Economics Letters 242 (2024) 111900

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

Do optimistic portfolios outperform pessimistic portfolios: Evidence from textual sentiment

Yudhvir Seetharam *, Kingstone Nyakurukwa

University of the Witwatersrand, School of Economics and Finance, Johannesburg, South Africa

ARTICLE INFO

JEL classification:
G12
G40

Keywords:
Asset pricing
Sentiment-augmented asset pricing model
Behavioural finance
Investor sentiment

ABSTRACT

We examine whether textual sentiment (from news and social media) explains the cross-section of stock returns. Sentiment scores are used to sort stocks into tercile portfolios daily. Various asset pricing models, including CAPM and Fama-French models, are used to assess the sentiment's impact on returns after accounting for traditional risk factors. Results show that portfolios with higher (optimistic) sentiment consistently yield better average returns across the two sentiment measures and the different asset pricing models.

1. Introduction

Shefrin (2008) proposed that the valuation of a security is influenced by the aggregation of both a fundamental component and a sentiment-based premium. As a result, academic researchers and investment professionals are frequently searching for novel investment tools that incorporate investor sentiment to refine existing models. In recent times, the significance of textual sentiment has gathered substantial interest from both investment professionals and academic researchers (Allen et al., 2019). Recent literature has shown that news content originating from conventional news sources (e.g. Allen et al., 2019) or social media platforms (e.g. Nyakurukwa and Seetharam, 2022) exerts an impact on investor sentiment, thereby influencing asset prices, asset volatility, and broader dynamics of financial markets.

The bounded rationality (Selten, 1990) of investors constrains their ability to process vast amounts of information, leading to cognitive and temporal limitations. As a result, investors often resort to heuristic decision-making to simplify the complex task of choosing among thousands of potential stock purchases. This attention-based heuristic allows investors to narrow down their choices and allocate their resources more efficiently, mitigating the effects of information overload (Barber and Odean, 2008). The advent of preprocessed sentiment scores from third-party suppliers has led to quicker ways to gauge the pulse of the market, offering investors an easy way of reducing information overload. In this study, we utilise two of such third-party sentiment scores (news sentiment and social media sentiment) to understand how they are related to stock returns. Our results show that higher sentiment portfolios tend to have higher average returns, and this pattern holds across the two sentiment measures and different asset pricing models used.

We proceed as follows; Section 2 outlines the methodology; Section 3 presents the results and Section 4 concludes.

2. Data and methods

2.1. Data

We get our social media and news media daily sentiment scores from Bloomberg. Our sample includes 29[1] constituents of the Dow Jones Industrial Average (DJIA). The DJIA was chosen because of the large size of the stocks, which means they are frequently mentioned on both news and social media platforms. Stock returns are estimated as the daily log returns of each sampled stock. Our sample period starts from 1 January 2016 to 30 April 2023, motivated by the availability of data. A full list of the sampled stocks and their respective sectors is shown in Table 1:

Firm-level news media and social media sentiment scores show the extent to which investors are optimistic or pessimistic about the prospects of a company estimated from online news articles and social media platforms respectively. Firm-level daily sentiment scores range from -1 to $+1$ with higher values showing investor optimism about the prospects

* Corresponding author.
E-mail addresses: Yudhvir.Seetharam@wits.ac.za (Y. Seetharam), knyakurukwa@gmail.com (K. Nyakurukwa).

[1] One stock was dropped because of insufficient data

https://doi.org/10.1016/j.econlet.2024.111900
Received 13 May 2024; Received in revised form 28 July 2024; Accepted 4 August 2024
Available online 5 August 2024

Table 1
List of sampled stocks.

	Symbol	Company	SECTOR
1	AAPL	Apple Inc.	Technology
2	AMGN	Amgen Inc.	Healthcare
3	AXP	American Express Company	Financials
4	BA	Boeing Company	Industrials
5	CAT	Caterpillar Inc.	Industrials
6	CRM	Salesforce Inc	Technology
7	CSCO	Cisco Systems Inc.	Technology
8	CVX	Chevron Corporation	Energy
9	DIS	Walt Disney Company	Communication Services
10	GS	Goldman Sachs Group Inc	Financials
11	HD	Home Depot Inc.	Consumer Discretionary
12	HON	Honeywell International Inc.	Industrials
13	IBM	International Business Machines	Technology
14	INTC	Intel Corporation	Technology
15	JNJ	Johnson & Johnson	Healthcare
16	JPM	JPMorgan Chase & Co	Financials
17	KO	Coca-Cola Company	Consumer Staples
18	MCD	McDonald's Corporation	Consumer Discretionary
19	MMM	3 M Company	Industrials
20	MRK	Merck & Co. Inc.	Healthcare
21	MSFT	Microsoft Corporation	Technology
22	NKE	NIKE Inc. Class B	Consumer Discretionary
23	PG	Procter & Gamble Company	Consumer Staples
24	TRV	Travelers Companies Inc	Financials
25	UNH	UnitedHealth Group Incorporated	Healthcare
26	V	Visa Inc. Class A	Financials
27	VZ	Verizon Communications Inc	Communication Services
28	WBA	Walgreens Boots Alliance Inc.	Healthcare
29	WMT	Walmart Inc.	Consumer Staples

