variable for National Bureau of Economic Research (NBER) recession periods.

Originally, Baker and Wurgler (2006) used a sixth sentiment proxy, which was NYSE Turnover. However, in a recent data release, they dropped it, arguing that turnover does not mean what it once did, given the explosion of institutional high-frequency trading and the migration of trading to a variety of venues. For this reason, I do not include this sentiment proxy in my sentiment indices.

Baker and Wurgler (2006) construct the sentiment index using principal component analysis (PCA), where the index represents the first principal component of the correlation matrix of the five sentiment proxies. Huang et al. (2015, p. 792) argue that “because all of the proxies may have approximation errors relative to the true but unobservable investor sentiment, and these errors are part of their variations, the first principal component can potentially contain

a substantial amount of common approximation errors that are not relevant for forecasting returns”. To address this issue, they suggest using the Partial Least Squares (PLS) method, originally developed by Wold (1966, 1975) and later extended by Kelly and Pruitt (2013, 2015), to compute the sentiment index by extracting the most relevant common component from the proxies. Given these concerns, I follow Huang et al. (2015) and adopt the PLS method to construct the industry-level sentiment index, as it explicitly addresses the risk of incorporating irrelevant approximation errors.

I use the industry definitions from the Global Industry Classification Standard (GICS), which includes 11 industries: Energy (10), Basic Materials (15), Industrials (20), Consumer Cyclical (Discretionary) (25), Consumer Defensive (Staples) (30), Healthcare (35), Financial Services (40), Technology (45), Communication Services (50), Utilities (55), and Real Estate (60). According to Morningstar’s sector structure, these industries are grouped into three major super sectors: Cyclical, Defensive, and Sensitive. This classification system enables investors to compare funds and portfolios based on their exposure to the three super sectors, while also allowing for a more detailed analysis at the individual industry level. The Cyclical super sector includes Basic Materials, Consumer Cyclical, Financial Services, and Real Estate, which are significantly impacted by economic shifts. During periods of economic growth, these industries generally experience expansion, whereas they contract in recessions, with their stocks usually exhibiting betas above one. The Defensive super sector includes Consumer Defensive, Healthcare, and Utilities, industries that tend to remain stable across economic cycles due to their provision of necessary goods and services regardless of economic conditions; their stocks are typically characterized by betas below one. The Sensitive super sector includes Communication Services, Energy, Industrials, and Technology, which tend to follow overall economic trends but with less volatility than cyclical industries. Although these industries are affected by economic downturns, their performance is less severely impacted compared to industries in the Cyclical super sector, and their stocks usually have

betas near one.

Let $x_{i,j,t}$ denote the i -th sentiment proxy ($i = 1, \dots, 5$) for industry j ($j = 1, \dots, 11$) at quarter t ($t = 1, \dots, T$). The sentiment index for each industry is constructed using Partial Least Squares (PLS) with two-step OLS regressions.

For each industry j , the first step involves running five time-series regressions, one for each sentiment proxy. Specifically, I regress the lagged sentiment proxy $x_{i,j,t-1}$ on a constant and the realized industry return $R_{j,t}$:

$$x_{i,j,t-1} = \pi_{i,j,0} + \pi_{i,j,1}R_{j,t} + u_{i,j,t-1}, \quad t = 1, \dots, T. \quad (1)$$

In the second step, I run T cross-sectional regressions. For each period t , I run a cross-sectional regression of $x_{i,j,t}$ on the corresponding loading $\hat{\pi}_{i,j}$ estimated from the first step.

$$x_{i,j,t} = c_{j,t} + SENT_{j,t}^{PLS} \hat{\pi}_{i,j} + v_{i,j,t}, \quad i = 1, \dots, 5; \quad j = 1, \dots, 11, \quad (2)$$

where $SENT_{j,t}^{PLS}$, the regression slope in the above equation, is the estimated sentiment score for industry j at time t .

The industry sentiment indices are constructed using the PLS approach with an expanding window. An expanding window allows the sentiment indices to adapt to changing market conditions over time and provides a more dynamic view of sentiment trends. It also ensures that investors can implement the strategy in real time.

2.2. Data

I use data from Thomson Reuters to identify IPOs in the United States from 1993 to 2024. I exclude American Depositary Receipts (ADRs), Real Estate Investment Trusts (REITs), unit and rights offerings, closed-end funds, and IPOs with an offer price below five dollars. By matching IPO data from Thomson Reuters with stock data from Center for Research in

Security Prices (CRSP), I calculate the first-day return as the difference between the IPO offer price and the first-day closing price. The first-day return is winsorized at the 1st and 99th percentile.

The equity share in new issues (S) and dividend premium (P^{D-ND}) are computed using stock data from CRSP and accounting data from Compustat. The sample includes all common stocks (share codes 10 and 11) listed on the New York Stock Exchange (NYSE), NASDAQ, and NYSE American (formerly AMEX) from 1993 to 2024, with stocks priced below five dollars excluded. Data used to compute closed-end fund discount ($CEFD$) is from Compustat Security Monthly (SECM) table. The sentiment proxies are measured quarterly from 1993 to 2024, as the $CEFD$ data is available starting in 1993.

The Herfindahl-Hirschman Index (HHI) is calculated using annual accounting data from Compustat. Institution ownership (IO) ratio is computed using data from the FactSet Ownership - Historical 13F Holding Data. Analyst coverage is measured using data from Institutional Brokers' Estimate System (I/B/E/S).

I use Compustat Segments - Customer to identify customer industries for a given industry. Market-level sentiment indices are from Baker and Wurgler (2024) and Huang et al. (2015). News sentiment scores at the company level are from RavenPack, with data on entity sentiment available for the period from 2000 to 2024. Industries are defined according to the Global Industry Classification Standard (GICS).

Figure 1 shows the average number of stocks per industry from 1993 to 2024. Financial Services has the highest average number of stocks, peaking at 729, which is substantially higher than in any other industry. Technology, Consumer Cyclical (Discretionary), and Industrials each have more than 500 stocks per year on average. Real Estate has the lowest average number of stocks per year, with only 27. The average number of stocks per industry highlights the disparity in stock representation across different sectors, suggesting a concentration of stocks in certain industries over the research period.

[Figure 1]

Table 1 presents the descriptive statistics for the proxies used to construct the industry-level sentiment indices.

[Table 1]

The average number of initial public offerings over the previous four quarters (*NIPO*) ranges from 0.28 to 13.84, with Technology (45) exhibiting the highest average *NIPO* at 13.84. For first-day returns of IPOs (*RIPO*), Technology (45) also has the highest average first-day returns (26.63%) and highest standard deviation (37.48%). Consumer Defensive (30) has the highest standard deviation regarding dividend premium ($P^D - ND$). Additionally, Real Estate (60) displays the highest mean closed-end fund discount (*CEFD*).

The descriptive statistics underscore significant variation in the sentiment proxies across industries, showing distinct industry-specific patterns in market activity and investor behavior. Specifically, the pronounced differences in the number of IPOs and first-day returns across industries suggest that investor sentiment manifests differently depending on industry characteristics.

Table 2 presents the correlation between the industry-level sentiment and the market-level sentiment indices, where PLS_MKT is the market-level sentiment index from Huang et al. (2015), and PCA_MKT is the market-level sentiment index from Baker and Wurgler (2024). The market-level sentiment data are available up to December 2023.

[Table 2]

Overall, the sentiment of cyclical industries, such as Basic Materials (15), Consumer Cyclical (25), Financial Services (40), and Real Estate (60), is positively correlated with the market sentiment indices. In particular, the Consumer Cyclical and Financial Services industries exhibit the highest positive correlations. The sentiment of the Consumer Cyclical industry shows a correlation of 0.38 with the market sentiment index from Huang et al. (2015)

and 0.63 with the market sentiment index from Baker and Wurgler (2024). Similarly, the sentiment of the Financial Services industry is correlated at 0.52 with Huang et al. (2015)'s market sentiment index and at 0.60 with Baker and Wurgler (2024)'s market sentiment index.

Among the defensive industries, the Consumer Defensive (30) industry shows a slightly negative correlation with the market sentiment indices from Huang et al. (2015) and Baker and Wurgler (2024), with values of -0.06 and -0.13, respectively. Meanwhile, the sentiment indices of the Healthcare (35) and Utilities (55) industries are slightly positively correlated with the market indices. Within the sensitive sector, the Industrials (20) and Technology (45) industries have the highest correlation with the market sentiment index, at 0.43 and 0.50 to 0.52, respectively.

The correlations indicate that industry-level sentiment does not always align with overall market sentiment. This underscores the importance of considering industry-specific sentiment for more effective assessment of risks and opportunities in portfolio.

3. Industry-Level Sentiment and Return Predictability

After constructing the sentiment index for each industry using the PLS approach, I examine whether these industry-level sentiment indices can effectively predict future market-level sentiment and industry returns.

3.1. Does Industry-Level Sentiment Predict Market-Level Sentiment?

Hong et al. (2007) find that the returns of industry portfolios can predict stock market movements. They argue that stock markets react with a delay to information contained in industry returns about their fundamentals and that this information diffuses only gradually across markets. This raises the question: if industry returns can predict stock market returns, can industry-level sentiment predict market-level sentiment?

To investigate whether industry-level sentiment can predict market-level sentiment, I first

apply PCA to the eleven industry-level sentiment scores and extract the first three principal components. This approach allows for a more manageable and interpretable analysis of the complex relationships between industry-specific sentiments and market sentiment. These components are linear combinations of the original industries' sentiments and are ordered by the amount of variance they explain in the data.

To assess the predictive power of these principal components, I regress the market-level sentiment index, as defined by Huang et al. (2015), on the three principal components derived from the industry-level sentiment scores. Table 3 reports the results of the regression of the market-level sentiment index from Huang et al. (2015) on these three principal components.

[Table 3]

The regression results show that only the second principal component (PC2) of industry-level sentiment strongly predicts market sentiment, while the first (PC1) and third (PC3) do not. PC1 likely represents a broad, average sentiment across all industries, which may already be priced into the market and thus lacks predictive power. PC2, on the other hand, captures a specific pattern of sentiment shared by industries that are particularly influential, such as those sensitive to economic cycles or investor risk appetite, making it a better indicator of overall market mood. PC3, which explains the least variance, appears to capture minor or idiosyncratic sentiment fluctuations with little impact. This result implies that market-level sentiment is not shaped equally by all industries.

To understand more about how industry-level sentiment can predict market-level sentiment, I perform a regression where the market-level sentiment from Huang et al. (2015) is regressed on the sentiment index of each industry.

[Table 4]

When market-level sentiment is regressed on industry-level sentiment scores, only the sentiment of three out of eleven industries significantly predicts market-level sentiment: Con-

sumer Cyclical (25), Technology (45), and Communication Services (50). The positive coefficients for these industries indicate a positive relationship between their sentiment and overall market sentiment. These are also the industries whose sentiment indices are highly correlated with market-level sentiment indices.

The [Appendix](#) presents graphs comparing the sentiment index of each industry with the market-level sentiment index from Huang et al. (2015) (available up to December 2023). The graphs indicate that the sentiment of industries such as Industrials (20), Consumer Cyclical (25), Financial Services (40), and Technology (45) closely tracks the market sentiment index.

3.2. Does Industry-Level Sentiment Explain Industry Returns?

3.2.1. Fama and MacBeth (1973) Regression

The previous analysis demonstrates that industry-level sentiment, particularly in specific industries, can predict market-level sentiment. Building on this, I now examine whether industry-level sentiment can also explain industry returns. I first run a Fama and MacBeth (1973) regression of the previous quarter’s sentiment scores ($\text{SENT}_{j,t-1}$) on the current quarter’s industry return ($R_{j,t}$):

$$R_{j,t} = \alpha + \beta_t^{\text{SENT}} \text{SENT}_{j,t-1} + e_{j,t}. \quad (3)$$

The results of this regression are presented in the first two columns of Table 5.

[Table 5]

Table 5 shows that lagged industry-level sentiment explains industry returns. The coefficients for $\text{SENT}_{j,t-1}$, which are 0.69 for equal-weighted and 0.64 for value-weighted portfolios, are statistically significant with t-statistics of 2.33 and 2.99, respectively. This indicates a positive relationship between sentiment scores and subsequent industry returns.

Moskowitz and Grinblatt (1999) demonstrate that industry momentum can explain industry returns. Thus, I aim to examine the extent to which industry momentum can predict industry returns. Columns (3) and (4) present the results of the following specification:

$$R_{j,t} = \alpha + \beta_t^R R_{j,t-1} + e_{j,t}, \quad (4)$$

where $R_{j,t-1}$ is the return of industry j in quarter $t - 1$ (industry momentum).

