

T

The Predictive Power of Twitter Sentiment Index on U.S. Stock Returns

Konpanas Dumrongwong^{1,*}

Received: April 22, 2022 Revised: May 27, 2022 Accepted: August 9, 2022

Abstract

Using a novel Twitter-based investor sentiment index, this research investigates whether investor sentiment from social media, as expressed in daily Twitter messages, has predictive power with respect to stock returns. Based on hierarchical regressions, the empirical results show that the Twitter sentiment index has additional predictive power for U.S. stock returns, which is not captured by traditional factors, such as market risk premium, firm size, book-to-market ratio, or momentum. The results suggest that investor sentiments from social media significantly affect short-term equity value. Thus, individual investors and fund managers should be aware of the impact social media sentiment can have on both their own portfolios and fund managers' investment strategies.

Keywords: behavioral finance, twitter, investor sentiment, empirical asset pricing

¹ Department of Finance, Thammasat University

* Corresponding author. E-mail: konpanas@tbs.tu.ac.th



ความสำคัญของโพสต์ Twitter ที่มีต่อผลตอบแทนการลงทุนในตราสารทุน

กนต์พนัส ดำรงวงศ์^{1,*}

วันรับบทความ: April 22, 2022 วันแก้ไขบทความ: May 27, 2022 วันตอบรับบทความ: August 9, 2022

บทคัดย่อ

งานวิจัยฉบับนี้ มุ่งศึกษาผลกระทบของอารมณ์ความรู้สึกของนักลงทุนต่อราคาหุ้นในตลาดรอง โดยอาศัยดัชนีบ่งชี้ความสุข ซึ่งชี้วัดจากข้อความในทวิตเตอร์ (Twitter sentiment index) เป็นดัชนีบ่งชี้อารมณ์ความรู้สึกของนักลงทุน อาศัยระเบียบวิธี hierarchical regressions หลักฐานเชิงประจักษ์ในงานวิจัยบ่งชี้ว่าดัชนีชี้วัดความสุขจาก Twitter ส่งผลกระทบอย่างมีนัยสำคัญต่อราคาหุ้นที่ซื้อขายในตลาดรองในสหรัฐอเมริกา ทั้งนี้ ปัจจัยผลกระทบดังกล่าว ไม่อาจอธิบายได้โดยตัวแปรทางการเงินที่มีอยู่เดิมในแบบจำลองราคาหลักทรัพย์ที่เป็นที่รู้จักดีในปัจจุบัน (อาทิเช่น ค่าชดเชยความเสี่ยงตลาด (Market Risk Premium), อัตราผลตอบแทนของสินทรัพย์ปราศจากความเสี่ยง (Risk-Free Rate), อัตราส่วนราคาทางบัญชีต่อราคาตลาด (B/M Ratio) หรือเทรดดิ้งโมเมนตัม) ดังนั้น ผู้มีส่วนได้เสียในตลาดหุ้น อาทิเช่น นักลงทุน ผู้จัดการกองทุน ควรมีความรู้ความเข้าใจถึงผลกระทบของอารมณ์ความรู้สึกของนักลงทุนที่สะท้อนจากสื่อโซเชียลมีเดีย ซึ่งอาจส่งผลกระทบต่อผลตอบแทนและความเสี่ยงของพอร์ตลงทุน

คำสำคัญ: การเงินเชิงพฤติกรรม ทวิตเตอร์ อารมณ์ความรู้สึกของนักลงทุน แบบจำลองราคาหลักทรัพย์

¹ Department of Finance, Thammasat University

* Corresponding author. E-mail: konpanas@tbs.tu.ac.th

Introduction

Financial scholars have long debated the possible effects of investor sentiment on asset prices. Traditional theoretical asset pricing models, such as the capital asset pricing model (CAPM), are generally unreliable in explaining changes in real-world stock returns and pose many challenges in practical application. Many versions of empirical asset pricing models exist and some of them are arguably better at explaining realized equity returns. Good examples include the famous Fama-French three-factor model (Fama & French, 1993), the Carhart four-factor model (Carhart, 1997) and the Fama-French five-factor model (Fama & French, 2015). Although these models are more successful at explaining equity returns, it is still unclear what risk (or risks) the empirical factors truly account for in these models. For example, what are the actual risk factors of a book-to-market (B/M) ratio (i.e., the ratio of the book value of a common stock to its market value)? Similar arguments apply to other empirical factors, such as firm size or momentum. Therefore, it remains inconclusive whether these recognized factors are the only ones relevant to stock returns, or whether there are other unknown factor(s) with additional predictive power.