Table 2
Alphas across the individual stock sentiment portfolios.

	1(Low)	2	3(High)	High-Low
Panel A: News sentiment				
Excess returns	−0.000455***	0.000446***	0.001047***	0.001502***
	(−8.244828)	(7.382992)	(20.877455)	(−19.683000)
CAPM alphas	−0.000371***	0.000534***	0.001124***	0.001495***
	(−6.828147)	(8.968067)	(22.739904)	(−19.88400)
FF3F alphas	−0.000366***	0.000537***	0.001124***	0.001490***
	(−6.735097)	(9.015377)	(22.736846)	(−19.820000)
FF5F alphas	−0.000335***	0.000576***	0.001154***	0.001489***
	(−6.164142)	(9.680111)	(23.370078)	(−19.818000)
Panel B: Social media sentiment				
Excess returns	0.000542*	0.000642**	0.001147***	0.001691***
	(−1.757320)	(2.312725)	(4.281782)	(−4.189000)
CAPM alphas	−0.000450	0.000732***	0.001216***	0.0016700***
	(−1.492500)	(2.684287)	(4.582861)	(−4.199000)
FF3F alphas	−0.000450	0.000734***	0.001219***	0.00167100***
	(−1.485100)	(2.691386)	(4.597034)	(−4.203000)
FF5F alphas	−0.000420	0.000763***	0.001248***	0.00166600***
	(−1.373940)	(2.797293)	(4.706912)	(−4.1910000)
Panel C: News sentiment-Social media sentiment				
Excess returns	0.000099	−0.000196	−0.000100	
	(0.298000)	(−0.663000)	(−0.341000)	
CAPM	0.0000790	−0.0001980	−0.0000920	
	(0.2850000)	(−0.6770000)	(−0.3130000)	
FF3F	0.0000840	−0.00019750	−0.0000950	
	(0.2970000)	(−0.6750000)	−0.327000	
FF5F	0.0000650	−0.0001870	−0.0000940	
	(0.2870000)	(−0.6400000)	−0.324000	

Notes: The table reports the time-series averages of excess returns as well as CAPM, FF3F and FF5F alphas for the individual stock sentiment portfolios. Newey and West's robust *t-statistics* are given in parentheses; ***, **, * show statistical significance at 1%, 5% and 10% respectively.

of a particular stock. Bloomberg utilises supervised machine-learning techniques to evaluate the sentiment of news articles and social media posts. Initially, human experts manually label each piece of content as positive, negative, or neutral based on its likely impact on an investor holding a long position in the mentioned security. This labelling answers whether the content would make the investor feel bullish, bearish, or neutral about their investment. These annotated datasets are used to train machine-learning models, including support vector machines. Once trained, these models can autonomously assign sentiment probabilities to new information, determining if it is positive, negative, or neutral. Bloomberg uses two platforms to estimate social media sentiment scores -X and StockTwits. The analyses from X[2] use only mentions with cashtags – which is a firm ticker symbol preceded by a dollar sign. This way, only financial *posts* related to listed stocks are used to come up with firm-level sentiment scores. The average firm-level daily sentiment and the weighted average story-level sentiment scores in the last 24 h.

2.2. Methods

At the end of each day, we sort our stocks into terciles based on their daily sentiment scores. The optimistic portfolio (Tercile 3) consists of stocks with the highest sentiment while the pessimistic portfolio (Tercile 1) shows the portfolio with the lowest sentiment. We compute average returns for the tercile portfolios and portfolio returns in excess of the risk-free rate daily. This allows us to find any pattern in the evolution of time series averaged returns across the tercile portfolios. Thus, if investor sentiment can explain cross-sectional returns, it is expected that average portfolio returns should increase monotonically from the lowest tercile (Tercile 1) to the highest tercile (Tercile 3). We also report the returns while controlling for the conventional risk factors[3] in the CAPM, Fama and French three-factor model (Fama and French, 1992) as well as the Fama and French five-factor model (Fama and French, 2015). The expectation is that if sentiment is a robust predictor of stock returns, there should be positive alphas left after controlling for the conventional risk factors.