The results show that industry momentum is statistically significant only in the equal-weighted portfolio, with a coefficient of 0.16 and a t-statistic of 2.81. In the value-weighted portfolio, the coefficient is 0.04 with a t-statistic of 1.02. It is noteworthy that the sentiment results are consistent across both equal-weighted and value-weighted portfolios, while industry momentum remains significant only in the equal-weighted portfolio. This finding aligns with prior literature, such as Grinblatt, Titman, and Wermers (1995) and Grundy and Martin (2001), which demonstrate that industry momentum tends to be strongest at shorter horizons, particularly at the one-month level. This may explain why the momentum effect weakens in value-weighted portfolios, where larger firms, often less sensitive to short-term momentum, dominate the returns. Notably, the sentiment results remain robust and consistent across both equal-weighted and value-weighted portfolios, highlighting the broader predictive power of sentiment relative to momentum.

To investigate whether industry-level sentiment provides additional predictive power beyond industry momentum, I include industry momentum as a control variable in the Fama and MacBeth (1973) regression:

$$R_{j,t} = \alpha + \beta_t^{\text{SENT}} \text{SENT}_{j,t-1} + \beta_t^R R_{j,t-1} + e_{j,t}, \quad (5)$$

where $\text{SENT}_{j,t-1}$ is the sentiment score of industry j in quarter $t - 1$ and $R_{j,t-1}$ is the return of industry j in quarter $t - 1$.

The results of this regression are presented in the last two columns of Table 5. The findings demonstrate that lagged industry-level sentiment explains industry returns for equal-weighted portfolios, even after controlling for industry momentum. However, when it comes to value-weighted results, only the lagged industry-level sentiment can explain industry returns. The coefficients for $\text{SENT}_{j,t-1}$ of 0.70 (t-statistic = 2.91) and 0.63 (t-statistic = 0.63) for equal-weighted and value-weighted results, are statistically significant. These coefficients indicate a positive relationship between the sentiment score and industry returns. This suggests that, on average, industries with higher sentiment scores in the previous quarter tend to experience higher returns in the current quarter, and vice versa. A one-standard-deviation increase in the sentiment score is associated with a 0.70% and 0.63% increase in quarterly equal- and value-weighted industry returns, respectively.

3.2.2. Pooled and Time-Series Regression Analyses

The pooled regression results presented in Table 6 examine the relationship between industry returns and lagged industry-level sentiment, while controlling for industry momentum, lagged market-level sentiment, and fixed effects.

[Table 6]

Across all specifications, the coefficient on lagged industry-level sentiment (SENT_{t-1}) remains consistently positive and statistically significant at the 1% level, with estimates ranging from 1.00% to 2.30%. This robustness suggests a strong and persistent predictive power of industry-level sentiment for subsequent industry returns. The inclusion of industry and time fixed effects generally reduces the magnitude of the sentiment coefficient slightly, indicating that a portion of the sentiment effect may be attributed to unobserved industry-specific or time-varying factors. However, the consistent significance of the sentiment coefficient after controlling for these fixed effects underscores the predictive power of industry-level sentiment.

The specifications that include industry momentum ($R_{j,t-1}$) as a control variable yields

mixed results, with specification (10) showing a marginally positive relationship, while others indicate no significant effect. This suggests that industry momentum, while relevant, does not fully account for the relationship of industry-level sentiment and industry returns.

The inclusion of lagged market sentiment (SENT_Market_{t-1}) indicates a negative but generally insignificant relationship with industry returns. This result implies that while industry-specific sentiment is a strong predictor of industry returns, broader market sentiment does not exhibit the same predictive power in this context.

Overall, these results provide evidence that industry-level sentiment is a robust and economically meaningful predictor of industry returns, even after controlling for industry momentum, market sentiment, and fixed effects. The consistency of the sentiment coefficient across specifications reinforces the idea that industry-level sentiment contains unique information not fully captured by traditional factors such as momentum or market-wide sentiment.

To examine whether industry-level sentiment can explain industry returns in the time series, I run regressions of industry returns on industry-level sentiment for each industry. The results are presented in Table 7.

[Table 7]

The first three columns in Table 7 present the results when the return of each industry is regressed on its one-quarter lagged industry-level sentiment index ($\text{SENT}_{j,t-1}$) and its one-quarter lagged return ($R_{j,t-1}$). The findings show that industry-level sentiment significantly predicts returns in eight out of eleven industries. Notably, all the industries for which industry-level sentiment fails to explain returns, namely, Basic Materials (15), Financial Services (40), and Real Estate (60), are classified as cyclical industries according to Morningstar’s definition.

The last three columns in Table 7 show the results when the return of each industry is regressed on its one-quarter lagged industry-level sentiment index ($\text{SENT}_{j,t-1}$) and the one-quarter lagged market sentiment ($\text{SENT}_{MKT,t-1}$). When market-level sentiment is included

as a control, industry-level sentiment continues to significantly predict returns for seven out of eleven industries. These seven industries are the same as those identified in the first specification, with the exception of Utilities (55), which loses significance when market-level sentiment is accounted for.

The pooled and time-series regression results collectively confirm that industry-level sentiment is a reliable predictor of industry returns. The predictive power of sentiment persists even after accounting for unobserved industry-specific and time-varying factors, as well as broader market-level sentiment.

3.2.3. Portfolios Sorted by Industry-Level Sentiment Indices

The pooled and time-series regression analyses confirm that industry-level sentiment is a robust predictor of industry returns. To further evaluate the economic significance of this relationship, I now examine a portfolio-based approach. By sorting industries into quintiles based on their lagged sentiment scores, I test whether sentiment-driven strategies can generate significant returns.

Each quarter, I sort the eleven industries into quintile portfolios based on their industry-level sentiment scores in the previous quarter. Each quintile contains two industries, except for the third quintile, which contains three. The high-minus-low industry portfolio involves buying industries with high sentiment and selling those with low sentiment. The first column of Table 8 reports the value-weighted returns from sorting industries into portfolios based on sentiment scores at quarters $t - 1$.

[Table 8]

Consistent with the Fama and MacBeth (1973) regression results, the high-minus-low portfolios generate significant excess returns, with quarterly value-weighted returns of 2.12% (t-statistic = 3.36) when sorted by sentiment scores at $t - 1$, which is equivalent to an annualized return of 8.48%. To investigate whether this return persistence extends over

time, I sort portfolios based on industry-level sentiment scores from quarters $t - 2$, $t - 3$, or $t - 4$.

Columns 2 to 4 of Table 8 report the value-weighted returns from sorting industries into portfolios based on sentiment scores at quarters $t - 2$, $t - 3$, and $t - 4$ (SENT_{t-1} , SENT_{t-2} , SENT_{t-3} , and SENT_{t-4}), respectively. The strategy remains effective when sorting based on sentiment scores from quarter $t - 2$, generating significant quarterly value weighted-returns of 1.93% (t-statistic = 2.93). However, when using sentiment scores from $t - 3$, $t - 4$, or further lags, the strategy is no longer significant.

At the market level, sentiment and future returns are negatively correlated because high sentiment often leads to overvaluation, which later corrects as prices revert to fundamentals. Conversely, low sentiment may signal undervaluation due to excessive pessimism, leading to future price rebounds. However, at the industry level, I find a positive short-term correlation between sentiment and returns. This aligns with behavioral mispricing: speculative demand initially inflates prices beyond fundamentals, but the effect fades as mispricing is arbitrated away. Delays in information processing also contribute to mispricing. The gradual-information-diffusion hypothesis (Hong et al., 2007) posits that information spreads slowly between markets due to two key factors: information travels with a lag, and investors, constrained by limited attention capacity, focus primarily on their specialized markets. High-sentiment industries often capture early signals of growth opportunities, which enhance industry-wide sentiment and lead to higher future returns. As this information is gradually absorbed by related sectors and the broader market, the relationship reverses, mirroring the pattern observed at the stock level. The positive correlation at the industry level thus reflects a temporary lead-lag effect, where sentiment anticipates fundamental improvements before they are fully priced in. The positive sentiment-return relationship for industries is also consistent with the result of Stambaugh et al. (2012), who find that when market sentiment is high, subsequent cross-sectional anomaly returns are positive.

3.2.4. Bivariate Portfolio Analyses

Fama and French (1997) demonstrate that industry momentum can explain industry returns. Hou and Robinson (2006) and Bustamante and Donangelo (2017) find that industry concentration is a significant predictor of industry returns. DeVault, Sias, and Starks (2019) show that sentiment metrics capture the demand shocks of institutional, rather than individual, investors. Lee and So (2017) show that analyst coverage proxies contain information about expected returns. While univariate portfolio analysis provides an initial understanding of how industry-level sentiment influences industry returns, it does not account for potential interactions or overlapping effects among these variables. To control for industry momentum, industry concentration, institutional ownership, and analyst coverage, I conduct conditional bivariate sort analyses.

In the first sort, the eleven industries are divided into two portfolios (High and Low) based on the value of the control variable. Since the number of industries is odd, the sixth industry in the ranking is assigned to both portfolios. In the second sort, each of these portfolios is further sorted into three sub-portfolios based on industry-level sentiment scores. The portfolios SENT_Low and SENT_High comprise the two industries with the lowest and highest industry-level sentiment scores, respectively. Table 9 presents the results of bivariate portfolio analyses.

[Table 9]

Industry momentum is defined as the industry’s return in the previous quarter. The results in Table 9 show that in high-momentum industries, the strategy of buying high-sentiment industries and selling low-sentiment industries generates an excess return of 1.55% with a t-statistic of 2.97. High-momentum industries often receive positive sentiment as investors assume strong past performance will continue, pushing prices higher. Unlike individual stocks where overvaluation from excessive optimism may eventually reverse, industry-level

momentum often persists longer. This persistence is driven by slower information diffusion among sectors and the tendency of institutional investors to herd.

I measure industry concentration using the Herfindahl-Hirschman Index (HHI). The HHI is defined as:

$$\text{HHI}_j = \sum_{i=1}^I s_{ij}^2, \quad (6)$$

where s_{ij} is the market share of firm i in industry j . Following Hou and Robinson (2006), the HHI is calculated annually for each industry and averaged over the past three years. Market share is measured using net sales, but alternative measures can include total assets or book equity. I calculate the HHI using these measures and find that they correlate highly with the HHI calculated based on net sales. Therefore, I use net sales as the measure of market share in my analysis.

When controlling for industry concentration, the results indicate that a trading strategy involving long positions in high-sentiment industries and short positions in low-sentiment industries produces significant excess returns in both high- and low-concentration industries. The strategy generates returns of 1.49% (t-statistic = 2.90) in the high-concentration portfolio and 2.05% (t-statistic = 3.43) in the low-concentration portfolio. In high-concentration industries, sentiment induces strong and synchronized price movements, as dominant firms amplify investor reactions, leading to momentum effects that persist until fundamentals are fully reflected in prices. In low-concentration industries, the effect is more substantial, likely due to limited analyst coverage and slower information dissemination, which prolongs the correction of sentiment-driven mispricing.

As in prior research, I define institutional ownership as the fraction of a firm's shares that are held by institutional investors in each quarter. When controlling for institutional ownership, the trading strategy of buying high-sentiment industries and selling low-sentiment industries remains effective only in portfolios with low institutional ownership, generating a

value-weighted return of 1.32% with the statistical significance at the 10% level. This finding aligns with established finance literature emphasizing the role of institutional investors in enhancing market efficiency and reducing sentiment-driven mispricing. According to Boone and White (2015), higher institutional ownership is associated with greater management disclosure, higher analyst coverage, and improved liquidity, all of which contribute to lower information asymmetry. In low-institutional-ownership industries, the lack of sophisticated capital allows sentiment effects to persist, creating opportunities for the trading strategy to generate alpha. Furthermore, Nagel (2005) demonstrates that return predictability is strongest for stocks with low institutional ownership, as short-sale constraints and limited arbitrage allow mispricing to persist.

Analyst coverage is measured by the average number of analysts covering an industry in each quarter. The last three columns in Table 9 show that the high-minus-low industry portfolio has significant returns in industries with lower analyst coverage with the quarterly return of 2.03% and t-statistic of 2.13. This occurs because, in these industries, information asymmetry is more pronounced, which slows the correction of sentiment-driven mispricing.

The bivariate portfolio analyses confirm that industry-level sentiment predicts returns even after controlling for industry momentum, industry concentration, institutional ownership, and analyst coverage. The sentiment effects are more pronounced in industries with limited market oversight, such as those with low institutional ownership or analyst coverage, where information asymmetry allows mispricing to persist. These findings indicate that investors underreact to sentiment-driven information, particularly in environments with weaker informational efficiency. In industries with low institutional ownership or limited analyst coverage, the reduced presence of sophisticated market participants delays the incorporation of sentiment signals into prices. This underreaction suggests that sentiment contains predictive information that is only gradually reflected in returns, consistent with models of gradual information diffusion proposed by Hong et al. (2007).

4. Industry-Level Sentiment and Supplier-Customer Relationships

Bustamante (2015) suggests that firms' expected returns depend on their strategic interactions with peers, creating cross-industry dynamics that differ from broader market trends. In addition, according to Menzly and Ozbas (2010), the gradual diffusion of information leads to strong cross-predictability effects between stock- and industry-level returns based on lagged returns in supplier and customer industries. Sharifkhani and Simutin (2021) further shows that industries are economically linked through customer-supplier trade flows, and that shocks propagating through this network can feed back into the originating industry, producing an "echo" effect that results in intermediate-term autocorrelation in returns.

