Journal of International Money and Finance 136 (2023) 102864

Contents lists available at ScienceDirect

Journal of International Money and Finance

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

News-based sentiment and the value premium

Francesco A. Fabozzi [a], Abdolreza Nazemi [b]

[a] Yale School of Management, International Center for Finance, New Haven, CT, USA
[b] School of Economics and Business Engineering, Karlsruhe Institute of Technology, Karlsruhe, Germany

A R T I C L E   I N F O

Article history:
Available online 13 May 2023

JEL classification:
G12

Keywords:
Asset pricing
Style investing
Value premium
News sentiment

A B S T R A C T

The literature has documented that growth stocks are long-duration assets that are sensitive to shocks to the market discount rate, and value stocks, being short-duration assets, are sensitive to shocks to future cash flows. However, there is an ongoing debate as to whether the co-movement of value stocks with other value stocks and growth stocks with other growth stocks is due to similarities in cash-flow characteristics (i.e., the fundamental-based view) or due to a time-varying discount rate applied to cash flows (i.e., the sentiment-based view). While most studies take a fundamental approach to answer this question, we provide a sentiment-based explanation of the value premium by using the sentiment of news articles pertaining to cash-flow risks and discount-rate shocks. Our findings are consistent with the literature for the main drivers of the value premium but in support of the sentiment-based view, not through a rejection of the fundamental view but through a novel method for quantifying the time-varying discount rate that investors apply to cash flows.

1. Introduction

Many of the leading equilibrium asset pricing models have had difficulty replicating the value premium (i.e., why value stocks tend to outperform growth stocks, as well as why these groups of stocks tend to move together). The basic, intuitive explanation of the value premium is that value stocks are cheap relative to their cash-flow fundamentals and growth stocks are expensive relative to their cash-flow fundamentals. Thus, value stocks are priced at a discount and, as a result, have higher expected returns compared to expensive, growth stocks. However, this has led academics to question why value stocks are cheap when these are typically well-proven stocks with strong fundamentals.

In replicating the value premium in their equilibrium asset pricing model, Lettau and Wachter (2007) explain that growth stocks are particularly sensitive to the market discount rate because these stocks have cash-flows weighted more heavily in the future. Thus, changes in the discount rate have a stronger effect on growth stocks, just as long-term bonds are more sensitive to changes in interest rates compared to short-term bonds. That is, there is a strong *duration effect* for growth stocks. In contrast, value stocks are more sensitive to shocks to the underlying cash flows. Since the cash flows of value stocks are more heavily weighted in the short-term, shocks to the discount rate have less of an impact on the pricing of these stocks.

Given that value and growth stocks are sensitive to the discount rate and cash-flow shocks, respectively, why is there a higher risk premium for value stocks? The answer is that investors are more fearful of adverse shocks to cash flows than shocks to the discount rate. Campbell and Vuolteenaho (2004) explain that when there is negative news about future cash flows, wealth decreases but the investment opportunity set remains unchanged. That is, when aggregate cash flows decrease,

E-mail addresses: francesco.fabozzi@yale.edu (F.A. Fabozzi), nazemi@kit.edu (A. Nazemi)

https://doi.org/10.1016/j.jimonfin.2023.102864
0261-5606/

the value of all stocks decline. Conversely, when there are negative shocks to the discount rate, current wealth decreases but expected returns in the future increase. For this reason, investors assign a higher risk premium to value stocks because they are more sensitive to cash-flow shocks, which investors fear more.

This has led to the conception of a two-beta model proposed by Campbell and Vuolteenaho (2004), where news about cash flows represent "bad beta" and news about discount rates represent "good beta". However, since cash-flow shocks can be idiosyncratic, why do value stocks covary more with other value stocks and growth stocks with other growth stocks?.

Campbell et al. (2010) differentiate between the two prevailing justifications for the co-movement of value stocks and the co-movement of growth stocks. The first camp takes a *fundamental-based view*, which argues that the cash-flow characteristics among value and growth stocks are similar, resulting in similar price movements. The other camp takes a *sentiment-based view* where value stocks are simply stocks that are out of favor or neglected by investors, whereas growth stocks are more in favor and popular among investors, resulting in the value premium. The sentiment-based view implies that there is a time-varying discount rate that investors apply to cash flows, while the fundamental-based view is more consistent with the view that cash flows are discounted at a constant rate where only changes in news about future cash flows result in the time-variation of the value premium.

While it is entirely possible that the time-variation in the value premium and co-movement among value and growth stocks can be due to a combination of these two views, van Binsbergen et al. (2012) show that the price of short-term cash flows exhibit excess volatility, meaning that prices of short-term cash flows are more volatile than their subsequent realizations. This finding supports the sentiment-based view of the value premium because it implies that there is a time-varying discount rate applied to cash flows which causes the excess volatility.

While much of the literature attempting to delineate between the fundamental- and sentiment-based view of the value premium using market-derived variables as proxies for news about future cash flows, these variables are either backwards-looking or, if they are observed by the market ex-ante, only measure expectations as opposed to shocks.

In this paper, we attempt to quantify the shocks to the discount rate applied to cash flows (i.e., the sentiment-based view) by using actual news articles that pertain to companies' cash flows, and the idiosyncratic component of discount-rate news by isolating news articles pertaining to price movements and returns for individual companies. By doing so, we are able to isolate the risk premium that investors apply to cash flows by measuring the difference between the return on stocks with positive news pertaining to cash flows and the return on stocks with negative news pertaining to cash flows. Similarly, we measure the risk premium associated with idiosyncratic discount-rate shocks by measuring the difference between returns on stocks experiencing positive discount-rate news and the return on stocks with negative discount-rate news.

By utilizing the sentiment of the news articles pertaining to cash-flow and discount-rate news, the tonality of the news article will inherently isolate the shocks to each of these components in comparison to the market's expectation. Therefore, we capture how shocks to discount rates and cash flows are priced by the market, irrespective of value and growth classifications. That is, we identify the overall risk-premium associated with both sources of risk and see how these two "betas" affect the value premium. In doing so, we find support for the sentiment-based view as the value premium is positively related to news about future cash flows and negatively related to discount-rate shocks.