As the search for a better empirical model continues, academic focus has shifted away from relying on the rationality assumption and has begun investigating the relationships between asset prices and investor sentiment. Noise traders and psychological biases are primary subjects in analyses of the impact of investor sentiment on stock prices. For example, De Long, Shleifer, Summers and Waldmann (1990) have shown that irrational noise trader behavior cannot be offset by limited arbitrageurs, and that a wide spectrum of sentiment from diverse investors could affect stock prices and result in higher expected returns. However, since it is not possible to directly observe investor sentiment without an intrusive survey, many studies have relied on indirect proxies, such as closed-end fund discounts (Lee, Shleifer & Thaler, 1991), bid-ask spreads and turnover (Baker & Stein, 2004), consumer confidence (Lemmon & Portniaguina, 2006) or a combination of all of these (Baker & Wurgler, 2006). Although these traditional measures of investor sentiment provide some useful insights into the relationship between asset prices and investor sentiment, they have their disadvantages. For example, market-based proxies may be affected by many confounding variables, and survey-based proxies cannot guarantee the quality of the response.

Most traditional asset pricing models and the efficient-market hypothesis (EMH) rely on the implicit assumption that asset prices ultimately incorporate new information. However, most of these models, whether theoretical or empirical (such as CAPM or the Fama-French models), were developed well before the Internet era. Although social media is not a source of information,

it nevertheless provides information about investor sentiment according to prior research (e.g., Da, Engelberg & Gao, 2015; 2020; Siganos, Vagenas-Nanos & Verwijmeren, 2014; Zhang, Wang, Li & Shen, 2018; Zhao, 2020). Moreover, given that any kind of information can be easily spread over the Internet (i.e., propagation through social media channels), it is particularly interesting to investigate whether the information contained in social media is actually relevant to stock prices. Accordingly, the present research aims to evaluate the relevance of social media and fill this research gap.

This contribution offers new insights into the issue of investor rationality by applying new data and techniques, employing a direct online measure of investor sentiment rather than relying on indirect proxies. This measure, observed passively, allows for a direct estimate of investor sentiment while minimizing the response quality problem. This paper is original in two ways. First, this research employs a novel proxy for investor sentiment constructed from social media (Twitter), which has the advantage of avoiding endogeneity and directly capturing investor sentiment. Unlike survey-based proxies, social media sentiment index can *reveal* attitudes rather than *inquire* about them. Consequently, the measure is much less prone to biases than are survey-based measures of sentiment. Second, the empirical results presented in this study are consistent with recent empirical theories that an investor sentiment index constructed from social media has significant predictive power with respect to U.S. stock returns. This predictive power has not previously been explained in the finance literature by popular risk factors, such as market risk premium, firm size, B/M ratio, or momentum. The empirical results also indicate that the inclusion of an online-search-based sentiment index increases the R-squares of traditional asset pricing models.

The remainder of the paper is structured as follows. The next section presents a review of the literature regarding investor sentiment and its implications for stock returns. The subsequent section describes the research hypotheses, data, and methodology. The empirical results are then presented and discussed. Finally, conclusions are presented, along with research limitations and suggestions for future research.

Literature Review

Several theoretical studies offer various behavioral-based models establishing the link between asset prices and investor sentiment. For example, it has been documented that investors may form erroneous beliefs, through either excessive optimism or pessimism, and may therefore incorrectly evaluate asset values, causing asset price movements (e.g., Baker, Wurgler & Yuan, 2012;

Black, 1986; Daniel, Hirshleifer & Subrahmanyam, 1998; De Long et al., 1990). Behavioral finance theory consistently suggests that the presence of noise traders in the stock market, with associated behavior, as well as limits on arbitrage, are restricting conditions that can lead investor sentiment to influence asset prices (e.g., Baker & Wurgler, 2006; Lee et al. 1991; Shleifer & Summers, 1990; Shleifer & Vishny, 1997). De Long et al. (1990), notably, modelled the influence of noise trading on equilibrium prices and showed that noise trading affects stock prices and that noise traders can earn higher expected returns. Shleifer and Vishny (1997) demonstrated that there are limits on risky arbitrage positions which can cause changes in security prices. Furthermore, Baker and Wurgler (2006) documented that investor sentiment contains functional predictive content about stock returns.

Baker and Wurgler (2007) showed that investor sentiment predictive content in relation to future market movements can act as valuable information for traders in formulating profitable trading strategies. They broadly defined investor sentiment as “investors’ belief about future cash flows and risk not justified by the facts at hand” and noted that, “Now, the question is no longer, as it was a few decades ago, whether investor sentiment affects stock prices, but rather how to measure investor sentiment and quantify its effects” (p. 130). The present research employs these definitions and extends the analysis of the link between investor sentiment and stock returns, using new data and a new methodology for measuring sentiment.