To investigate whether industry-level sentiment is linked to these customer-supplier relationships, I first assess whether customer industry returns predict supplier industry returns, thereby validating existing findings in the literature. I then examine whether customer industry-level sentiment can predict supplier industry-level sentiment, extending the analysis to the role of sentiment in these economic linkages.

For this analysis, I use the Compustat Segments - Customer file to identify supplier-customer relationships, rather than the Input-Output data from the Bureau of Economic Analysis (BEA) used by Menzly and Ozbas (2010). While the BEA Input-Output data capture the flow of goods and services across the entire economy, my study focuses exclusively on firms listed on the NYSE, NASDAQ, and NYSE American (formerly AMEX). The Compustat Segments - Customer file is therefore more suitable for my sample.

Following Cohen and Frazzini (2008), I extract the identities of firms' principal customers from the Compustat Segments - Customer file. For each firm, I consider only customers listed on the NYSE, NASDAQ, and NYSE American (formerly AMEX) exchanges. When customer names do not exactly match those in the Compustat database, I employ a fuzzy

string matching algorithm to generate a list of potential matches. I then manually verify and link each customer to the correct GVKEY by cross-referencing firm names and industry sectors. Customers that cannot be uniquely matched are excluded from the sample. I then aggregate the supplier-customer data at the industry level.

I follow the methodology of Menzly and Ozbas (2010) to construct the customer portfolio. For example, if industry 1 supplies customer industries 2 through 11, I first calculate the weight of each customer industry (2 to 11) as a proportion of industry 1's total sales. Using these weights, I multiply them by the return of each corresponding customer industry and sum the results to obtain the customer portfolio return. I also multiply the same weights by the sentiment of each corresponding customer industry and sum the results to obtain the customer portfolio sentiment.

I then estimate two Fama and MacBeth (1973) regressions to examine the relationship between supplier and customer industries. In the first regression, the dependent variable is the quarterly supplier industry return, and the independent variable is the lagged customer industry return. In the second regression, the dependent variable is the quarterly supplier industry sentiment, and the independent variable is the lagged customer industry sentiment.

$$R_t^{\text{Supplier}} = \alpha_t + \beta_t^{\text{Customer}} R_{t-1}^{\text{Customer}} + e_t \quad (7)$$

$$\text{SENT}_t^{\text{Supplier}} = \alpha_t + \beta_t^{\text{Customer}} \text{SENT}_{t-1}^{\text{Customer}} + e_t \quad (8)$$

Panel A of Table 10 presents the regression results of Equation (7), showing that lagged customer industry returns ($R_{t-1}^{\text{Customer}}$) positively predict focal industry returns. The coefficient of 1.38, with a t -statistic of 6.53 in the Fama and MacBeth (1973) regression, indicates a strong predictive relationship. This result aligns with the findings of Menzly and Ozbas (2010). The industry-level regression results further demonstrate that customer industry returns predict supplier industry returns in nine out of eleven industries. All coefficients are positive and statistically significant, suggesting that positive returns in customer industries

are associated with subsequent positive returns in supplier industries.

[Table 10]

Panel B of Table 10 presents the regression results of Equation (8), where supplier industry sentiment is regressed on lagged customer industry sentiment ($\text{SENT}_{t-1}^{\text{Customer}}$). The Fama and MacBeth (1973) regression results indicate that lagged customer industry sentiment positively predicts focal industry sentiment, with a highly statistically significant coefficient of 6.75. Industry-specific regressions further demonstrate that customer industry sentiment predicts supplier industry sentiment in all eleven industries. All coefficients are positive and statistically significant, suggesting that positive sentiment in customer industries is associated with subsequent positive sentiment in supplier industries.

These results indicate that customer industries significantly influence their suppliers. The positive lagged relationship occurs because supplier industries rely on their customers for demand. When customer industries show positive sentiment in one quarter, suppliers become more optimistic in the next quarter as they expect higher orders. Investors adjust their expectations for suppliers after seeing strong performance in customer industries. Together, these findings illustrate how closely linked industries shape each other’s sentiment through both economic fundamentals and market expectations.

Given that customer industry sentiment in quarter $t - 1$ explains supplier industry sentiment in quarter t , and that supplier industry sentiment in quarter t predicts supplier industry returns in quarter $t + 1$, a natural question arises: Can customer industry sentiment directly explain supplier industry returns?

Panel C of Table 10 presents the regression results of the following equation:

$$R_t^{\text{Supplier}} = \alpha_t + \beta_t^{\text{Customer}} \text{SENT}_{t-2}^{\text{Customer}} + e_t, \quad (9)$$

where customer industry sentiment is lagged by two quarters. This specification assumes

that customer sentiment in quarter $t - 2$ first predicts supplier sentiment in quarter $t - 1$, which in turn predicts supplier returns in quarter t .

The results in Panel C of Table 10 include both cross-sectional and time-series analyses. The cross-sectional Fama and MacBeth (1973) regression indicates that customer industry sentiment ($\text{SENT}_{t-2}^{\text{Customer}}$) positively predicts supplier industry returns (R_t^{Supplier}), with a coefficient of 14.35 and a t -statistic of 2.94. In the time series, industry-specific regressions show that customer industry sentiment predicts supplier returns in five industries: Consumer Cyclical (25), Financial Services (40), Technology (45), Communication Services (50), and Real Estate (60). One plausible explanation for this result is that the strength of customer-supplier relationships varies across industries. In industries with stronger or more direct customer-supplier linkages, customer sentiment has a more pronounced impact on supplier returns. Conversely, industries with weaker or less immediate dependencies may not exhibit the same predictive relationship, underscoring how the intensity of economic linkages influences sentiment transmission.

5. Robustness Check

Constructing industry-level sentiment indices using the sentiment proxies from Baker and Wurgler (2006) is a top-down approach. To validate my industry-level sentiment measures and assess the robustness of my findings, I construct alternative sentiment measures in this section using a bottom-up approach based on RavenPack sentiment data and a sentiment index derived from textual analysis of Morningstar industry reports. I further investigate whether changes in industry-level sentiment, in addition to sentiment levels, exhibit predictive power for future returns.

5.1. RavenPack Sentiment Data

RavenPack is a leading commercial provider of financial news sentiment scores, widely used in both academic and industry research. However, its proprietary methodology makes it a “black box”. Nevertheless, I use the RavenPack data, which is available since 2000, to construct a bottom-up industry-level sentiment measure to validate my top-down approach. By comparing results across these two distinct methodologies, I assess whether the findings remain consistent or if discrepancies arise due to differences in sentiment construction.

The RavenPack data sample includes company-level news content sourced from Dow Jones Newswires and the Wall Street Journal, focusing exclusively on business and economy news in the United States. I use the Composite Sentiment Score, which combines five sentiment analytics: Stock Tone Sentiment, Earnings Tone Sentiment, Commentary Sentiment, Mergers & Acquisitions Sentiment, and Corporate Actions Sentiment. I merged RavenPack data with my current sample, which includes all common stocks (share codes 10 and 11) listed on the New York Stock Exchange, NYSE American (formerly AMEX), and NASDAQ.

RavenPack’s sentiment scores range from -1.00 to +1.00. These scores represent the sentiment of individual news stories about a company, derived from an analysis of the language used in each sentence mentioning the company. RavenPack also provides a Sentiment Confidence Score, which ranges from 0.00 to +1.00. This confidence score reflects the reliability of the sentiment measurement by indicating the level of agreement in the sentiment distribution. Specifically, it measures the degree of consistency in the sentiment score’s magnitude and direction. A lower confidence score (below 0.20) signifies a high degree of uncertainty, whereas a higher confidence score (above 0.80) indicates high reliability. Therefore, I exclude observations with a sentiment confidence score below 0.80 from the analysis. I then aggregate the news sentiment scores quarterly at the industry level using the Global Industry Classification Standard (GICS).

Table 11 and 12 present analyses using news sentiment scores derived from RavenPack

data, aggregated at the industry level. To make these analyses comparable with the results obtained using the industry-level sentiment index I have constructed, I also run the same analyses for the period from 2000 to 2024 using the sentiment index constructed by the PLS method with an expanding window.

[Table 11]

Table 11 presents the univariate results of portfolios sorted by the RavenPack and PLS sentiment indices for the period from 2000 to 2024. Each quarter, eleven industries are sorted into quintile portfolios based on their industry-level sentiment scores at quarter $t - 1$, $t - 2$, $t - 3$, or $t - 4$. The trading strategy involves buying industries with high sentiment and selling those with low sentiment. The result table presents the value-weighted returns from sorting industries into portfolios based on sentiment scores at quarters $t - 1$, $t - 2$, $t - 3$, and $t - 4$ (SENT_{t-1} , SENT_{t-2} , SENT_{t-3} , and SENT_{t-4}), respectively. Panel A presents the value-weighted results for portfolios sorted by the industry-level sentiment index constructed using RavenPack data, which is available from 2000. Panel B displays the results for portfolios sorted by the industry-level sentiment index constructed using the PLS method with an expanding window. Similar to the results obtained using the sentiment index constructed via PLS with an expanding window, the strategy using industry-level sentiment aggregated from RavenPack data yields quarterly value-weighted returns of 2.45% (t-statistic = 2.97) and 3.29% (t-statistic = 3.17) when sorting portfolios based on industry-level sentiment scores from quarters $t - 1$ and $t - 2$, respectively. The strategy earns a weakly significant result when sorting by sentiment indices at $t - 3$, with value-weighted quarterly returns of 1.15% (t-statistic = 1.73). However, the effect disappears when portfolios are sorted based on industry-level sentiment scores from quarter $t - 4$. This indicates that investors can generate excess returns in the short term by taking long positions in high-sentiment industries and short positions in low-sentiment industries. Additionally, the findings in Panel B confirm that the results in Section 3 remain robust when using only the most recent sample.

[Table 12]

Table 12 shows the value-weighted results of the bivariate portfolio analyses from 2000 to 2024. In the first sort, the eleven industries are divided into two portfolios (High and Low) based on momentum, industry concentration, institutional ownership, or analyst coverage. Industry momentum is defined as the industry’s return in the previous quarter. Industry concentration is measured using the Herfindahl-Hirschman Index (HHI). Institutional ownership is the fraction of a firm’s shares held by institutional investors in each quarter. Analyst coverage is measured by the average number of analysts covering an industry in each quarter. Since the number of industries is odd, the sixth industry in the ranking is assigned to both portfolios. In the second sort, each of these portfolios is further sorted into three sub-portfolios based on industry-level sentiment scores.

Overall, the results derived from using the RavenPack sentiment score are consistent with those obtained using the sentiment index constructed via the PLS method with an expanding window. Specifically, when industry-level sentiment is measured using the RavenPack score, a trading strategy that involves buying high-sentiment industries and selling low-sentiment industries produces notable returns. This strategy generates a quarterly value-weighted return of 1.49% in high-momentum industries and a value-weighted return of 1.68% in low-momentum industries, with t-statistics of 2.41 and 2.32, respectively.

Furthermore, the trading strategy continues to yield positive quarterly returns even after controlling for industry concentration. In high-concentration industries, the strategy yields a value-weighted return of 1.94% with a t-statistic of 2.62, while in low-concentration industries, it generates a return of 2.29% with a t-statistic of 2.01. However, when controlling for institutional ownership, the strategy of buying high-sentiment industries and selling low-sentiment industries generates positive quarterly returns of 2.50% (t-statistic = 2.27) only for industries with low institutional ownership. Similarly, when controlling for analyst coverage, this strategy yields positive quarterly returns of 2.95% (t-statistic = 2.89) only for industries

with low analyst coverage.

5.2. Sentiment Analysis of Industry Reports Using Textual Analysis

To construct the industry-level sentiment index using textual analysis, I employ a dataset of 1,662 sector reports from Morningstar Direct, spanning the years 2010 to 2024. Since some reports address multiple industries, each report is included in the analysis for all industries it covers. As a result, the number of reports analyzed varies across industries: Basic Materials (194), Consumer Cyclical (597), Financial Services (450), Real Estate (203), Consumer Defensive (317), Healthcare (365), Utilities (199), Communication Services (417), Energy (285), Industrials (505), and Technology (502).

To measure the sentiments, I use both unigrams and bigrams from either the sentiment dictionary of Loughran and McDonald (2011) or the machine learning dictionary of finance words from Garcia et al. (2023). The sentiment of each document is calculated as follows:

$$\text{Sentiment Score} = \frac{\text{No. of positive words} - \text{No. of negative words}}{\text{No. of positive words} + \text{No. of negative words}}.$$

The sentiment score ranges from -1 to +1, where -1 reflects fully negative sentiment, +1 reflects fully positive sentiment, and 0 indicates neutrality. Because industry reports are published at different frequencies (weekly, monthly, or quarterly), I aggregate the sentiment scores on a quarterly basis for each industry.