A sentiment-based explanation of the value premium would help reconcile two views on asset pricing in financial markets: the classical pricing model and behavioral finance. The classical pricing model states that market prices reflect all available information and it's impossible to beat the market through analysis or prediction. In contrast, behavioral finance highlights the importance of investors' irrational decisions and the emotional and psychological factors that influence the stock market and the perception of individual stocks by investors. Adopting a sentiment-based view of the value premium reconciles these two perspectives by suggesting that the value premium is influenced by discount-rate shocks driven by the emotional and psychological biases of investors and their tendency to overlook or undervalue value stocks.

In terms of practical implications, a sentiment-based explanation of the value premium suggests that investors can benefit from actively monitoring the market and investor sentiment to identify opportunities to invest in value stocks when they are undervalued. By incorporating news sentiment analysis into their investment strategy, investors can have a more active investment strategy and potentially capitalize on market inefficiencies caused by sentiment in order to enhance their returns.

The remainder of this paper is organized as follows. In Section 2, we provide a review of the literature and the associated methodologies that have been used to capture the sources of the value premium. Section 3 discusses the methodology used to capture both the "good" and "bad" beta, as well as the empirical tests used to quantify the relationship between cash-flow and discount-rate shocks with the value premium. This is followed by a discussion of the data sources and implications of using such data in Section 4. The results are presented in Section 5 and a discussion of the main findings and implications in Section 6. Our conclusions are provided in Section 7.

2. Literature review

In this section, we first discuss the general role of sentiment in different fields of financial economics and then focus on the effects of sentiment on the value premium.

Sentiment in financial economics refers to market participants' overall emotional state or attitude towards a particular asset, market, or economic environment. While there is no consensus among economists on a specific definition of sentiment, they have used various techniques and proxies to measure it and understand its impact on financial markets. These market sentiment measures can help analysts and investors gain valuable insights into the emotional state of market participants, which can help predict the potential direction of financial markets.

The underlying proxies for sentiment can be broadly divided into market-based, survey-based, and text-based data. Market-based measures estimate investor sentiment using financial market-related measures such as trading volume, extreme one-day returns, and implied volatility (e.g., Baker and Wurgler (2007) and Barber and Odean (2007)). However, these measures may not be available for all markets. Survey-based methods use survey data to create a sentiment index for a market (e.g., studies by Brown and Cliff (2004), Lemmon and Portniaguina (2006), and Qiu and Welch (2006)). However, these indexes have limitations, including high cost, availability latency, replication difficulty, lack of incentives for honest responses, and inability to be backtested (i.e., their historical accuracy cannot be verified). The third type of sentiment involves categorizing texts from sources such as financial news, social trading networks, media, blogs, and Twitter as positive, negative, or neutral. This type of analysis can be performed using lexicon-based or machine-learning methods. Lexicon-based methods are efficient but have limitations such as subjectivity, inability to account for multiple meanings of words, and slow reaction to new sentiment-related terms. Additionally, most systems that use sentiment dictionaries only measure sentiment for individual words and struggle to handle negations or consecutive groups of words effectively.

Due to advancements in computing resources and technologies, the use of machine-learning techniques in text analysis has grown in popularity in recent years. These methods mostly use basic features like word counts or frequency-based features and, by significantly reducing the need for human judgment, make it easier to utilize novel data sources and improve the replicability of the results (see Algaba et al. (2020) and Gentzkow et al. (2019)). With the development of Natural Language Processing (NLP) techniques (e.g., word embeddings, semantic analysis, and sentiment), sentiment analysis has evolved to incorporate more sophisticated features that capture the context and meaning of words. Integrating NLP techniques into sentiment analysis also allows for processing larger amounts of data, such as social media and news articles, to gain insights into public sentiment (see, for example, Jakubik et al. (2023),Frankel et al. (2022) and Hochberg et al. (2022)).

In our study, we aim to gain a deeper understanding of the value premium using news sentiment data from RavenPack. RavenPack is a specialized service that uses NLP methods for financial market sentiment analysis, which makes it a crucial tool for finance professionals. Its technology provides up-to-date and accurate sentiment data that traders, portfolio managers, and other finance professionals can use to make more informed investment decisions. Additionally, RavenPack's technology has the capability to identify and track emerging risks and opportunities in the financial market, which can help investors proactively manage their exposure to financial markets, thereby reducing the risk of losses.

Because of the broad scope of its coverage and various analytics, RavenPack's databases have been used in various areas of financial economics. For example, Ke et al. (2019) present evidence that news sentiment, as measured by RavenPack, predicts future stock returns. Jeon et al. (2022) investigates the connection between stock returns and global news events and discovers evidence for the relationship between the content and frequency of news and sudden changes in stock prices over the last decades. Hirshleifer and Sheng (2022) use RavenPack to identify attention events (e.g., executive-related news, scandals, legal issues, etc.) and examine their relationship to earnings announcements.

In the literature on the value premium, the two main explanations given are based on fundamental factors and sentiment. Papers advocating for the fundamental explanation examine the cash-flow characteristics of growth and value stocks. They argue that if the cash-flow fundamentals are similar within these categories, the movement of these stocks can be attributed to the similarity in their cash flows, not to a fluctuating discount rate applied to them. On the other hand, those arguing for the sentiment-based explanation of the value premium suggest that the co-movements of stocks are due to a varying discount rate applied to cash flows or that some variable is used as a substitute for cash-flow shocks instead of the similarities in the cash-flow fundamentals.

Cao (2015) empirically examines the influence of adjustment costs, operating leverage, and financial leverage on the value premium. She reports that financial leverage significantly influences the value premium, but the operating leverage does not drive the value premium. Whereas, Donangelo (2021) finds that operating leverage explains almost 50%, profitability and growth-based drivers explain the other 50% of the value premium. Gerakos and Linnainmaa (2018) find that book-to-market loses its predictive power for the value premium when they control for five years of changes in firm size. Klemola (2020) applies a simple model for studying the relationship between Internet search-based sentiment and value premium. Beccalli et al. (2022) study the effect of macroeconomic variables on the value premium. They report macroeconomic risk can model around 17% of the value premium variation.