3. Results

We present the results of the evolution of time series averages of excess returns[4] as well as alphas from the different asset pricing models in Table 2. Panel A and Panel B show the excess returns and factor alphas using news and social media sentiment while Panel C shows the difference between news and social media for all portfolio terciles. We observe that the excess returns of the tercile portfolios increase monotonically from the tercile with the least sentiment to the tercile with the highest sentiment. The pattern is the same whether online sentiment used in the univariate sorts is news sentiment (Panel A) or social media sentiment (Panel B). For portfolios sorted using news sentiment (Panel A), the difference of excess returns between the highest and lowest tercile portfolios (column 5) is 0.15% per day. The tercile alphas also increase monotonically from the lowest news sentiment tercile portfolio to the highest news sentiment tercile portfolio. The same pattern observed for univariate sorts using news sentiment can also be observed with univariate portfolio sorting using social media sentiment (Panel B). Tercile alphas increase monotonically from the lowest social media sentiment tercile portfolio to the highest tercile portfolio. In Fig. 1, we visualise the results presented in Table 2 graphically to clearly show the monotonic increase in the alphas across the tercile portfolios.

[2] Throughout the study, we use X, Twitter and social media interchangeably
[3] Data is obtained from the Fama and French website https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html
[4] The risk-free rate is derived from the Fama and French website which uses the returns of a 1-month Treasury Bill, with these returns presented on a monthly, weekly, or daily basis.