A growing body of research has documented the relationship between investor sentiment and asset prices. For example, Baker et al. (2012) investigated stock prices in Canada, France, Germany, Japan, the United Kingdom, and the United States and found that sentiment is correlated with stock prices of listed companies in these major stock markets. Dergiades (2012) studied the U.S. stock indexes from 1965 to 2007 and showed that investor sentiment holds significant predictive power with respect to stock returns. In addition, Kaplanski, Levy, Veld and Veld-Merkoulova, (2015) used a survey to investigate sentiment among 900 investors and found that, on average, more positive investor sentiment is accompanied by higher return expectations and greater intentions to buy stocks. In their study, they also found that investor sentiment affects expected returns more intensely than does expected risk.

In this context, prior research has also revealed that online sources are known to contain information regarding investor sentiment, which is correlated with stock returns. For example, Siganos et al. (2014) examined investor sentiment as expressed in Facebook posts, and found that this sentiment has a positive contemporaneous association with stock returns, and that there

exists a one-way causality from online sentiment to stock returns. Kim and Kim (2014) investigated investor sentiment as measured by Yahoo! Finance message board postings, and found that investor sentiment has predictive power for stock returns. Da et al. (2015) documented a sentiment index constructed from Google search volumes and showed a correlation between sentiment as expressed in Google searches and asset prices. They found that negative words expressed in Google searches correspond with low market returns on the same day. Zhang et al. (2018) and Zhao (2020) found that the Twitter-based sentiment index Granger-causes stock returns in their causality tests. Finally, Naeem et al. (2020) documented that Twitter significantly causes the future volatility of their sample countries. These prior studies have provided an important foundation for the hypothesis in this research in two principal ways: first, they established the one-way causality from social-media sentiment to stock returns; and, second, their findings implied that investor sentiment constructed from Twitter should be a relevant factor for explaining stock returns.

Although prior evidence has established that social-media sentiment has a correlation with stock return movement, it is unclear if this evidence offers any additional explanatory power regarding stock returns over the known risk factors (for example, it is possible that the explanatory power of social-media sentiment is already captured by well-known factors documented in the finance literature, such as market risk premium, firm size, B/M ratio, or momentum). Prior studies (e.g., Da et al., 2015; Naeem et al. 2020; Siganos et al., 2014; Zhang et al., 2018) focused solely on the causality tests and correlation tests while the correlation between online investor sentiment and the aforementioned well-known risk factors was largely ignored. Therefore, the analysis of the research gap is the primary focus of the present paper. To the best knowledge of the author, this study is the first to study the explanatory power of investor sentiment from social media relative to the predictive power of common risk factors.

Because A) human sentiment such as mood or happiness, as expressed in Twitter messages, is unlikely to be explained by systematic factors in popular asset pricing models, such as market-risk premium, firm size, or Book-to-Market ratio, and because B) prior research has consistently established that a Twitter-based sentiment index Granger-causes stock returns (e.g., Zhang et al., 2018), intuitively, the logical way to express a testable hypothesis in order to reconcile these two observations is that Twitter-based sentiment index should have predictive power with respect to stock returns in addition to those systematic factors. This is the main hypothesis investigated in the present paper. Unlike prior research (e.g., Baker & Stein, 2004; Baker & Wurgler, 2006) which employed indirect proxies for investor sentiment, this study applies

new techniques by utilizing a novel Twitter-based sentiment index. This measure provides unique advantages by allowing for a direct estimate of investor sentiment at high frequency, while minimizing the problem of response quality.

Because this research also involves the investigation of asset prices, it is particularly important to consider prior contributions in the field of empirical asset pricing. Notable among the empirical asset pricing models is research by Fama and French (1993) who proposed their Fama-French three-factor model. This model garnered much attention from academics, and several studies have extended the model by including additional factors. Notably, Carhart (1997) argued that momentum should also be an important factor for determining asset prices. Moreover, Fama and French (2015) extended their own three-factor model, adding profitability and investment as additional factors. It is important to note that although many other versions of empirical asset pricing models exist, this research primarily focuses on the Fama-French three-factor model and the Carhart four-factor model because they are among the most widely known models in academia.

Methodology

1. Data

The Twitter happiness index was constructed from Hedonometer.org, which is generated from Twitter's Decahose API feed database of over 50 million daily Twitter post observations. The daily index is formulated by scoring nearly 10,000 sentiment-related words found in the database. Each of these words is then scored on a nine-point scale of happiness: (1) sad ... to (9) happy, following the Dodds, Harris, Kloumann, Bliss and Danforth (2011) methodology. Because of data availability, the study period ranges from September 2008 to January 2021. All daily risk factors (market risk premium, risk-free rate, HML, and SMB measures) are observed from the French data library (French, 2021). All data sources accord with prior research, providing a reliable base for comparisons.