He and Leippold (2020) state that value stocks have a larger exposure to both short-run and long-run risk than growth stocks. They find that short-run risk has a prediction power for predicting the value premium in-sample and out-of-sample. Fama and French (2021) cannot reject the hypothesis that expected premiums are significantly lower than expected in the second half of the July 1963-June 2019 period. Fama and French (2021) report an apparent decline in the value premium forecasting with book-to-market ratios. Gonçalves and Leonard (2023) mention that book equity, that is defined as the present value of cash flows under a common discount rate, is not a good proxy for fundamental equity for modeling the value premium. They find that the fundamental-to-market ratio captures the value premium better than alternative drivers used in the literature. After controlling for the fundamental-to-market ratio, the value premium and book equity have a statistically insignificant relation. Their results motivated us to introduce the new news-sentiment drivers related to cash-flow risks and discount-rate shocks that can be informative for estimating the value premium during extreme changes in the corporate environment.

Lakonishok et al. (1994) study 5-year returns on growth and value portfolios in relation to the aggregate market fundamentals and business cycles. While the fundamental-based view would suggest that value returns covary more with aggregate market fundamentals and the business cycle than that of growth returns, the authors find little evidence of this relationship. The conclusions imply a sentiment-based view, indicating that the excess return of value stocks is due to investor irrationality.

In contrast to the findings of the previous study, Bansal et al. (2005) and Hansen et al. (2008) find that the cash-flow fundamentals of value stocks are more sensitive to aggregate consumption growth than that of growth stocks. Thus, this finding supports the fundamental-based view that the value risk premium is due to similarities in cash-flow characteristics among growth stocks and value stocks.

While similarities in cash-flow fundamentals may indicate support for the fundamental-based view, the aforementioned studies do not provide support against the sentiment-based view. The similarities in cash flows only tells one part of the story. Although cash flows can be similar, this does not imply the absence of a time-varying discount rate being applied to those cash flows. van Binsbergen et al. (2012) provides a novel methodology for recovering the prices of short-term dividend payments using index options data. Although the authors focus on short-term dividend pricing of the aggregate market, and not growth and value stocks individually, they find excess volatility in the pricing of short-term cash distributions. The excess volatility supports the sentiment-based view because a time-varying discount rate is needed in order to justify the price volatility in relation to the volatility of subsequent dividend realizations.

An alternative strand of the literature justifies the value premium by investigating a duration-based explanation of the value premium. Cornell (1999) and Lettau and Wachter (2007) explain that growth stocks have cash flows that occur in the more distant future, and as a result, are more sensitive to shocks to the market discount rate because they affect how those cash flows are discounted. While these findings do not support or reject the fundamental-based view, it does suggest that discount rates cause the different risk characteristic in growth and value stocks. However, since these equilibrium models do not include the possibility of a separate discount factor applied to the cash flows, it is difficult to draw any conclusions regarding the sentiment-based view.

Further evidence against the fundamental-based view is provided by Barberis and Shleifer (2003) and Barberis et al. (2005) who find, again, that the co-movement in value stocks and co-movement in growth stocks is not due to commonalities in the cash-flow fundamentals of these firms. The conclusions drawn are similar to that of Lakonishok et al. (1994) in that the value premium is due to value stocks being neglected by investors and growth stocks being more popular or, in their words, "glamorous" among investors.

Campbell et al. (2010) set out to empirically test whether the value premium is a result of the fundamental-based view or the sentiment-based view. The stated goal of the paper is to directly test both of these views. Using return of equity to proxy for news about cash flows and price-earnings ratios to proxy for discount-rate shocks, the authors find that the strong relationship between growth stock returns and market discount-rate shocks and the strong relationship between value stock returns and market cash-flow shocks are determined by the cash-flow fundamentals of growth and value stocks.

While the stated intention of the previous research is to provide a direct, empirical investigation of the fundamental-based view against the sentiment-based view, the research simply provides an alternative formulation for measuring the cash-flow fundamentals of growth stocks and value stocks, neglecting to provide an empirical methodology for quantifying the sentiment-based view. Similar to the other studies discussed in this section, support (rejection) of the sentiment-based view is solely based on a rejection (support) for the fundamental-based view. Therefore, the focus has not been on identifying an effective methodology for quantifying the sentiment-based view, but rather to find unique ways to compare cash-flow fundamentals between growth and value stocks and that of the aggregate market.

The underlying intuition behind the sentiment-based view is that there is a separate, time-varying discount rate that is applied to the cash flow, not to be confused with the market discount rate. While most studies attempting to quantify the risk premium of discount-rate news and cash-flow news by regressing the market return on variables that are presumed to measure these shocks, we take an alternative approach to measure the risk premia associated with discount-rate shocks and cash-flow shocks.

By using sentiment on news articles pertaining to these shocks, we are able to identify the stocks in a given period that experience negative and positive news pertaining to discount rates and cash flows. By capturing the difference in the returns of stocks with negative and positive news for each shock, we can capture how the market prices these risks, allowing us to identify the time-variation in these premia. Thus, this approach attempts to contribute to the literature regarding the sentiment-based view, not through a rejection of the fundamental-based view, but rather through a novel approach in measuring shocks to these variables, which are inherently disseminated through news. If the sentiment-based view is to hold, we will be able to identify that the value premium is positively related to cash-flow shocks and negatively related to discount-rate shocks.

3. Methodology

In this section, we discuss the methodology used to capture the risk premium associated with cash-flow shocks and discount-rate shocks. That is, we aim to capture how the market prices the cross-section of stocks experiencing positive news about future cash-flow shocks and negative news about future cash-flow shocks. Similarly for discount rates, we aim to capture how the market prices the cross-section of stocks that have experienced positive news about future discount rates and negative news about future discount rates. For both sources of risk, we formulate long-short portfolios to capture the premium associated with each.

The first step in formulating the long-short portfolios is to isolate news articles that pertain to each source of risk. For cash-flow shocks, we isolate the set of news articles that pertain to forecasts or opinions regarding a company's future earnings or dividends, which is made convenient by the news provider dataset used (discussed in the next section). For discount-rate shocks, we isolate the set of news articles pertaining to stock prices. We use news about stock prices to characterize discount-rate shocks because sharp changes in prices (i.e., price changes that warrant news coverage) reflect changes in how the stock will be discounted in the future.