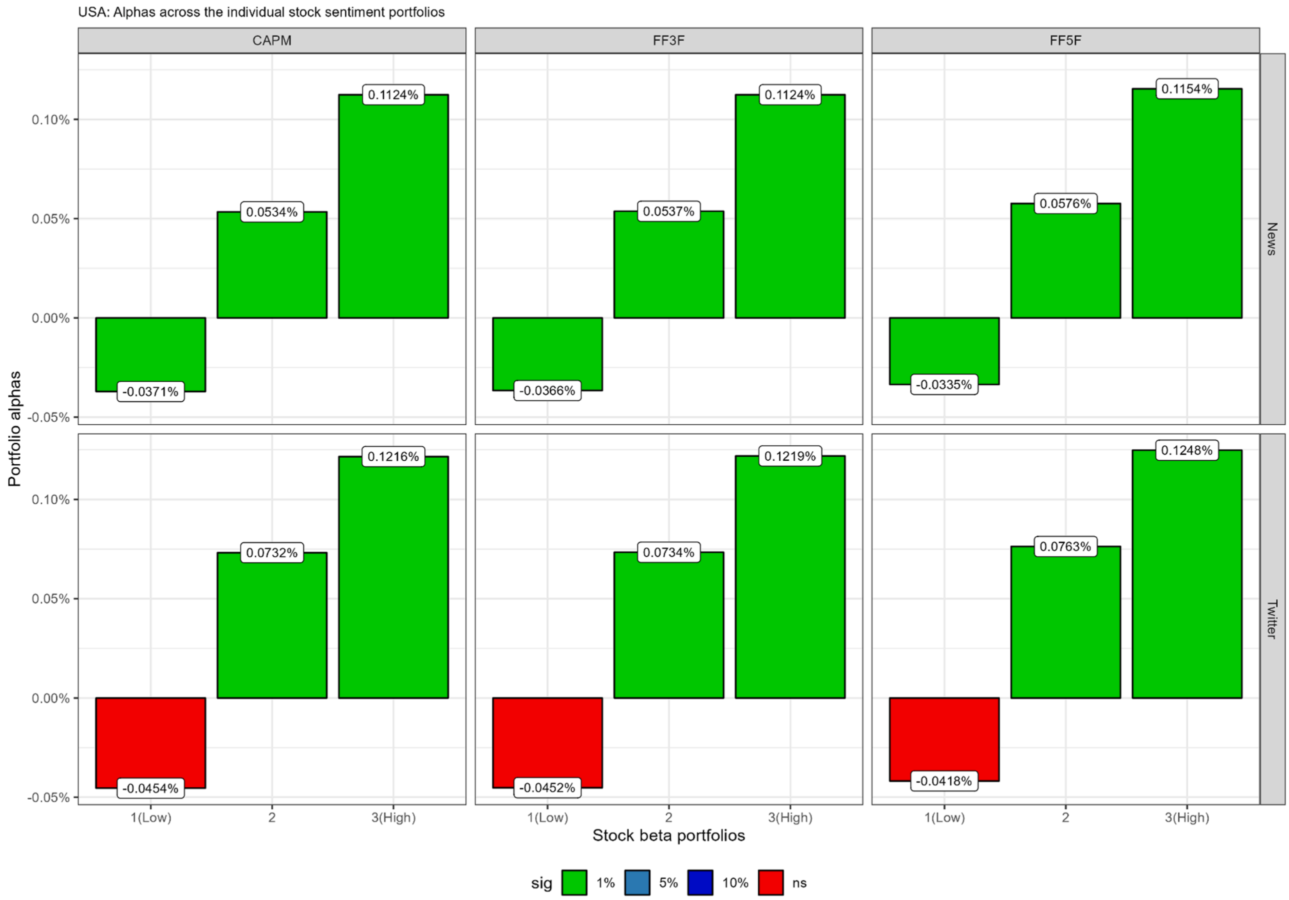

Fig. 1. Alphas for portfolios sorted using sentiment scores.

In Fig. 1, green, steel-blue and blue colours represent alphas that are statistically different from 0 at the 1%, 5% and 10% level of significance respectively, while the red colour represents statistically insignificant alphas. The monotonic increase in the alphas can be clearly seen using the bar graphs. For robustness, rather than using univariate portfolio sorts based on raw sentiment scores, we also sort portfolios using sentiment changes. A sentiment change today is defined as a change in online sentiment from the previous day. Even with sentiment changes, higher tercile sentiment portfolios outperform low tercile sentiment portfolios and the differences in the alphas between the highest sentiment change tercile portfolio and the lowest sentiment change tercile portfolio are positive and statistically significant across all the asset pricing models (see Fig. 2).

In our attempt to understand the role of online sentiment in explaining cross-sectional returns, we used two proxies of online investor sentiment. It would be ideal to know which one between the two performs better in terms of the tercile portfolios created. For each tercile portfolio constructed, we test whether the differences in the excess returns and alphas from the different asset pricing models are statistically different between news sentiment and social media sentiment. Panel C of Table 2 shows that the differences in alphas and excess returns are not statistically significantly different between news media sentiment sorted portfolios and social media sentiment sorted portfolios.

Incorporating sentiment into asset pricing models enhances their explanatory power by capturing behavioural dimensions of market activity that traditional factors might miss. Our findings demonstrate that sentiment is a robust predictor of stock returns, influencing them beyond what is explained by market risk, size, value, profitability, and investment factors. This interplay between sentiment and other factors shows the complexity of market dynamics and the importance of considering behavioural insights in financial analysis and investment strategies. Arteaga-Garavito et al. (2024) observe a notable rise in information diffusion (from news and social media) specifically during the hours when financial markets are active. These empirical findings indicate that monitoring news spread *via* news platforms is an effective method for capturing investors' information at a high frequency. In expanding the findings we get from this study, future studies could therefore incorporate intra-day data to establish if textual sentiment helps in generating alphas.

4. Conclusion

The study examines whether online investor sentiment can explain variations in cross-sectional stock returns. Stocks are sorted into portfolios based on their daily sentiment scores, creating optimistic (highest tercile) and pessimistic (lowest tercile) portfolios. The results show that higher sentiment portfolios tend to have higher average returns, and this pattern holds across the two sentiment measures and three asset pricing models. The study also compares sentiment from news media and social media (X) and finds that differences in excess returns and alphas are not statistically significant between the two sources. These findings can be of value to investment professionals seeking to incorporate sentiment-based strategies into their decision-making processes. The incorporation of sentiment into various asset pricing models, such as the CAPM, Fama-French three-factor model, and Fama-French five-factor model, could enhance the explanatory power of these models.

Data availability

The authors do not have permission to share data.