I then use these sentiment scores to analyze portfolio performance. Table 13 shows the univariate results of portfolios sorted by the industry-level sentiment index derived from the textual analysis of industry reports.

[Table 13]

The high-minus-low portfolio constructed using the industry-level sentiment index from textual analysis of industry reports with the sentiment dictionary of Loughran and McDonald

(2011) yields a quarterly equal-weighted return of 1.55%, with a t-statistic of 2.25 and a value-weighted return of 1.63%, with a t-statistic of 2.12. However, constructing portfolios using the machine learning dictionary of finance words from Garcia et al. (2023) does not lead to excess returns. While the results generally support the main findings, the use of industry reports comes with important limitations. The reports are only available from 2010 to 2024, and even within this period, coverage is inconsistent, with many industries lacking frequent updates. Between 2012 and 2016, for example, several industries had no sector reports at all. Moreover, since some reports address multiple industries simultaneously, it is difficult to extract precise, industry-specific sentiment. These limitations likely contribute to the weaker predictive power of sentiment measures derived from textual analysis of sector reports. For these reasons, the main analysis relies on the PLS measures, which benefit from a longer sample period, more comprehensive data, and a transparent construction process.

5.3. The Change in Industry-level Sentiment

Given that the industry-level sentiment index can explain industry returns, it is reasonable to investigate whether changes in industry-level sentiment also has return predictability. As noted by Baker and Wurgler (2007), the levels index is derived as the first principal component of the sentiment proxies. Rather than directly using changes in the sentiment levels index, a more robust method involves constructing the change index based on the first principal component of the changes in the sentiment proxies. This approach is preferred because the proxies exhibit varying degrees of noisiness when going from levels to changes. Employing this method, DeVault et al. (2019) demonstrate that changes in sentiment primarily capture aggregate demand shocks driven by institutional investors, rather than individual investors.

To explore the return predictability of the change in industry-level sentiment, I construct an industry-level change sentiment index using both the PCA approach and the PLS approach applied to changes in sentiment proxies.

[Table 14]

However, the results in Table 14 suggest that changes in industry-level sentiment do not exhibit significant predictive power for industry returns. The finding that industry-level sentiment, but not changes in industry-level sentiment, can predict returns suggests that industry returns are more closely tied to sustained sentiment trends rather than short-term sentiment shifts.

6. Conclusion

This paper demonstrates that industry-level investor sentiment is a significant and robust predictor of industry returns, contributing to the literature on return predictability and behavioral finance. Using a partial least squares (PLS) approach with an expanding window, I construct industry-level sentiment indices and find that industries with high investor sentiment outperform those with low sentiment by 8.48% annually. This predictive relationship persists even after controlling for industry momentum, market sentiment, and fixed effects, and it is most pronounced in industries with low institutional ownership or analyst coverage. The outperformance remains economically meaningful for up to two quarters and is corroborated by alternative sentiment measures, such as RavenPack news sentiment and MorningStar industry reports textual analysis, reinforcing the robustness of the findings.

Beyond the direct effect of industry-level sentiment on industry returns, the analysis also shows sentiment spillovers between supplier and customer industries. Customer industry sentiment positively predicts supplier industry sentiment, and in some cases, supplier returns, highlighting the transmission of sentiment through economic linkages. While the predictive relationship between customer sentiment and supplier returns is not uniform across all industries, the documented spillover effects underscore the importance of considering inter-industry dependencies in return forecasting and investment strategies. These findings align with the gradual information diffusion hypothesis (Hong et al., 2007) and extend prior work

on economic linkages (Bustamante, [2015](#); Cohen and Frazzini, [2008](#); Menzly and Ozbas, [2010](#)) by demonstrating how sentiment propagates through supplier-customer networks.

The results have important implications for both investors and corporate managers. For investors, industry-level sentiment provides a valuable tool for identifying mispricing opportunities and enhancing portfolio construction, particularly in industries with low institutional ownership or analyst coverage. For corporate managers, understanding sentiment spillovers can help in anticipating demand shocks and optimizing supply chain and investment decisions. The study also contributes to the broader literature by showing that industry-level sentiment contains unique information not fully captured by traditional factors like momentum or market-wide sentiment.

To further expand the analysis of industry-level sentiment, I plan to incorporate additional sentiment measures, such as news sentiment from major outlets and aggregated earnings call sentiment at the industry level. I will also investigate how information disseminates across industries, specifically examining whether production networks influence the transmission of information between sectors. Future research could further examine the role of industry sentiment in different economic environments or across international markets.

References

- Baker, M., Wurgler, J., 2006. Investor sentiment and the cross-section of stock returns. *The Journal of Finance*. 61, 1645–1680.
- Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. *Journal of Economic Perspectives*. 21, 129–151.
- Baker, M., Wurgler, J., 2024. Investor sentiment data, 196507-202312 [WWW Document]. Retrieved July 23, 2024, from <https://pages.stern.nyu.edu/~jwurgler/>
- Baker, M., Wurgler, J., Yuan, Y., 2012. Global, local, and contagious investor sentiment. *Journal of Financial Economics*. 104, 272–287.
- Boone, A. L., White, J. T., 2015. The effect of institutional ownership on firm transparency and information production. *Journal of Financial Economics*. 117, 508–533.
- Boudoukh, J., Richardson, M., Whitelaw, R. F., 1994. Industry returns and the fisher effect. *The Journal of Finance*. 49, 1595–1615.
- Bustamante, M. C., 2015. Strategic investment and industry risk dynamics. *The Review of Financial Studies*. 28, 297–341.
- Bustamante, M. C., Donangelo, A., 2017. Product market competition and industry returns. *The Review of Financial Studies*. 30, 4216–4266.
- Chan, L. K., Karceski, J., Lakonishok, J., 1999. On portfolio optimization: Forecasting covariances and choosing the risk model. *The Review of Financial Studies*. 12, 937–974.
- Chen, A. Y., Velikov, M., 2023. Zeroing in on the expected returns of anomalies. *Journal of Financial and Quantitative Analysis*. 58, 968–1004.
- Cohen, L., Frazzini, A., 2008. Economic links and predictable returns. *The Journal of Finance*. 63, 1977–2011.

- Corhay, A., Kung, H., Schmid, L., 2020. Competition, markups, and predictable returns. *The Review of Financial Studies*. 33, 5906–5939.
- Da, Z., Engelberg, J., Gao, P., 2015. The sum of all fears investor sentiment and asset prices. *The Review of Financial Studies*. 28, 1–32.
- Daniel, K., Mota, L., Rottke, S., Santos, T., 2020. The cross-section of risk and returns. *The Review of Financial Studies*. 33, 1927–1979.
- DellaVigna, S., Pollet, J. M., 2007. Demographics and industry returns. *American Economic Review*. 97, 1667–1702.
- DeVault, L., Sias, R., Starks, L., 2019. Sentiment metrics and investor demand. *The Journal of Finance*. 74, 985–1024.
- Driscoll, J. C., Kraay, A. C., 1998. Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*. 80, 549–560.
- Fama, E. F., French, K. R., 1997. Industry costs of equity. *Journal of Financial Economics*. 43, 153–193.
- Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy*. 81, 607–636.
- Feng, G., Giglio, S., Xiu, D., 2020. Taming the factor zoo: A test of new factors. *The Journal of Finance*. 75, 1327–1370.
- Garcia, D., Hu, X., Rohrer, M., 2023. The colour of finance words. *Journal of Financial Economics*. 147, 525–549.
- Green, J., Hand, J. R., Zhang, X. F., 2017. The characteristics that provide independent information about average us monthly stock returns. *The Review of Financial Studies*. 30, 4389–4436.
- Grinblatt, M., Titman, S., Wermers, R., 1995. Momentum investment strategies, portfolio performance, and herding: A study of mutual fund behavior. *American Economic Review*. 85, 1088–1105.

- Grundy, B. D., Martin, J. S. M., 2001. Understanding the nature of the risks and the source of the rewards to momentum investing. *The Review of Financial Studies*. 14, 29–78.
- Harvey, C. R., Liu, Y., Zhu, H., 2016. ... and the cross-section of expected returns. *The Review of Financial Studies*. 29, 5–68.
- Hong, H., Torous, W., Valkanov, R., 2007. Do industries lead stock markets? *Journal of Financial Economics*. 83, 367–396.
- Hou, K., Robinson, D. T., 2006. Industry concentration and average stock returns. *The Journal of Finance*. 61, 1927–1956.
- Hou, K., Xue, C., Zhang, L., 2015. Digesting anomalies: An investment approach. *The Review of Financial Studies*. 28, 650–705.
- Hou, K., Xue, C., Zhang, L., 2020. Replicating anomalies. *The Review of Financial Studies*. 33, 2019–2133.
- Huang, D., Jiang, F., Tu, J., Zhou, G., 2015. Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies*. 28, 791–837.
- Jiang, F., Lee, J., Martin, X., Zhou, G., 2019. Manager sentiment and stock returns. *Journal of Financial Economics*. 132, 126–149.
- Kelly, B., Malamud, S., Pedersen, L. H., 2023. Principal portfolios. *The Journal of Finance*. 78, 347–387.
- Kelly, B., Pruitt, S., 2013. Market expectations in the cross-section of present values. *The Journal of Finance*. 68, 1721–1756.
- Kelly, B., Pruitt, S., 2015. The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*. 186, 294–316.
- Lee, C. M., So, E. C., 2017. Uncovering expected returns: Information in analyst coverage proxies. *Journal of Financial Economics*. 124, 331–348.
- Loughran, T., McDonald, B., 2011. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*. 66, 35–65.

- Menzly, L., Ozbas, O., 2010. Market segmentation and cross-predictability of returns. *The Journal of Finance*. 65, 1555–1580.
- Moskowitz, T. J., Grinblatt, M., 1999. Do industries explain momentum? *The Journal of Finance*. 54, 1249–1290.
- Nagel, S., 2005. Short sales, institutional investors and the cross-section of stock returns. *Journal of Financial Economics*. 78, 277–309.
- Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*. 55, 703–708.
- Sharifkhani, A., Simutin, M., 2021. Feedback loops in industry trade networks and the term structure of momentum profits. *Journal of Financial Economics*. 141, 1171–1187.
- Soo, C. K., 2018. Quantifying sentiment with news media across local housing markets. *The Review of Financial Studies*. 31, 3689–3719.
- Stambaugh, R. F., Yu, J., Yuan, Y., 2012. The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*. 104, 288–302.
- Tetlock, P. C., 2007. Giving content to investor sentiment: The role of media in the stock market. *The Journal of Finance*. 62, 1139–1168.
- Wold, H., 1966. Estimation of principal components and related models by iterative least squares. *Multivariate Analysis*, 391–420.
- Wold, H., 1975. Path models with latent variables: The nipals approach. In Quantitative sociology (pp. 307–357). Elsevier.

Figure 1: Average Number of Stocks per Industry (1993 - 2024)