Once isolated to this set of news articles, we use the sentiment associated with each news article as our sorting variable. Because we rebalance portfolios monthly, we take the average sentiment score of the news articles pertaining to a given stock in the previous month. Based on the average monthly sentiment score of each category of risk (cash flow and discount rate), we formulate two equally-weighted portfolio of the stocks in the highest decile and quintile and the stocks in the lowest decile and quintile of monthly sentiment scores. The risk premium is then captured by taking the difference in the return of the portfolio with the highest sentiment scores and the return on the portfolio with the lowest sentiment scores. Again, this is done separately for the cash-flow portfolios and discount-rate portfolios.

To understand the relationship between cash-flow shocks and discount-rate shocks with the value premium, we estimate two regressions. The first regression attempts to understand how the value factor loads on the cash-flow shock risk premium and the discount-rate shock risk premium. If we let R_{hml} denote the return on the value premium, R_{cs} denote the return of the long-short portfolio formulated on the sentiment pertaining to news about cash-flow shocks, and R_{ps} denote the return on the long-short portfolio formulated on news about discount-rate shocks, the regression can be written as follows:

$$R_{hml,t} = \alpha + \beta_{cs} R_{cs,t} + \beta_{ps} R_{ps,t} + \epsilon_t \tag{1}$$

The regression given by Eq. 1 will help to understand the value premium's loadings on news about cash-flow shocks and news about discount rates. Again, we expect that the value portfolio will have a positive loading on the cash sentiment portfolio and a negative loading on the price sentiment portfolio.

The following two regressions are

$$R_{cs,t} = \alpha + \beta_{mkt} R_{mkt,t} + \beta_{smb} R_{smb,t} + \beta_{hml} R_{hml,t} + \epsilon_t \tag{2}$$

$$R_{ps,t} = \alpha + \beta_{mkt} R_{mkt,t} + \beta_{smb} R_{smb,t} + \beta_{hml} R_{hml,t} + \epsilon_t \tag{3}$$

Eqs. 2 and 3 serve to understand the relationship between the cash and price sentiment portfolios and the Fama–French three factor model to determine whether these shocks influence other systematic risk factors, as well as determine whether the relationship between value and the sentiment portfolios stand when accounting for the other well-known risk factors.

4. Data

In this section, we discuss the data sources used for this work. We gather news sentiment data from RavenPack, individual equity returns from Bloomberg, and factor returns from the Kenneth French Data Library.

RavenPack is one of the leading data analytics providers for financial services, designed to cover all major news sources from around the world, in multiple languages, and across different sectors. It provides a wide range of data and analytics services, including sentiment analysis, entity extraction, topic classification, and event detection. These services can be used to analyze news articles, social media posts, press releases, and other forms of digital content to provide actionable insights that can help organizations make better investment decisions, identify emerging trends and risks, and improve their overall performance.

Our dataset of cash-flow news consists of over six million news articles from 7,141 distinct sources (e.g., Dow Jones Newswires, Wall Street Journal, CNN, etc.). The news data consist of news articles that pertain to the 500 largest U.S. equities, which is a predefined universe provided by Ravenpack. RavenPack also provides further classifications of news items within each group, which we will refer to as "news sub-categories". For example, cash-flow news related to earnings and dividends has 39 sub-categories, such as earnings expectations, earnings per share, dividend guidance, etc. Our price-news database has more than five million news articles from 8,039 sources with no further sub-categories.

In order to validate our results and conduct robustness tests, we use a number of additional datasets. We use the US Economic Policy Uncertainty Index (EPU)[1] created by Baker et al. (2016) as a measure of the level of uncertainty in the US economy based on various sources of information, including newspaper articles, expert surveys, and legal disputes. The index is intended to capture the degree of policy-related uncertainty faced by businesses and consumers and has been used in various academic studies to analyze the effects of uncertainty on economic outcomes such as investing, spending, and hiring. Equity Market

[1] Downloaded from http://www.policyuncertainty.com/

Volatility Index (EMV)[2] developed by Baker et al. (2019) is another dataset we use in our robustness tests. This index uses articles about stock market volatility in leading US newspapers to construct a measure of volatility in equity returns.

We use the VIX[3] index as a control variable for market-based implied volatility. This index, created by the Chicago Board Options Exchange (CBOE) using a combination of options prices on the S&P 500 Index, reflects market participants' expectations for the stock market's volatility over the next 30-day period. We also use the Financial Uncertainty Index (FUI)[4] developed by Ludvigson et al. (2021), which is an index that measures the level of uncertainty in the financial markets. It is created using high-frequency financial data such as stock prices, interest rates, and exchange rates. The FUI is constructed using a factor-augmented vector autoregression (FAVAR) model, which incorporates information from multiple financial and macroeconomic variables.

The last dataset we use for our robustness checks is the University of Michigan Consumer Sentiment Index (MCSI)[5]. This index reflects overall consumer sentiment based on answers to questions about the financial situation, expectations, and opinions on the economy and buying conditions. Although MCSI is not directly used for evaluating the value premium, both consumer sentiment and spending play a role in determining the performance of the stock market and affecting the relative performance of value and growth stocks. Higher consumer confidence and spending can drive up economic activity and growth stock valuations, while a decline can slow down the economy and impact the stock market, including value stocks.

The correlation matrix in Table 1 demonstrates that there is only a limited correlation between the sentiment measures that we use and the other variables used for robustness tests, suggesting the distinctive information content of cash sentiment and price sentiment.

5. Results

In this section, we report summary statistics and figures regarding the cash and price sentiment portfolios in relation to the value factor as well as the other common systematic risk factors. In particular, we pay attention to how value loads on the price and cash sentiment. Similarly, we focus on how price and cash sentiment portfolios load on Fama–French factor returns. It is important to note that the purpose of this paper is not to argue for the cash and price sentiment portfolios as alpha-generating strategies. Instead, the purpose of constructing these portfolios is to quantify the risk premium associated with cash-flow risk and discount-rate risk.

We present the cumulative returns of the sentiment portfolios along with the cumulative returns of the value factor in Fig. 1. The left side of Fig. 1 shows the sentiment portfolios formulated using quintile sorts and the right side shows the sentiment portfolios using decile sorts. The cumulative returns from the value factor (HML) are the same on both sides, as again, these returns are taken from the Kenneth French Data Library. The quintile sorts show the relationship that is to be expected if the sentiment-based view of the value premium is to hold. That is, the returns on the price sentiment portfolio (proxying for news about future discount rates) is almost the inverse of the value factor returns and the cash sentiment portfolio is more positively related to the returns of the value factor. The results are less clear for the decile sorts, but a similar case could be made about these portfolios as well.