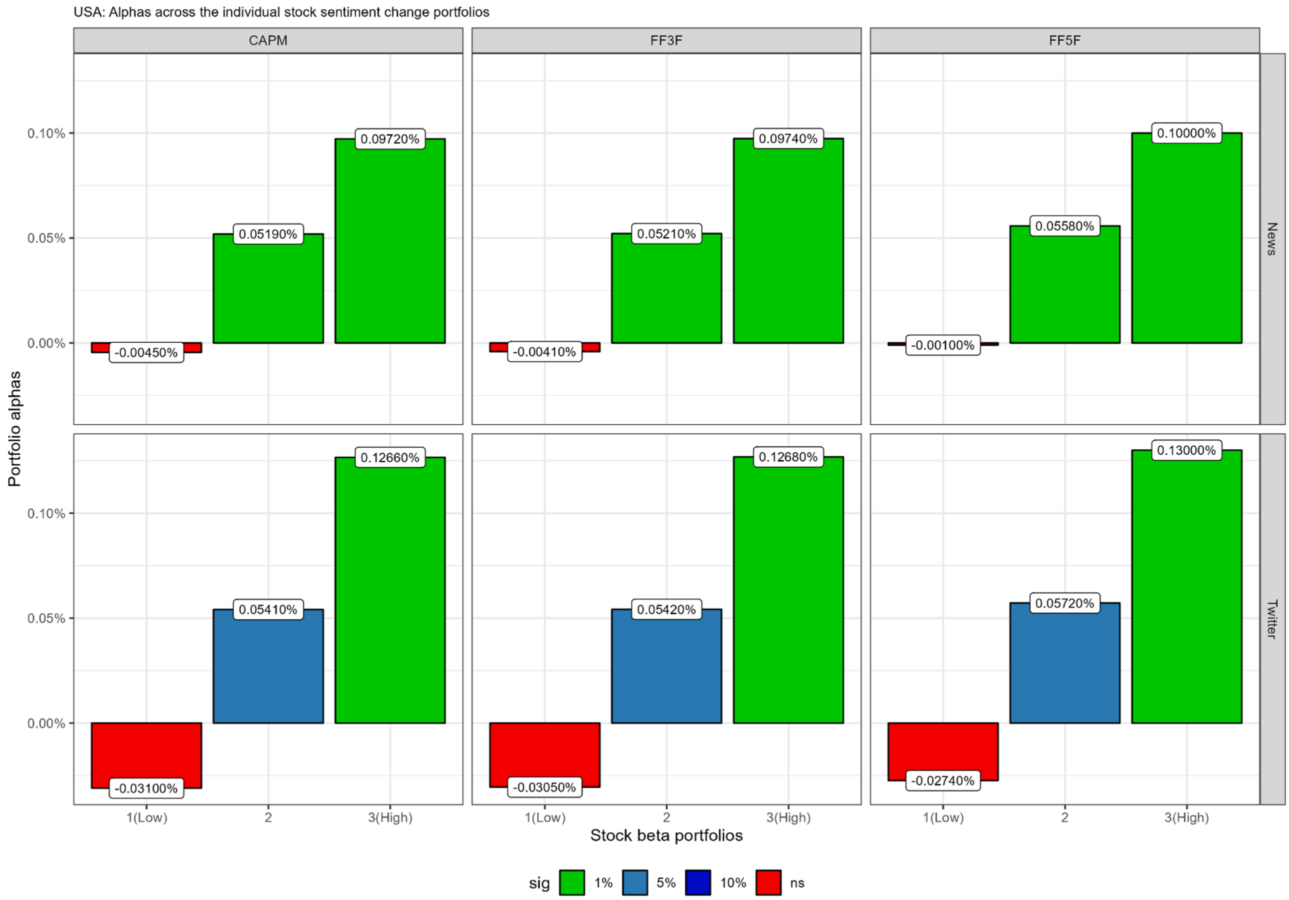

Fig. 2. Alphas across portfolios sorted using sentiment changes.

References

Allen, D.E., McAleer, M., Singh, A.K., 2019. Daily market news sentiment and stock prices. Appl. Econ. 51 (30), 3212–3235. https://ideas.repec.org//a/taf/applec/v51y2019i30p3212-3235.html.

Arteaga-Garavito, M.J., Croce, M.M., Farroni, P., Wolfskeil, I., 2024. When the markets get CO.V.I.D: contagion, viruses, and information diffusion. J. Financ. Econ. 157, 103850 https://doi.org/10.1016/j.jfineco.2024.103850.

Barber, B.M., Odean, T., 2008. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev. Financ. Stud. 21 (2), 785–818. https://doi.org/10.1093/rfs/hhm079.

Fama, E.F., French, K.R., 1992. The cross-section of expected stock returns. J. Finance 47 (2), 427–465. https://doi.org/10.2307/2329112.

Fama, E.F., French, K.R., 2015. A five-factor asset pricing model. J. Financ. Econ. 116 (1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010.

Nyakurukwa, K., Seetharam, Y., 2022. The wisdom of the Twitter crowd in the stock market: evidence from a fragile state. Afr. Rev. Econ. Finance 14 (1), 203–228. https://doi.org/10.10520/ejc-aref_v14_n1_a7.

Selten, R., 1990. Bounded rationality. J. Inst. Theor. Econ. 146 (4), 649–658. https://www.jstor.org/stable/40751353.

Shefrin, H., 2008. *A Behavioral Approach to Asset Pricing* [Elsevier Monographs]. Elsevier. https://econpapers.repec.org/bookchap/eeemonogr/9780123743565.htm.