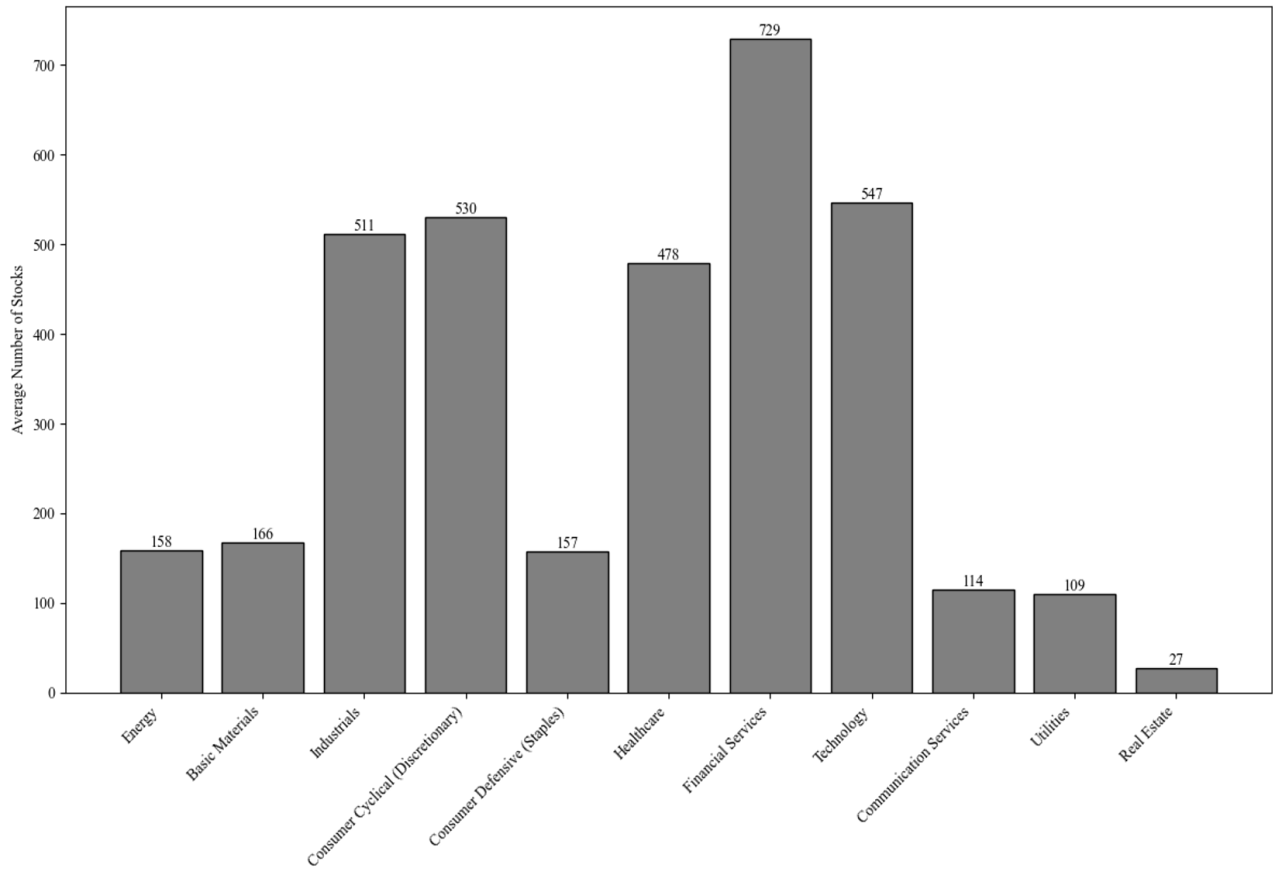


Table 1: Descriptive Statistics

This table presents the descriptive statistics of the variables used to construct the industry-level sentiment indices. *NIPO* is the total number of initial public offerings over the previous four quarters. *RIPO* is the NIPO-weighted average of first-day returns of initial public offerings over the previous four quarters. P^{D-ND} is the quarter-end log ratio of the value-weighted average market-to-book ratios of dividend payers to non-payers. *S* is the total volume of equity issues over the prior four quarters divided by the total volume of equity and debt issues over the prior four quarters. *CEFD* is the closed-end fund discount, defined as the average difference between the net asset values (NAV) of closed-end stock fund shares and their market prices. All proxies are calculated at the industry level with quarterly intervals. The industries are defined according to the Global Industry Classification Standard (GICS) as follows: Energy (10), Basic Materials (15), Industrials (20), Consumer Cyclical (Discretionary) (25), Consumer Defensive (Staples) (30), Healthcare (35), Financial Services (40), Technology (45), Communication Services (50), Utilities (55), and Real Estate (60).

Industry	(10)	(15)	(20)	(25)	(30)	(35)	(40)	(45)	(50)	(55)	(60)
<u>Panel A. <i>NIPO</i></u>											
Mean	2.84	1.17	5.36	7.73	1.43	11.73	6.58	13.84	2.45	0.28	1.05
SD	2.26	0.95	5.28	7.42	1.35	8.54	4.02	16.23	2.40	0.32	0.79
Min	0.00	0.00	0.00	0.75	0.00	0.25	0.50	0.25	0.00	0.00	0.00
P25	0.75	0.50	2.25	3.25	0.50	5.00	3.50	4.75	1.00	0.00	0.50
P50	2.25	1.00	3.00	4.75	1.00	9.25	5.75	7.75	2.00	0.25	1.00
P75	4.50	1.50	6.50	8.50	1.75	17.00	8.75	14.25	3.25	0.50	1.50
Max	8.00	4.00	22.50	31.75	6.50	37.50	16.75	76.00	13.50	1.25	3.75
<u>Panel B. <i>RIPO</i></u>											
Mean	5.84	7.81	16.16	21.88	6.18	14.88	11.01	26.63	12.74	1.66	3.60
SD	8.24	16.29	22.39	19.43	11.87	13.50	11.25	37.48	18.57	5.25	9.90
Min	-5.27	-21.93	-19.70	-3.33	-42.71	-2.22	-2.76	-45.23	-53.48	-17.58	-10.44
P25	0.46	0.29	5.51	10.74	0.73	6.82	4.09	13.99	1.46	0.00	0.00
P50	3.27	3.69	10.49	17.44	5.49	9.85	8.75	21.25	8.81	0.00	1.57
P75	8.09	8.84	17.91	23.86	11.07	18.48	12.78	26.41	20.00	2.91	4.04
Max	37.91	95.44	116.32	118.96	33.44	80.50	63.71	361.41	92.94	19.00	73.71
<u>Panel C. P^{D-ND}</u>											
Mean	-0.38	0.18	0.01	0.39	-0.62	0.59	-0.39	0.39	0.91	0.23	0.18
SD	0.42	0.46	0.58	0.34	0.71	0.52	0.47	0.31	0.47	0.50	0.54
Min	-2.00	-0.44	-1.62	-0.69	-4.42	-0.26	-2.14	-0.56	-0.06	-1.13	-1.42
P25	-0.58	-0.17	-0.38	0.20	-0.70	0.25	-0.55	0.21	0.69	-0.06	-0.13
P50	-0.37	0.01	-0.02	0.44	-0.55	0.55	-0.31	0.38	0.89	0.21	0.20
P75	-0.14	0.45	0.50	0.62	-0.39	0.85	-0.11	0.60	1.14	0.41	0.46
Max	0.77	1.61	1.06	1.09	0.91	2.38	0.77	1.26	2.83	2.24	1.52
<u>Panel D. <i>S</i></u>											
Mean	0.92	0.97	0.96	0.91	0.94	0.93	0.86	0.98	1.00	0.98	0.70
SD	0.85	0.62	0.61	0.69	0.65	0.48	0.96	0.36	0.65	0.49	1.03
Min	-3.73	-3.73	-3.73	-3.73	-3.73	-3.73	-3.73	-2.68	-3.70	-1.96	-3.73
P25	0.84	0.92	0.89	0.89	0.88	0.95	0.70	0.96	0.88	0.93	0.55
P50	0.94	1.00	0.97	0.99	0.99	0.99	0.93	0.99	0.96	1.02	0.80
P75	1.05	1.08	1.07	1.06	1.05	1.01	1.10	1.02	1.04	1.11	0.97
Max	3.72	3.72	3.72	3.72	3.72	1.64	3.72	1.83	3.72	2.35	3.72
<u>Panel E. <i>CEFD</i></u>											
Mean	-0.98	0.92	-3.97	-0.92	0.05	0.40	1.43	-0.21	0.00	-0.16	1.60
SD	1.83	4.43	5.15	3.35	0.30	1.68	1.07	1.09	0.00	0.93	1.96
Min	-5.52	-3.10	-24.57	-21.65	0.00	-4.88	-2.42	-6.35	0.00	-5.23	0.00
P25	-2.39	0.00	-7.80	0.00	0.00	0.00	1.07	0.00	0.00	0.00	0.49
P50	-0.30	0.00	-2.30	0.00	0.00	0.00	1.39	0.00	0.00	0.00	0.88
P75	0.00	0.00	0.00	0.00	0.00	0.20	2.09	0.00	0.00	0.00	2.09
Max	1.65	24.93	4.92	0.27	1.69	9.53	3.32	2.57	0.00	0.37	9.10

Table 2: Correlation between Industry-Level Sentiment and Market-Level Sentiment Indices

This table presents the correlation between each industry-level sentiment index and the market-level sentiment indices. PLS_MKT is the market index from Huang, Jiang, Tu, and Zhou (2015), and PCA_MKT is the market index from Baker and Wurgler (2024). The industries are defined according to the Global Industry Classification Standard (GICS) as follows: Energy (10), Basic Materials (15), Industrials (20), Consumer Cyclical (Discretionary) (25), Consumer Defensive (Staples) (30), Healthcare (35), Financial Services (40), Technology (45), Communication Services (50), Utilities (55), and Real Estate (60).

Industry	Correlation with	
	PLS_MKT	PCA_MKT
10	−0.17	0.24
15	0.13	0.08
20	0.43	0.43
25	0.38	0.63
30	−0.06	−0.13
35	0.11	0.13
40	0.52	0.60
45	0.50	0.52
50	0.17	0.34
55	0.10	0.17
60	0.21	0.19

Table 3: Regression Result of Market-Level Sentiment on the Principal Components of Industry-Level Sentiment

I first run PCA on the industry-level sentiment scores to obtain three principal components. This table presents the regression results where the market-level sentiment from Huang, Jiang, Tu, and Zhou (2015), which is available up to 2023, is regressed on these three principal components. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: $\text{SENT}_{MKT,t} = \alpha + \sum_{i=1}^3 \beta_i \text{PC}_{i,t-1} + \epsilon_t$	
	Coefficient
PC1 _{t-1}	0.06 (1.40)
PC2 _{t-1}	0.28*** (5.68)
PC3 _{t-1}	0.02 (0.45)
Adj. R ²	0.205
Panel B: $\text{SENT}_{MKT,t} = \alpha + \sum_{i=1}^3 \beta_i \text{PC}_{i,t} + \epsilon_t$	
	Coefficient
PC1 _t	0.03 (0.69)
PC2 _t	0.22*** (4.95)
PC3 _t	0.21*** (4.37)
Adj. R ²	0.250

Table 4: Regression Results of Industry-Level Sentiment on Market-Level Sentiment

This table reports regressions where the market-level sentiment from Huang, Jiang, Tu, and Zhou (2015), which is available up to 2023, is regressed on the sentiment scores of each industry. The industries are defined according to the Global Industry Classification Standard (GICS) as follows: Energy (10), Basic Materials (15), Industrials (20), Consumer Cyclical (Discretionary) (25), Consumer Defensive (Staples) (30), Healthcare (35), Financial Services (40), Technology (45), Communication Services (50), Utilities (55), and Real Estate (60). The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Industry	Slope	t-stat	Adj. R ²
10	0.09	(1.31)	0.014
15	−0.03	(−0.39)	0.001
20	−0.09	(−1.32)	0.014
25	0.39***	(6.01)	0.230
30	0.08	(1.01)	0.008
35	0.09	(1.27)	0.013
40	−0.01	(−0.12)	0.000
45	0.35***	(5.50)	0.200
50	0.25***	(3.64)	0.099
55	0.04	(0.58)	0.003
60	0.04	(0.49)	0.002

Table 5: Fama and MacBeth (1973) Regressions of Industry Return on Industry-Level Sentiment

This table reports the results of Fama and MacBeth (1973) regressions of quarterly industry returns ($R_{j,t}$) on lagged industry-level sentiment score ($SENT_{j,t-1}$) and industry momentum ($R_{j,t-1}$). The first two columns present estimates of the following specification:

$$R_{j,t} = \alpha + \beta_t^{SENT} SENT_{j,t-1} + e_{j,t}.$$

Column (3) and (4) present the results of the following specification:

$$R_{j,t} = \alpha + \beta_t^R R_{j,t-1} + e_{j,t}.$$

The last two columns extend the specification:

$$R_{j,t} = \alpha + \beta_t^{SENT} SENT_{j,t-1} + \beta_t^R R_{j,t-1} + e_{j,t}.$$

All standard errors are adjusted using Newey and West (1987) with four lags, and t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
	Equal-weighted	Value-weighted	Equal-weighted	Value-weighted	Equal-weighted	Value-weighted
$SENT_{j,t-1}$	0.69** (2.33)	0.64*** (2.99)			0.70*** (2.91)	0.63*** (3.16)
$R_{j,t-1}$			0.16*** (2.81)	0.04 (1.02)	0.16*** (3.10)	0.04 (1.13)
Adj. R^2	0.021	0.011	0.125	0.083	0.136	0.087

Table 6: Pooled Regressions of Industry Return on Industry-Level Sentiment

This table reports pooled regression results of industry returns on industry-level sentiment using data from 1993 to 2024. The dependent variable is the quarterly industry return. All variables are at quarterly frequency. The t-statistics are computed using Driscoll and Kraay (1998) standard errors. Coefficient estimates are reported with t-statistics in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
SENT _{<i>t</i>-1}	2.24*** (4.36)	1.00** (2.37)	2.26*** (4.14)	1.02** (2.29)					2.25*** (4.23)	0.90** (2.55)	2.30*** (4.00)	0.95** (2.46)	2.22*** (4.33)	2.25*** (4.13)		
R _{<i>j,t</i>-1}					0.01 (0.07)	0.13* (1.79)	-0.01 (-0.08)	0.10 (1.33)	-0.01 (-0.13)	0.12* (1.74)	-0.03 (-0.29)	0.09 (1.26)				
SENT_Market _{<i>t</i>-1}													-0.19 (-0.23)	-0.18 (-0.22)	-0.51 (-0.55)	-0.51 (-0.55)
Industry Fixed Effects	No	No	Yes	Yes	No	No	Yes	Yes	No	No	Yes	Yes	No	Yes	No	Yes
Time Fixed Effects	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes	No	No	No	No
Observations	1,364	1,364	1,364	1,364	1,364	1,364	1,364	1,364	1,364	1,364	1,364	1,364	1,364	1,364	1,364	1,364
Adj. <i>R</i> ²	0.030	0.665	0.044	0.679	0.000	0.666	0.013	0.677	0.030	0.670	0.044	0.681	0.030	0.044	0.001	0.014