The cumulative returns plots help to visualize the relationship between the value factor and risk premia associated with news about cash-flow shocks and news about discount-rate shocks. Table 2 shows the correlation between the sentiment portfolios and the three other risk factors in the Fama–French three-factor model. The first column shows the correlation between the value factor and the other two Fama–French factors. We include this column to show that the value factor is negatively correlated with the market returns as well as the size factor. As a result, we would expect the cash sentiment portfolios to show a similar relationship and the price sentiment portfolio to show the opposite relationship.

The results from the quintile sorts portfolios are more in line with our expectations. The return on the cash sentiment portfolio is also negatively correlated with the market return and the return on the size factor, similar to the relationships noted with the value factor. Additionally, we see that the cash portfolio has a positive relationship with the value factor itself. The price sentiment portfolio, as expected, shows the opposite relationship. The price sentiment portfolio has a strong, positive relationship with the market return, a positive relationship with the size portfolio, and a negative relationship with the value factor itself. Thus, the results suggest that our proxies for news about future cash flows and news about future discount rates are able to identify the positive relationship between the cash-flow news and the value factor and the negative relationship between discount-rate news and the value factor.

To formalize this relationship and to identify the contribution of cash-flow news and discount-rate news within the value premium, we estimate the regression given by Eq. 1. The results of this regression are presented in Table 3 with the first column showing the regression results when using the quintile sorting methodology for the sentiment portfolios and the second column showing the results when regressing on the decile sort methodology. The results using the quintile sorts indicate that there is a significant, positive relationship between the value factor and the cash sentiment portfolio and a significant, negative relationship for the price sentiment portfolio.

---

[2] Downloaded from http://www.policyuncertainty.com/

[3] Downloaded from https://www.cboe.com/tradable_products/vix/

[4] Downloaded from https://www.sydneyludvigson.com/uncertainty-indexes

[5] Downloaded from https://data.sca.isr.umich.edu/data-archive/mine.php

Table 1
Correlation matrix of news sentiment data and other variables used for robustness tests.

	Cash Sentiment	Price Sentiment	EMV	EPU	VIX	FIU	MCSI
Cash Sentiment	1.00	-	-	-	-	-	-
Price Sentiment	0.18	1.00	-	-	-	-	-
EMV	-0.01	-0.10	1.00	-	-	-	-
EPU	-0.08	0.01	0.50	1.00	-	-	-
VIX	0.05	-0.07	0.72	0.45	1.00	-	-
FIU	0.10	-0.04	0.61	0.47	0.81	1.00	-
MCSI	0.07	-0.10	-0.31	-0.25	-0.45	-0.30	1.00

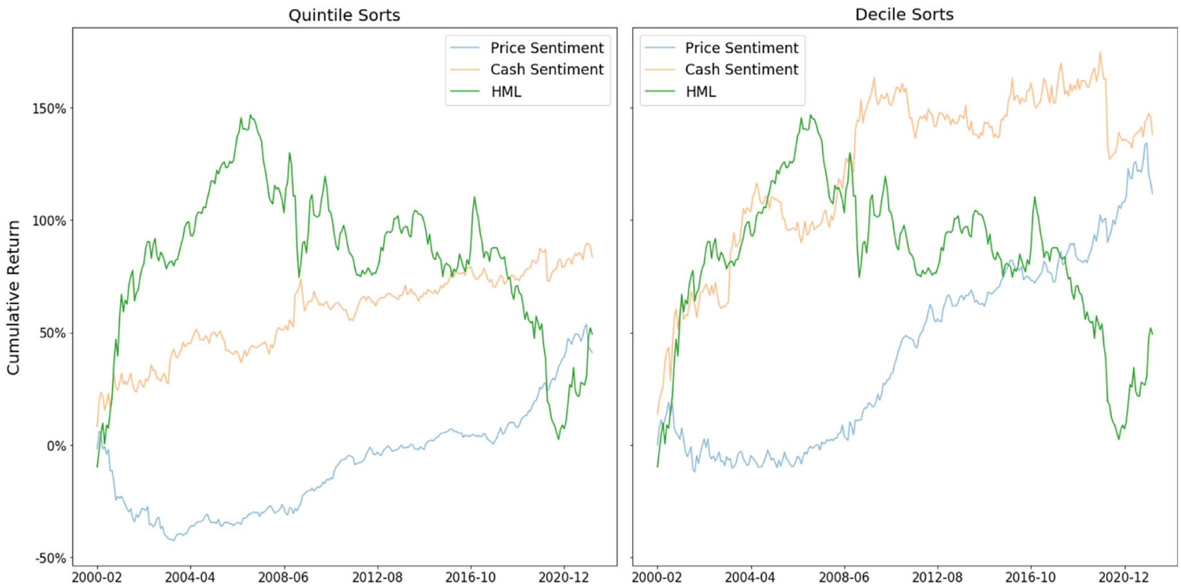

Fig. 1. Returns on value and sentiment portfolios.

Table 2
Return correlations.

	Quintile Sorts			Decile Sorts	
	R_{hml}	R_{cs}	R_{ps}	R_{cs}	R_{ps}
R_{mkt}	−2.56%	−7.25%	31.24%	−2.83%	28.62%
R_{smb}	−1.88%	−1.24%	3.64%	12.41%	3.56%
R_{hml}	-	16.11%	−15.93%	13.38%	−2.59%
R_{cs}	-	-	18.56%	-	20.45%
R_{ps}	-	18.56%	-	20.45%	-

Table 3
Value regressions on cash and price portfolios.

	$R_{hml} = \alpha + \beta_{cs}R_{cs} + \beta_{ps}R_{ps} + \epsilon$ Quintile Sorts	Decile Sorts
α	0.0017	0.0016
	(0.0021)	(0.0021)
β_{cs}	0.3414**	0.1890*
	(0.1050)	(0.0812)
β_{ps}	−0.3054**	−0.0777
	(0.0947)	(0.0871)
$R - Squared$	6.3%	2.08%

Standard errors are shown in parentheses.
* Statistically significant at the 5% level.
** Statistically significant at the 1% level.