Table 7: Time-Series Regressions of Industry Return on Industry-Level Sentiment

This table reports the results of time-series regressions of industry returns on industry-level sentiment for each industry. The first three columns present the results when the return of each industry is regressed on its one-quarter lagged industry-level sentiment index and its one-quarter lagged return. The last three columns show the results when the return of each industry is regressed on its one-quarter lagged industry-level sentiment index and the one-quarter lagged market sentiment. The sample period spans from 1993 to 2024. The industry-level sentiment index is constructed using Partial Least Squares with an expanding window. The market-level sentiment is from Huang, Jiang, Tu, and Zhou (2015). The industries are defined according to the Global Industry Classification Standard (GICS) as follows: Energy (10), Basic Materials (15), Industrials (20), Consumer Cyclical (Discretionary) (25), Consumer Defensive (Staples) (30), Healthcare (35), Financial Services (40), Technology (45), Communication Services (50), Utilities (55), and Real Estate (60). All standard errors are adjusted using Newey and West (1987) with four lags, and t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Industry	(1)			(2)		
	SENT _{<i>j,t-1</i>}	R _{<i>j,t-1</i>}	Adj. R^2	SENT _{<i>j,t-1</i>}	SENT _{<i>MKT,t-1</i>}	Adj. R^2
10	3.03* (1.90)	-0.07 (-0.78)	0.031	3.21* (1.96)	1.11 (0.51)	0.031
15	0.70 (0.65)	0.05 (0.53)	0.006	0.64 (0.58)	-0.28 (-0.19)	0.003
20	2.44** (2.51)	-0.06 (-0.74)	0.051	2.45** (2.49)	-0.07 (-0.05)	0.049
25	3.12*** (2.59)	-0.02 (-0.17)	0.051	4.03*** (2.94)	2.34 (1.37)	0.068
30	1.18* (1.77)	-0.11 (-1.22)	0.036	1.18* (1.71)	0.48 (0.53)	0.026
35	2.97** (2.49)	-0.12 (-1.35)	0.059	2.87** (2.28)	-0.76 (-0.47)	0.043
40	1.34 (1.64)	0.03 (0.32)	0.023	1.32 (1.49)	0.47 (0.43)	0.021
45	5.88*** (4.00)	-0.10 (-1.15)	0.115	6.72*** (3.86)	1.95 (0.86)	0.121
50	2.56** (2.12)	0.07 (0.78)	0.046	2.58** (2.02)	-1.95 (-1.17)	0.058
55	0.96* (1.66)	-0.01 (-0.08)	0.022	0.91 (1.56)	-0.10 (-0.13)	0.021
60	0.37 (0.35)	0.03 (0.31)	0.002	0.39 (0.37)	-1.08 (-0.77)	0.006

Table 8: Univariate Results of Portfolios Sorted by Industry-Level Sentiment Index

The sample period spans from 1993 to 2024. Each quarter, eleven industries are sorted into quintile portfolios based on their industry-level sentiment scores at quarter $t-1$, $t-2$, $t-3$, or $t-4$. Each quintile contains two industries, except for the third quintile, which contains three. The trading strategy involves buying industries with high sentiment and selling those with low sentiment. The result table presents the value-weighted returns from sorting industries into portfolios based on sentiment scores at quarters $t-1$, $t-2$, $t-3$, and $t-4$ (SENT_{t-1} , SENT_{t-2} , SENT_{t-3} , and SENT_{t-4}), respectively. α_{CAPM} , α_{FF4} , and α_{FF6} represent the CAPM alpha, the Fama-French three-factor plus momentum alpha, and the Fama-French five-factor plus momentum alpha for the high-minus-low portfolios, respectively. All standard errors are adjusted using Newey and West (1987) with four lags, and t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	SENT_{t-1}	SENT_{t-2}	SENT_{t-3}	SENT_{t-4}
Low (L)	1.38* (1.66)	1.45* (1.68)	2.29*** (2.86)	2.26*** (2.91)
2	2.16*** (2.92)	2.41*** (3.34)	1.61* (1.94)	2.49*** (2.83)
3	2.70*** (3.70)	2.84*** (3.96)	2.96*** (3.94)	2.33*** (3.16)
4	2.69*** (3.48)	2.08** (2.41)	2.87*** (3.76)	3.05*** (3.88)
High (H)	3.51*** (5.08)	3.38*** (4.61)	2.51*** (3.74)	2.21*** (3.57)
H-L	2.12*** (3.36)	1.93*** (2.93)	0.21 (0.39)	-0.05 (-0.09)
α_{CAPM}	2.39*** (3.44)	2.46*** (3.46)	0.72 (1.16)	0.24 (0.45)
α_{FF4}	2.14*** (3.63)	2.43*** (3.64)	1.11* (1.75)	0.89 (1.58)
α_{FF6}	2.05*** (3.55)	2.41*** (3.29)	0.93 (1.30)	0.38 (0.58)

Table 9: Bivariate Portfolio Analyses of Industry-Level Sentiment

This table shows the value-weighted results of the bivariate portfolio analyses. The sample period spans from 1993 to 2024. In the first sort, the eleven industries are divided into two portfolios (High and Low) based on momentum, industry concentration, institutional ownership, or analyst coverage. Since the number of industries is odd, the sixth industry in the ranking is assigned to both portfolios. Industry momentum is defined as the industry's return in the previous quarter. Industry concentration is measured using the Herfindahl-Hirschman Index (HHI). Institutional ownership is the fraction of a firm's shares held by institutional investors in each quarter. Analyst coverage is measured by the average number of analysts covering an industry in each quarter. In the second sort, each of these portfolios is further sorted into three sub-portfolios based on industry-level sentiment scores. The portfolios SENT_Low and SENT_High comprise the two industries with the lowest and highest industry-level sentiment scores, respectively. α_{CAPM} , α_{FF4} , and α_{FF6} represent the CAPM alpha, the Fama-French three-factor plus momentum alpha, and the Fama-French five-factor plus momentum alpha for the high-minus-low portfolios, respectively. All standard errors are adjusted using Newey and West (1987) with four lags, and t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Control							
	Industry Momentum		Industry Concentration		Institutional Ownership		Analyst Coverage	
	MOM_High	MOM_Low	HHI_High	HHI_Low	IO_High	IO_Low	AC_High	AC_Low
SENT_Low	2.17*** (2.83)	1.82** (2.36)	1.96*** (2.64)	1.71** (2.16)	2.28*** (2.78)	1.68* (1.84)	2.80*** (3.95)	0.97 (1.10)
SENT_Mid	2.34*** (3.06)	2.39*** (3.18)	2.54*** (2.96)	2.60*** (4.09)	2.46** (2.38)	2.26** (2.28)	2.44** (2.49)	2.19*** (2.61)
SENT_High	3.72*** (4.95)	2.86*** (3.71)	3.45*** (4.49)	3.75*** (5.10)	2.63*** (2.76)	3.01*** (3.55)	2.13** (2.35)	3.00*** (3.53)
SENT_High-Low	1.55*** (2.97)	1.05 (1.58)	1.49*** (2.90)	2.05*** (3.43)	0.35 (0.59)	1.32* (1.66)	-0.67 (-0.96)	2.03** (2.13)
α_{CAPM}	1.62*** (3.27)	1.06 (1.32)	1.15** (2.23)	2.04*** (2.94)	0.20 (0.34)	1.92* (1.96)	-1.24 (-1.65)	2.45** (2.25)
α_{FF4}	1.52*** (3.13)	0.71 (0.99)	1.13** (2.32)	1.75*** (2.62)	0.30 (0.41)	1.78** (1.98)	-1.28 (-1.62)	2.85*** (2.90)
α_{FF6}	1.93*** (3.50)	0.46 (0.68)	1.27** (2.07)	2.04*** (2.67)	-0.04 (-0.06)	2.26** (2.45)	-1.17 (-1.47)	2.76*** (2.89)

Table 10: Supplier-Customer Relationships: Returns and Sentiment Spillovers Across Industries

The sample period spans from 1993 to 2024. Panel A presents the regression results of supplier industry returns on lagged customer industry returns. Panel B reports the regression analysis of supplier industry sentiment on lagged customer industry sentiment. Panel C shows the regression results where the dependent variable is supplier industry returns in quarter t and the independent variable is customer industry sentiment in quarter $t - 2$. The industries are defined according to the Global Industry Classification Standard (GICS) as follows: Energy (10), Basic Materials (15), Industrials (20), Consumer Cyclical (Discretionary) (25), Consumer Defensive (Staples) (30), Healthcare (35), Financial Services (40), Technology (45), Communication Services (50), Utilities (55), and Real Estate (60). The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: $\mathbf{R}_t^{\text{Supplier}} = \alpha + \beta R_{t-1}^{\text{Customer}} + \epsilon_t$

FMB		Regression Results by Industry										
	Regression	10	15	20	25	30	35	40	45	50	55	60
Coefficient	1.38***	4.21***	3.65**	2.69***	4.31***	1.94**	1.80**	3.13*	3.02***	2.40***	-0.75	1.31
t-statistics	(6.53)	(2.63)	(2.56)	(2.61)	(3.76)	(2.52)	(2.36)	(1.66)	(4.07)	(4.77)	(-1.56)	(0.81)
Adj. R^2	0.030	0.049	0.047	0.048	0.095	0.045	0.04	0.02	0.11	0.145	0.018	0.005

Panel B: $\mathbf{SENT}_t^{\text{Supplier}} = \alpha + \beta \mathbf{SENT}_{t-1}^{\text{Customer}} + \epsilon_t$

FMB		Regression Results by Industry										
	Regression	10	15	20	25	30	35	40	45	50	55	60
Coefficient	6.75***	12.98***	10.27***	6.42***	10.34***	5.50***	7.16***	6.32**	5.35***	3.04***	2.23***	11.18*
t-statistics	(20.13)	(10.07)	(5.46)	(3.65)	(8.16)	(8.64)	(8.49)	(2.01)	(5.93)	(3.02)	(2.64)	(1.91)
Adj. R^2	0.227	0.448	0.193	0.096	0.348	0.374	0.366	0.031	0.22	0.068	0.053	0.028

Panel C: $\mathbf{R}_t^{\text{Supplier}} = \alpha + \beta \mathbf{SENT}_{t-2}^{\text{Customer}} + \epsilon_t$

FMB		Regression Results by Industry										
	Regression	10	15	20	25	30	35	40	45	50	55	60
Coefficient	14.35***	28.95	18.20	28.02	35.19**	7.31	4.65	61.18**	24.37*	28.20***	6.05	12.11*
t-statistics	(2.94)	(1.32)	(0.76)	(1.53)	(2.13)	(1.34)	(0.61)	(2.03)	(1.85)	(3.15)	(0.80)	(1.72)
Adj. R^2	0.021	0.014	0.005	0.018	0.035	0.014	0.003	0.032	0.027	0.074	0.005	0.023

Table 11: Univariate Results of Portfolios Sorted by Industry-Level Sentiment Index Using RavenPack Data

Each quarter, eleven industries are sorted into quintile portfolios based on their industry-level sentiment scores at quarter $t-1$, $t-2$, $t-3$, or $t-4$. The sample period spans from 2000 to 2024. The trading strategy involves buying industries with high sentiment and selling those with low sentiment. The result table presents the value-weighted returns from sorting industries into portfolios based on sentiment scores at quarters $t-1$, $t-2$, $t-3$, and $t-4$ (SENT_{t-1} , SENT_{t-2} , SENT_{t-3} , and SENT_{t-4}), respectively. Panel A presents the value-weighted results for portfolios sorted by the industry-level sentiment index constructed using RavenPack data, which is available from 2000. Panel B displays the results for portfolios sorted by the industry-level sentiment index constructed using the PLS method with an expanding window. α_{CAPM} , α_{FF4} , and α_{FF6} represent the CAPM alpha, the Fama-French three-factor plus momentum alpha, and the Fama-French five-factor plus momentum alpha for the high-minus-low portfolios. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: RavenPack				
	SENT_{t-1}	SENT_{t-2}	SENT_{t-3}	SENT_{t-4}
Low (L)	1.01 (0.96)	0.58 (0.51)	1.35 (1.28)	1.91* (1.71)
2	1.38 (1.29)	1.64 (1.45)	2.49** (2.44)	2.20** (2.18)
3	2.45*** (2.82)	2.20*** (2.63)	2.39*** (2.84)	2.25** (2.54)
4	2.64*** (3.42)	2.48*** (3.09)	1.57* (1.74)	1.16 (1.22)
High (H)	3.46*** (4.59)	3.87*** (5.00)	2.51** (2.48)	2.38** (2.38)
H-L	2.45*** (2.97)	3.29*** (3.17)	1.15* (1.73)	0.47 (0.69)
α_{CAPM}	3.06*** (3.77)	3.82*** (3.72)	1.33* (1.87)	0.72 (0.99)
α_{FF4}	2.55*** (3.40)	3.58*** (3.56)	1.28** (2.19)	0.70 (1.02)
α_{FF6}	2.12** (2.48)	3.02*** (3.14)	0.83 (1.59)	0.45 (0.58)

Panel B: PLS				
	SENT_{t-1}	SENT_{t-2}	SENT_{t-3}	SENT_{t-4}
Low (L)	1.22 (1.17)	0.94 (0.88)	1.73* (1.76)	1.56* (1.74)
2	1.76** (2.02)	2.46*** (2.88)	1.34 (1.31)	2.11* (1.90)
3	2.57*** (2.91)	2.50*** (2.98)	2.85*** (3.11)	2.07** (2.29)
4	2.35** (2.49)	1.76 (1.64)	2.43*** (2.60)	3.06*** (3.16)
High (H)	2.98*** (3.69)	2.97*** (3.57)	2.27*** (2.79)	1.92** (2.55)
H-L	1.76** (2.49)	2.03*** (2.76)	0.55 (0.83)	0.36 (0.56)
α_{CAPM}	2.07*** (2.65)	2.56*** (3.41)	0.76 (1.06)	0.48 (0.75)
α_{FF4}	1.95*** (2.91)	2.74*** (3.99)	1.04 (1.45)	0.93 (1.44)
α_{FF6}	1.97*** (2.95)	2.78*** (3.63)	1.05 (1.22)	0.37 (0.50)

Table 12: Bivariate Portfolio Analysis of Industry-Level Sentiment Using RavenPack Data

This table shows the value-weighted results of the bivariate portfolio analyses from 2000 to 2024. In the first sort, the eleven industries are divided into two portfolios (High and Low) based on industry momentum, industry concentration, institutional ownership, or analyst coverage. Since the number of industries is odd, the sixth industry in the ranking is assigned to both portfolios. Industry momentum is defined as the industry's return in the previous quarter. Industry concentration is measured using the Herfindahl-Hirschman Index (HHI). Institutional ownership is the fraction of a firm's shares held by institutional investors in each quarter. Analyst coverage is measured by the average number of analysts covering an industry in each quarter. In the second sort, each of these portfolios is further sorted into three sub-portfolios based on industry-level sentiment scores. Panels A and B present the results for portfolios sorted by the industry-level sentiment index based on RavenPack data and the PLS sentiment index with an expanding window, respectively. α_{CAPM} , α_{FF4} , and α_{FF6} represent the CAPM alpha, the Fama-French three-factor plus momentum alpha, and the Fama-French five-factor plus momentum alpha for the high-minus-low portfolios, respectively. All standard errors are adjusted using Newey and West (1987) with four lags, and t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Panel A: RavenPack

	Industry Momentum		Industry Concentration		Institutional Ownership		Analyst Coverage	
	MOM_High	MOM_Low	HHI_High	HHI_Low	IO_High	IO_Low	AC_High	AC_Low
SENT_Low	1.61 (1.49)	1.50 (1.36)	1.26 (1.15)	1.40 (1.24)	2.05* (1.73)	0.68 (0.53)	1.89* (1.68)	0.20 (0.16)
SENT_Mid	2.45*** (3.32)	1.73* (1.76)	2.25** (2.33)	2.17** (2.42)	2.11* (1.68)	1.88 (1.59)	2.21* (1.77)	2.01* (1.88)
SENT_High	3.10*** (3.57)	3.18*** (4.13)	3.19*** (4.05)	3.70*** (4.53)	2.45** (2.42)	3.18*** (3.79)	2.64*** (3.05)	3.15*** (3.84)
SENT High-Low	1.49** (2.41)	1.68** (2.32)	1.94*** (2.62)	2.29** (2.01)	0.40 (0.55)	2.50** (2.27)	0.75 (0.76)	2.95*** (2.89)
α_{CAPM}	1.93*** (3.06)	2.11*** (3.05)	2.32*** (3.15)	2.95** (2.53)	0.23 (0.29)	3.33*** (3.02)	1.14 (1.18)	3.42*** (3.11)
α_{FF4}	1.73** (2.52)	1.99*** (2.79)	2.01*** (2.93)	2.48** (2.37)	0.00 (0.00)	2.93*** (2.75)	0.86 (0.91)	3.04*** (2.62)
α_{FF6}	1.33** (1.96)	1.81** (2.07)	1.32* (1.83)	2.26* (1.92)	0.39 (0.48)	2.83** (2.39)	0.88 (0.83)	2.43** (2.00)

Panel B: PLS

	Industry Momentum		Industry Concentration		Institutional Ownership		Analyst Coverage	
	MOM_High	MOM_Low	HHI_High	HHI_Low	IO_High	IO_Low	AC_High	AC_Low
SENT_Low	1.82* (1.93)	1.71* (1.77)	1.66* (1.80)	1.25 (1.29)	1.56 (1.46)	0.60 (0.59)	2.56*** (2.91)	0.46 (0.44)
SENT_Mid	2.14** (2.30)	2.04** (2.19)	2.05** (2.05)	2.53*** (3.27)	2.18* (1.75)	1.74 (1.41)	2.31* (1.88)	2.12** (2.03)
SENT_High	3.18*** (3.70)	2.55*** (2.66)	2.99*** (3.39)	3.48*** (3.93)	2.64** (2.36)	3.06** (2.44)	1.89* (1.72)	2.78*** (2.68)
SENT High-Low	1.36** (2.33)	0.85 (1.07)	1.33*** (2.61)	2.23*** (3.25)	1.08* (1.89)	2.46** (2.04)	-0.67 (-0.79)	2.32** (2.15)
α_{CAPM}	1.47*** (2.62)	0.96 (1.02)	1.08* (1.94)	2.40*** (3.19)	1.01* (1.75)	2.72* (1.89)	-1.02 (-1.17)	2.71** (2.14)
α_{FF4}	1.47*** (2.93)	0.77 (1.02)	1.01** (2.08)	2.32*** (3.35)	1.27** (2.26)	2.42** (1.98)	-1.07 (-1.19)	3.00*** (2.68)
α_{FF6}	1.92*** (3.28)	0.75 (1.06)	1.23** (1.97)	2.73*** (3.54)	0.86 (1.63)	3.05** (2.40)	-0.87 (-0.94)	3.43*** (3.19)

Table 13: Univariate Results of Portfolios Sorted by Industry-Level Sentiment Index from Industry Report Textual Analysis

Each quarter, eleven industries are sorted into quintile portfolios based on the previous quarter's industry-level sentiment scores. The trading strategy involves buying industries with high sentiment and selling those with low sentiment. The first two columns of the table present equal-weighted and value-weighted results for portfolios sorted by the industry-level sentiment index, constructed via textual analysis of industry reports using the sentiment dictionary of Loughran and McDonald (2011) (SENT_LM). The next two columns display results for portfolios sorted by the industry-level sentiment index, constructed via textual analysis of industry reports using a machine learning dictionary of finance words from Garcia, Hu, and Rohrer (2023) (SENT_GHR). α_{CAPM} , α_{FF4} , and α_{FF6} represent the CAPM alpha, the Fama-French three-factor plus momentum alpha, and the Fama-French five-factor plus momentum alpha for the high-minus-low portfolios. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	SENT_LM		SENT_GHR	
	Equal-weighted	Value-weighted	Equal-weighted	Value-weighted
Low (L)	3.78*** (2.86)	2.05** (2.56)	4.30*** (3.28)	2.47*** (3.31)
2	3.78*** (2.91)	2.77*** (3.00)	4.70*** (4.07)	2.89*** (3.90)
3	6.44*** (4.43)	3.87*** (4.81)	5.91*** (4.05)	3.79*** (4.09)
4	5.31*** (5.03)	3.29*** (4.51)	5.45*** (4.32)	3.74*** (4.57)
High (H)	5.33*** (3.91)	3.67*** (3.71)	4.54*** (3.64)	2.79*** (2.72)
H-L	1.55** (2.25)	1.63** (2.12)	0.24 (0.28)	0.32 (0.38)
α_{CAPM}	1.34* (1.87)	1.14 (1.23)	0.57 (0.60)	-0.09 (-0.08)
α_{FF4}	1.83** (2.43)	1.36* (1.83)	1.23 (1.30)	0.46 (0.46)
α_{FF6}	2.29*** (3.04)	1.68** (2.35)	1.51 (1.62)	0.53 (0.50)

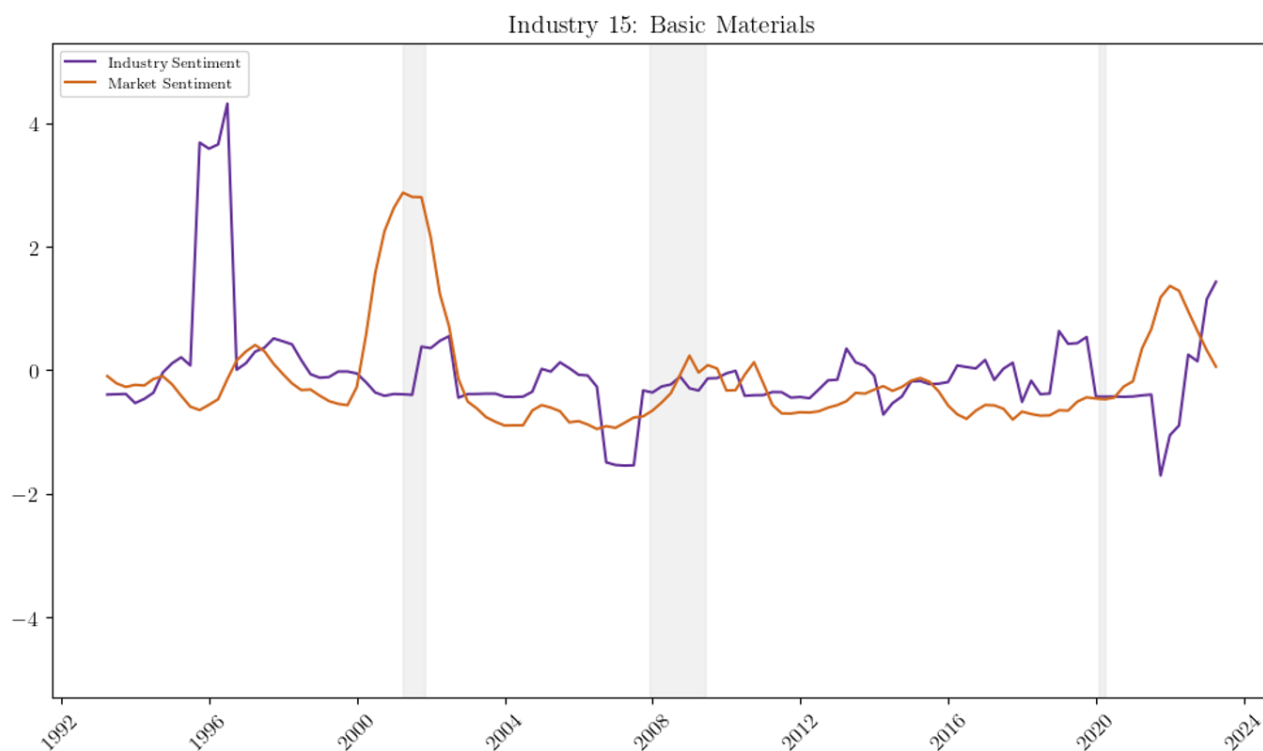
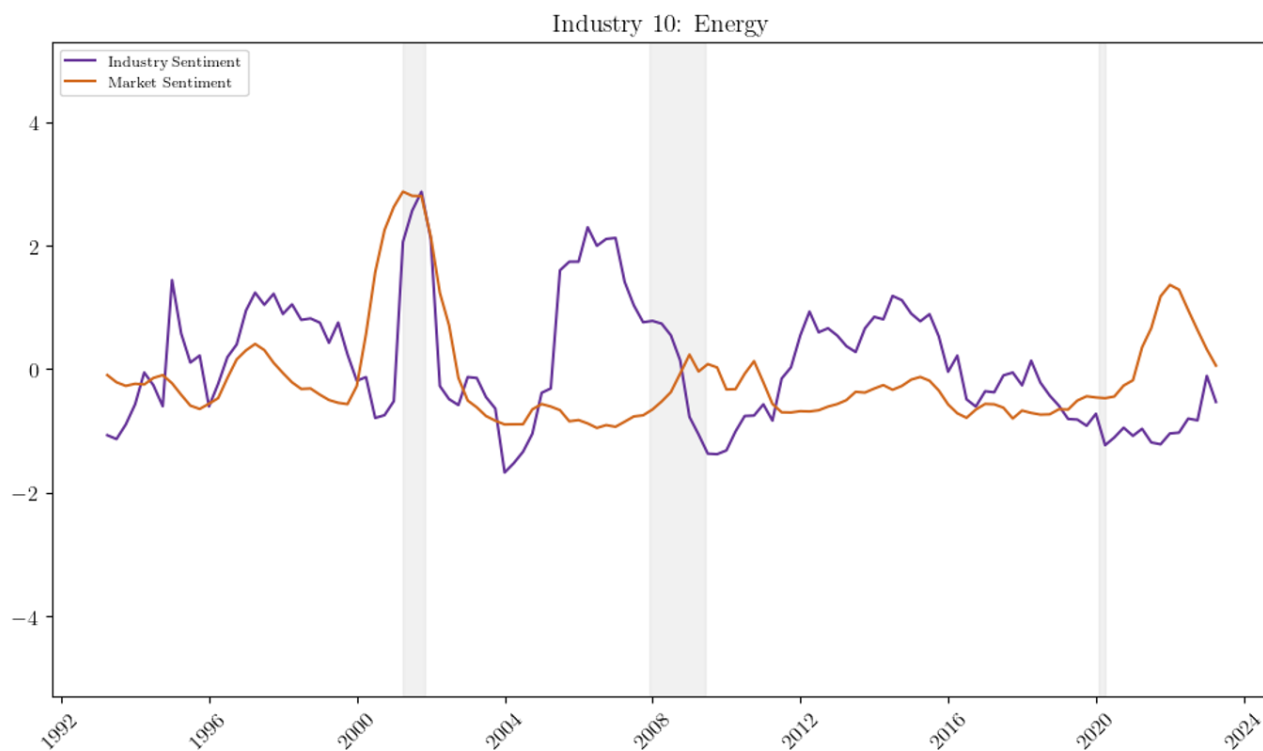
Table 14: Univariate Results of Portfolios Sorted by Changes in Sentiment Index