We notice the same relationship between the value returns and the sentiment portfolios using decile sorts. However, the coefficient on the price sentiment portfolio is statistically insignificant and has a much lower magnitude for the relationship with value. The loading on the cash sentiment portfolio is only significant at the 5% level, compared to the cash-sentiment portfolio using quintile sorts being significant at the 1% level.

Table 4 presents the results of regressing the sentiment portfolios on the three Fama–French factor returns. The first two columns show the results with the quintile sorting methodology, and the following two columns show the results with the decile sorting methodology. The first column shows results of regressing the cash-flow news portfolio on the Fama–French factor returns. The results suggest that the cash-flow news portfolio has a significant positive loading on the value factor and insignificant loadings on the size and market returns. The intercept (representing portfolio alpha) is positive and significant; however, as stated previously, we are not arguing for the cash-flow portfolio as an alpha-generating strategy. The results are similar for the cash-flow portfolio generated using decile sorts, but there is a significant and positive loading on the size factor as well.

The price-sentiment portfolio regressions are shown in the second and fourth columns. For the quintile sort portfolio, we notice a negative and statistically significant loading on the value factor, as would be expected. However, the decile sort portfolio, while still having a negative relationship with value, is not significant. For both sorting methodologies, we notice that the price-sentiment portfolios have a similar magnitude in the positive loading on the market returns. This is an interesting result, but not out of line with expectations. Because the price-sentiment portfolio is constructed to measure shocks to the discount rate, it is not surprising that the return on this portfolio has a positive and significant loading on the market.

To gain a more comprehensive understanding of the relationship between sentiment and the value premium, and increase the validity of the results, we conduct robustness checks with variables such as EPU, EMV, VIX, FUI, and MCSI. These variables are relevant for performing robustness checks for the value premium because they capture different aspects of economic sentiment and financial uncertainty. EPU considers various sources such as the frequency of policy-related newspaper articles, expert surveys, and legal disputes to determine the level of uncertainty related to policy changes that businesses and consumers face. EMV calculates stock market volatility based on articles in leading US newspapers. The VIX index is also included as a proxy for implied volatility in the market and is calculated using option prices on the S&P 500 Index. Lastly, we utilize FUI, which measures the level of uncertainty in the financial markets using high-frequency financial data like stock prices, interest rates, and exchange rates.

In Panel A of Table 5, we present robustness results for cash and price sentiment portfolios using quintile sorting and other selected control variables, and Panel B depicts results based on decile sorting. From the results in Panel A, it can be seen that our suggested news sentiment indexes perform better than the other control variables, with an R-squared value of 6.3%. This suggests that the sentiment indexes are more effective in explaining the variation in the value premium compared to the other control variables. However, the results in Panel B show a slight decrease in the explaining power of the cash and price decile portfolios compared to the EPU.

To identify the primary factors affecting the value premium, we use the random forest algorithm as a machine learning method to assess the significance of variables in our robustness tests, excluding the MCSI[6]. The random forest method is an efficient way to rank variables as it considers their interactions and can reveal intricate non-linear relationships. By using this method, we can identify the most important predictors and eliminate those that are less significant, streamlining the analysis and reducing computational demands. This results in a more focused and straightforward examination of the data, providing clearer insight into the factors contributing to the value premium.

The findings presented in Table 6 illustrate the relative importance scores of the predictors. The predictor that had the greatest impact was assigned a score of 100, and the scores of the remaining predictors were adjusted accordingly. The results emphasize that our news sentiment indexes are the most valuable predictors in terms of predicting the value premium compared to other control variables like EPU, FIU, VIX, and EMV.

6. Discussion

The results in the previous section are consistent with the findings in the literature that value stocks are more sensitive to news about future cash flows and growth stocks are more sensitive to news about future discount rates (Lettau and Wachter (2007);Campbell and Vuolteenaho (2004) and Campbell et al. (2010)). Since we find that the value factor, which is long value stocks and short growth stocks, has a positive loading on the cash-sentiment portfolio and a negative loading on the price-sentiment portfolio, our results contribute to the literature regarding the drivers of the value premium.

Although the drivers of the value premium have been well-established in the literature, there is still an ongoing debate as to whether the co-movement in returns of value stocks with other value stocks and growth stocks with other growth stocks is due to similarities in cash-flow fundamentals (i.e., fundamental-based view) or due to a time-varying discount rate that investors apply to cash flows (i.e., the sentiment-based view).

While the results in this study seem to suggest that the sentiment-based view holds, one must be careful making this conclusion. Because we study the sentiment of the news pertaining to cash flows, we inherently capture the discount rate that investors apply to cash flows as well as future cash-flow fundamentals. That is, in the unlikely case that the sentiment

[6] This is because the results in Table 5 indicate that MCSI has no predictive power for the value premium.

Table 4
Sentiment return loadings on Fama–French Factors.

	Quintile Sorts		Decile Sorts	
	R_{cp}	R_{ps}	R_{cp}	R_{ps}
α	0.0024*	0.0009	0.0033*	0.0023
	(2.0847)	(0.7563)	(2.1479)	(1.6360)
β_{mkt}	−0.0350	0.1619**	−0.0426	0.1629**
	(-1.2017)	(5.5623)	(-1.1121)	(4.9459)
β_{smb}	0.0253	−0.0598	0.1427**	−0.0407
	(0.5743)	(-1.4790)	(2.6822)	(-0.9299)
β_{hml}	0.0965**	−0.1074**	0.1266**	−0.0203
	(2.6390)	(-2.8798)	(2.6312)	(-0.5327)

The t-values are shown in parentheses.
* Statistically significant at the 5% level.
** Statistically significant at the 1% level.

Table 5
R-squared results of robustness checks for control variables.

Panel A: Quintile Sorts							
α	0.0017	0.0018	0.0136	0.0104	0.0107	−0.0172	0.0144
	(0.0021)	(0.0003)	(0.0048)	(0.0056)	(0.0103)	(0.0149)	(0.0212)
β_{cs}	0.3414**	-	-	-	-	-	0.3210**
	(0.1050)						(0.1077)
β_{ps}	−0.3054**	-	-	-	-	-	−0.3225**
	(0.0947)						(0.0966)
β_{EMV}	-	−0.0006*	-	-	-	-	−0.0003
		(0.0003)					(0.0005)
β_{EPU}	-	-	−0.0001*	-	-	-	-0.0001
			(0)				(0)
β_{VIX}	-	-	-	−0.0004	-	-	-0.0005
				(0.0003)			(0.0005)
β_{FIU}	-	-	-	-	−0.0092	-	0.0185
					(0.0108)		(0.0187)
β_{MCSI}	-	-	-	-	-	0.0002	−0.0001
						(0.0002)	(0.0002)
$R-Squared$	6.3%	1.3%	2.6%	0.9%	0.2%	0%	9.2%
Panel B: Decile Sorts							
α	0.0016	0.0113	0.0136*	0.0104*	0.0107	-0.0172	0.0041
	(0.0021)	(0.0052)	(0.0048)	(0.0056)	(0.0103)	(0.0149)	(0.0216)
β_{cs}	0.1890*	-	-	-	-	-	0.1552*
	(0.0812)						(0.0841)
β_{ps}	−0.0777	-	-	-	-	-	−0.0951
	(0.0871)						(0.0883)
β_{EMV}	-	−0.0006*	-	-	-	-	−0.0002
		(0.0003)					(0.0005)
β_{EPU}	-	-	−0.0001*	-	-	-	-0.0001*
			(0)				(0)
β_{VIX}	-	-	-	-0.0004	-	-	-0.0004
				(0.0003)			(0.0005)
β_{FIU}	-	-	-	-	−0.0092	-	0.0188
					(0.0108)		(0.0193)
β_{MCSI}	-	-	-	-	-	0.0002	0
						(0.0002)	(0.0002)
$R-Squared$	2.08%	1.3%	2.6%	0.9%	0.2%	0%	4.8%

Standard errors are shown in parentheses.
* Statistically significant at the 5% level.
** Statistically significant at the 1% level.

variables used in this study to proxy cash-flow shocks perfectly correlates with future cash-flow fundamentals, then it would be fair to make the conclusion that the fundamental-based view holds.

Conversely, if the sentiment variables used are not good indicators of future cash flows, then we are picking up on the discount rate that investors are applying to *perceived* cash-flow risks. That is, given positive news coverage regarding cash flows and negative news coverage regarding cash flows, how does the market price this difference in perceived risk? Thus, if this were the case, the results would strongly support the sentiment-based view.

Table 6
Ranking importance of sentiment portfolios and control variables.

	Quintile Sorts Importance	Decile Sorts Importance
Price	100	100
Cash	77	65
VIX	70	43
FIU	38	45
EPU	3	7
EMV	0	0

In all likelihood, the sentiment-based proxies are likely to pick up on the market's perception of a stock's cash-flow risks while also tracking the future cash flows. In this case, the approach used in this study provides a blend between the sentiment-based view and the fundamental-based view. Therefore, both views are valid so long as the market's perception of cash-flow risks actually correlate with subsequent cash flows, which would explain the excess volatility of the short-term assets documented by van Binsbergen et al. (2012).

Future work can aim to investigate whether this correlation between market perception of a stock's cash-flow risks and future cash flows holds in other markets, and for different asset classes since it may not always exist and various factors such as market sentiment, investor behavior, and economic conditions can impact the excess volatility of short-term assets. Additionally, further studies can explore the potential for incorporating sentiment-based proxies in investment strategies and portfolio construction. For instance, investors may consider adjusting their portfolios depending on the market sentiment—investing in value stocks when the sentiment is negative and switching to growth stocks when the sentiment is positive. This type of active investment strategy can allow investors to benefit from market inefficiencies created by shifts in sentiment.

Our paper introduces a new type of value premium's driver that is closely related to value investing. As Fama and French (2021) show, book-to-market ratios as an imperfect measure of fundamental-to-market ratios, our suggested news sentiment measures can help asset pricing models in the new corporate markets over the rapid changes in recent years.

7. Conclusion

This paper takes a novel approach in quantifying the sentiment-based view of the value premium. While most studies justify or reject the sentiment-based view by either rejecting or confirming the fundamental-based view, the approach in this paper is able justify the sentiment-based view without the need for extracting cash-flow fundamentals of value and growth firms. This is done by using sentiment from news articles pertaining to cash-flow and discount-rate risks, separately, in order to understand how the market prices these two sources of risk. The benefit of this approach is that news articles are where "news" about future cash flows and discount rates are disseminated, as well as the tonality of the news articles being able to isolate whether these shocks are out of line with the market's expectations.

We find that the value factor has significant and positive loadings on news about future cash flows and a significant and negative relationship with news about discount-rate shocks. We also find that the cash-flow risk premium has a positive loading on the value factor when regressed on the Fama–French factors and does not have any significant loadings on the other factors. In terms of shocks to the discount rate, we notice a significant, positive loading on the market factor and a significant, negative loading on the value factor.

These results are in line with the literature that value stocks are more sensitive to news about future cash flows and growth stocks are more sensitive to news about future discount rates. However, these findings are not the result of measuring the cash-flow fundamentals of value and growth stocks, which is where most of the support, or lack thereof, for the value premium comes from, but rather from a direct analysis of the news sentiment pertaining to these shocks. Thus, these results shed new light on the underlying drivers of the value premium, suggesting a sentiment-based view. That is, the value premium is driven by a time-varying discount rate applied to cash flows in addition to shocks to the discount rate, which are not necessarily the result of similarities in the cash-flow fundamentals of value and growth stocks.

There are a few limitations of this analysis that are left for future research. First, the criticism of other papers attempting to distinguish between the fundamental- and sentiment-based view is that they default to accepting or rejecting either view by rejecting or accepting the fundamental-based view. We do the opposite in this paper. We accept the sentiment-based view but do not provide any specific analysis to accept or reject the fundamental-based view. However, this does not mean that we are rejecting the fundamental-based view because it is possible that our proxy variable picks up on the realized fundamentals, which can be explored in future iterations.

The next criticism is that we use sentiment to proxy for shocks to the discount rate and cash flows, which is subject to estimation errors. As compared to other sentiment-based approaches in finance, we isolate the news articles pertaining to specific events, which have tonalities that are much less ambiguous than the entire scope of news. For example, "Apple's earnings are expected to drop this quarter" is much easier to quantify tonality than, "Apple is set to release the new iPhone." Nonetheless, it is important to better understand how sentiment correlates with subsequent cash flows and verify whether sentiment is a good proxy for these risks.