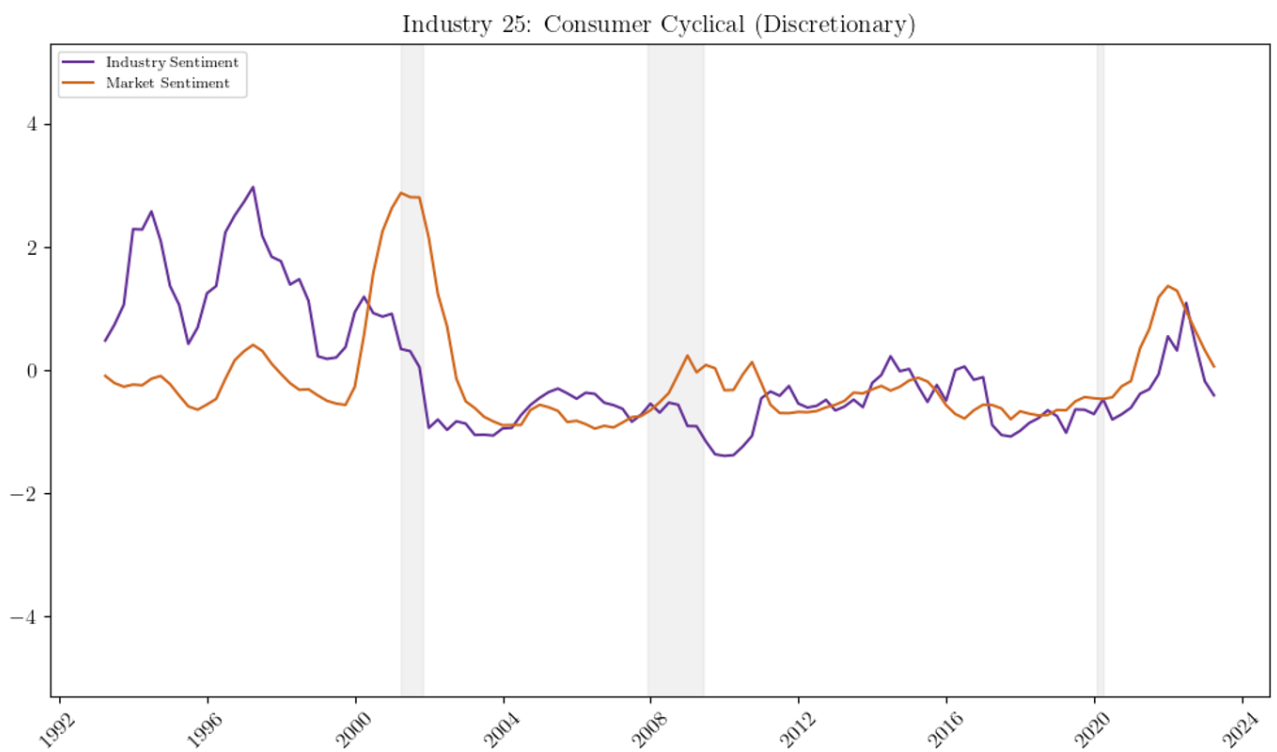
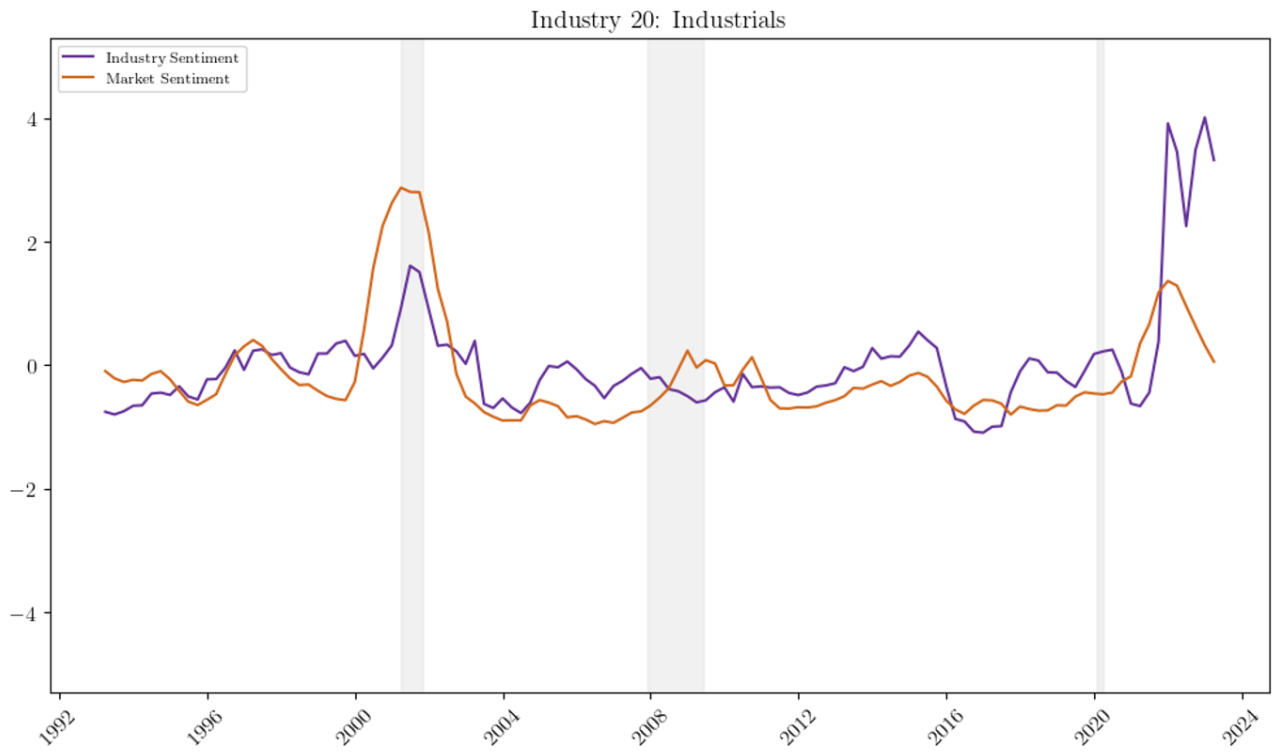
Each quarter, the eleven industries are sorted into quintile portfolios based on the previous quarter's change in industry-level sentiment scores. The trading strategy involves buying industries with the highest change in sentiment and selling those with the lowest change in sentiment. The first two columns of the table present equal-weighted and value-weighted results for portfolios sorted by the change in industry-level sentiment (Change_PLS), where the change is computed using the PLS approach. The next two columns display results for portfolios sorted by the change in the industry-level sentiment index (Change_PCA), where the change is constructed using the PCA approach. α_{CAPM} , α_{FF4} , and α_{FF6} represent the CAPM alpha, the Fama-French three-factor plus momentum alpha, and the Fama-French five-factor plus momentum alpha for the high-minus-low portfolios. The t-statistics are reported in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Change_PLS		Change_PCA	
	Equal-weighted	Value-weighted	Equal-weighted	Value-weighted
Low (L)	5.47*** (6.15)	2.35*** (3.36)	5.03*** (5.34)	2.04*** (2.59)
2	4.51*** (4.84)	2.30*** (2.71)	5.39*** (6.06)	2.95*** (3.87)
3	4.47*** (5.05)	2.33*** (3.05)	4.79*** (4.86)	2.55*** (3.29)
4	5.75*** (6.37)	2.48*** (3.80)	5.00*** (5.96)	2.49*** (3.66)
High (H)	5.39*** (5.27)	3.16*** (4.19)	5.21*** (5.21)	2.33*** (2.92)
H-L	-0.08 (-0.12)	0.80 (1.23)	0.18 (0.24)	0.29 (0.43)
α_{CAPM}	-0.10 (-0.16)	0.85 (1.28)	0.24 (0.31)	0.42 (0.56)
α_{FF4}	-0.12 (-0.18)	0.55 (0.94)	0.35 (0.46)	0.30 (0.44)
α_{FF6}	-0.03 (-0.05)	0.67 (1.20)	0.18 (0.23)	0.12 (0.16)

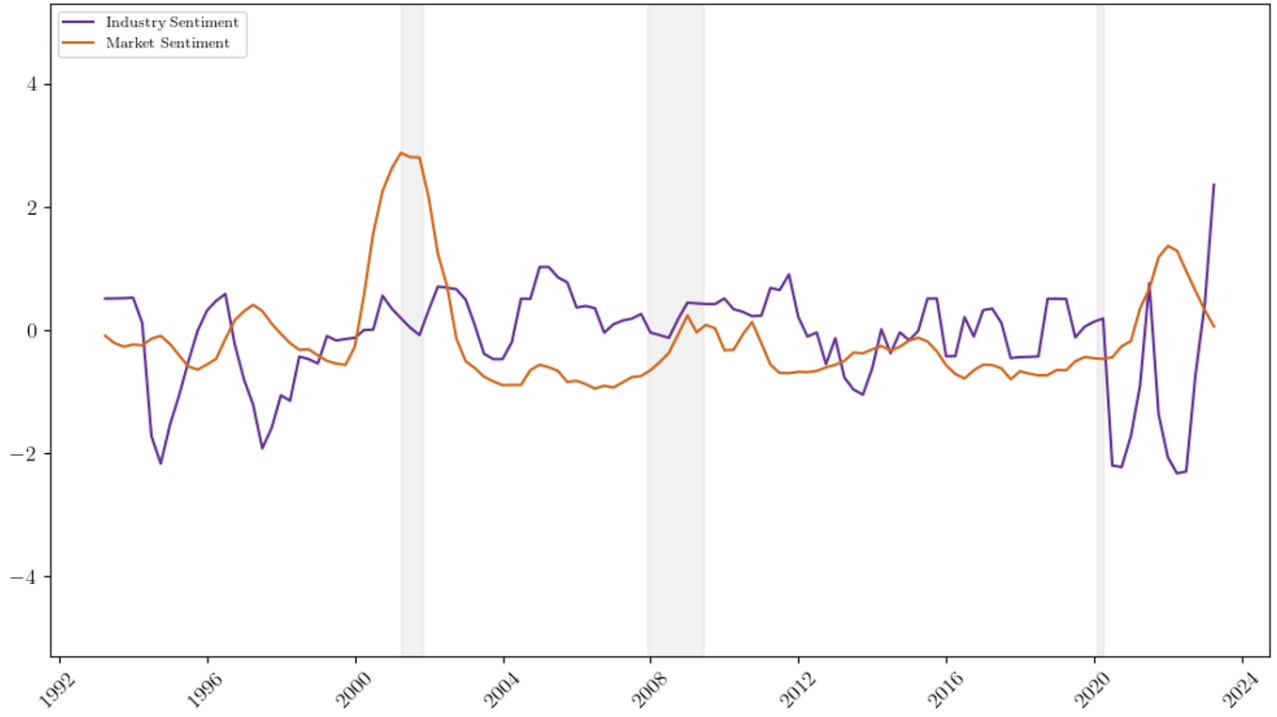
Appendix. Comparison of Industry-Level and Market-Level Sentiment

The figures below compare the sentiment index of each industry with the market-level sentiment index from Huang et al. (2015) (available up to December 2023). The 11 industries are defined according to the Global Industry Classification Standard (GICS). Each industry's sentiment index is constructed using the Partial Least Squares method with an expanding window. The shaded areas indicate National Bureau of Economic Research (NBER) recession periods.





Industry 30: Consumer Defensive (Staples)



Industry 35: Healthcare