Thus, future research in this area must verify the integrity of the sentiment variables used. More importantly, future research, in describing the value premium, should focus on identifying methodologies for exploring the sentiment-based view, or more specifically, whether there exists a separate and time-varying discount rate that is applied cash flows in addition to the market discount rate.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Algaba, A. et al, 2020. Econometrics meets sentiment: an overview of methodology and applications. J. Econ. Surveys 34 (3), 512–547.
Baker, M., Wurgler, J., 2007. Investor sentiment in the stock market. J. Econ. Perspect. 21 (2), 129–152. https://doi.org/10.1257/jep.21.2.129.
Baker, S.B. et al., 2019. Policy news and equity market volatility. NBER working paper 25720.
Baker, S.R., Bloom, N., Davis, S.J., 2016. Measuring economic policy uncertainty. Quart. J. Econ. 131 (4), 1593–1636.
Bansal, R., Dittmar, R.F., Lundblad, C.T., 2005. Consumption, dividends, and the cross section of equity returns. J. Finance 60 (4), 1639–1672. https://doi.org/10.1111/j.1540-6261.2005.00776.x.
Barber, B.M., Odean, T., 2007. All that glitters: the effect of attention and news on the buying behavior of individual and institutional investors. Rev. Finan. Stud. 21 (2), 785–818. https://doi.org/10.1093/rfs/hhm079.
Barberis, N., Shleifer, A., 2003. Style investing. J. Financ. Econ. 68 (2), 161–199. https://doi.org/10.1016/S0304-405X(03)00064-3.
Barberis, N., Shleifer, A., Wurgler, J., 2005. Comovement. J. Financ. Econ. 75 (2), 283–317. https://doi.org/10.1016/j.jfineco.2004.04.003.
Beccalli, E., Doninelli, N., Orsini, C., 2022. Value premium and macroeconomic variables. Eur. Finan. Manage., Forthcoming.
Brown, G.W., Cliff, M.T., 2004. Investor sentiment and the near-term stock market. J. Empirical Finance 11 (1), 1–27. https://doi.org/10.1016/j.jempfin.2002.12.001.
Campbell, J.Y., Polk, C., Vuolteenaho, T., 2010. Growth or glamour? Fundamentals and systematic risk in stock returns. Rev. Finan. Stud. 23 (1), 305–344. https://doi.org/10.1093/rfs/hhp029.
Campbell, J.Y., Vuolteenaho, T., 2004. Bad beta, good beta. Am. Econ. Rev. 94 (5), 1249–1275. https://doi.org/10.1257/0002828043052240.
Cao, V.N., 2015. What explains the value premium? The case of adjustment costs, operating leverage and financial leverage. J. Bank. Finance 59, 350–366.
Cornell, B., 1999. Risk, duration, and capital budgeting: New evidence on some old questions. J. Bus. 72 (2), 183–200. https://doi.org/10.1086/209609.
Donangelo, A., 2021. Untangling the value premium with labor shares. Rev. Finan. Stud. 34 (1), 451–508.
Fama, E.F., French, K.R., 2021. The value premium. Rev. Asset Pricing Stud. 11 (1), 105–121.
Frankel, R., Jennings, J., Lee, J., 2022. Disclosure sentiment: machine learning vs. Dictionary methods. Manage. Sci. 68 (7), 5514–5532. https://doi.org/10.1287/mnsc.2021.4156.
Gentzkow, M., Kelly, B., Taddy, M., 2019. Text as data. J. Econ. Literat. 57 (3), 535–574.
Gerakos, J., Linnainmaa, J.T., 2018. Decomposing value. Rev. Finan. Stud. 31 (5), 1825–1854.
Gonçalves, A.S., Leonard, G., 2023. The fundamental-to-market ratio and the value premium decline. J. Financ. Econ. 147 (2), 382–405.
Hansen, L.P., Heaton, J.C., Li, N., 2008. Consumption strikes back? Measuring long-run risk. J. Polit. Econ. 116 (2), 260–302. https://doi.org/10.1086/588200.
He, Y., Leippold, M., 2020. Short-run risk, business cycle, and the value premium. J. Econ. Dynam. Control 120, 103993.
Hirshleifer, D., Sheng, J., 2022. Macro news and micro news: complements or substitutes? J. Financ. Econ. 145 (3), 1006–1024.
Hochberg, Y.V. et al, 2022. Lead inventor gender and patent citations: evidence from causal BERT Estimation. SSRN Electron. J. https://doi.org/10.2139/ssrn.4269703.
Jakubik, J. et al, 2023. Incorporating financial news for forecasting Bitcoin prices based on long short-term memory networks. Quant. Finance 23 (2), 335–349.
Jeon, Y., McCurdy, T.H., Zhao, X., 2022. News as sources of jumps in stock returns: Evidence from 21 million news articles for 9000 companies. J. Financ. Econ. 145 (2), 1–17.
Ke, Z.T., Kelly, B., Xiu, D., 2019. Predicting Returns With Text Data. https://doi.org/10.3386/w26186.
Klemola, A., 2020. Internet search-based investor sentiment and value premium. Finance Res. Lett. 33, 101224.
Lakonishok, J., Shleifer, A., Vishny, R.W., 1994. Contrarian investment, extrapolation, and risk. J. Finance 49 (5), 1541–1578. https://doi.org/10.2307/2329262.
Lemmon, M., Portniaguina, E., 2006. Consumer confidence and asset prices: some empirical evidence. Rev. Finan. Stud. 19 (4), 1499–1529.
Lettau, M., Wachter, J.A., 2007. Why is long-horizon equity less risky? A duration-based explanation of the value premium. J. Finance 62 (1), 55–92. https://doi.org/10.1111/j.1540-6261.2007.01201.x.
Ludvigson, S.C., Ma, S., Ng, S., 2021. Uncertainty and business cycles: exogenous impulse or endogenous response? Am. Econ. J.: Macroecon. 13 (4), 369–410.
Qiu, L.X., Welch, I., 2006. Investor Sentiment Measures. SSRN Electron. J. https://doi.org/10.2139/ssrn.589641.
van Binsbergen, J., Brandt, M., Koijen, R., 2012. On the timing and pricing of dividends. Am. Econ. Rev. 102 (4), 1596–1618. https://doi.org/10.1257/aer.20131416.