2. Methodology and Hypotheses

Prior studies have established that online sentiment Granger-causes stock returns and that there is a linear relationship between a Twitter-based investor sentiment index and stock returns (e.g., Siganos et al., 2014; Zhang et al., 2018; Zhang, Li, Shen & Teglio, 2016). Therefore, the present study utilizes a conventional linear framework, based on prior research, to investigate this relationship. Accordingly, Twitter-based investor sentiment was treated as an independent

variable in conventional linear regression, again consistent with prior research. Because Da et al. (2015) showed that daily negativity (or positivity) in online posts corresponds to low (high) market-level returns on the same day, the effect of Twitter-based sentiment toward stock returns is therefore assumed to be observable on the same day. The Standard & Poor's 500 Index (S&P 500) and the Dow Jones Industrial Average Index (DJIA) were selected to represent the U.S. stock market because the former is one of the most commonly followed indexes, while the latter is the oldest U.S. stock index. Additional tests using Gibbons, Ross and Shanken's methodologies (henceforth GRS) (Gibbons, Ross & Shanken, 1989) on U.S. portfolios (2x3 and 5x5 formed on size and B/M ratios) were also conducted and are reported in the robustness checks section.

As discussed previously, the main focus of investigation of this study was the predictive power of Twitter-based sentiment, in addition to the well-known factors documented in the finance literature. This involved testing for the significance of the coefficient (β) of the Twitter-based sentiment (DH_t)² in the following time-series models. It is important to note here that although GRS tests are commonly used for time-series regressions, the following equations were considered more relevant to the research questions, because they are capable of directly testing the relevance of DH_t as an additional variable, rather than indirectly testing the magnitude of alpha. Nevertheless, GRS tests on 2x3 and 5x5 portfolios are briefly discussed in the robustness checks section (see also Fama and French (2020) for an alternative method).

$$R_t + R_{ft} = \alpha_t + \beta_1 MRP_t + \beta_2 DH_t = \varepsilon_t \quad (1)$$

$$R_t + R_{ft} = \alpha_t + \beta_1 MRP_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_4 DH_t = \varepsilon_t \quad (2)$$

$$R_t + R_{ft} = \alpha_t + \beta_1 MRP_t + \beta_2 SMB_t + \beta_3 HML_t + \beta_3 UMD_t + \beta_4 DH_t = \varepsilon_t \quad (3)$$

where R_t represents the daily stock returns at the end of day t , DH_t represents the Twitter-based sentiment index on day t , MRP_t represents the market risk premium on day t , SMB_t represents the size premium (Small Minus Big) on day t , HML_t represents the value premium (High Minus Low) on day t , UMD_t represents the momentum factor on day t , R_{ft} represents the risk-free rate, ε_t represents the error terms and α_t represents the intercept of the regression.

² The explanatory variable used is DH_t rather than DH_{t-1} , in line with Da et al. (2015) who concluded that daily negativity (or positivity) in online posts corresponds to low (or high) market-level returns on the same day. Hence, the daily stock returns R_t , measured at the end of day t , are assumed to be adjusted for investor sentiment during the day.

Accordingly, the following two working hypotheses were formulated:

H1: Twitter-based sentiment contains additional predictive power with respect to stock returns, which is not explained by factors in the Fama-French three-factor model.

H2: Twitter-based sentiment contains additional predictive power with respect to stock returns, which is not explained by factors in the Carhart four-factor model.

The above hypotheses were employed because they involve investigating empirical factors from the Fama-French three-factor model and the Carhart four-factor model, which are among the most popular asset pricing models in the finance literature. Although none of these factors are related to CAPM, the main tests also included a test using CAPM to provide basic contextual material for the investigation. Kaplanski et al. (2015) noted that sentiment can drive stock return expectations. However, the expected returns at the time of writing Twitter messages cannot be precisely measured; therefore, there is no testable hypothesis for CAPM. Tests for newer empirical models, such as Fama-French's five-factor model (Fama & French, 2015) and GRS tests on additional portfolios were also conducted, and are reported in the robustness checks section³. Finally, since the hypotheses testing involves model comparisons, Hierarchical Regression (HR) is also conducted to investigate whether the Twitter-based sentiment add explanatory power to the models.

Results

All considered time series were tested for being stationary using Augmented Dickey–Fuller (Dickey & Fuller, 1979) and Phillips–Perron (Phillips & Perron, 1988) methodologies (henceforth, ADF and PP, respectively). The null hypothesis of a unit root for all considered series was rejected at 1% significance level. Table 1 presents the summary statistics for stock returns, independent variables, and the happiness sentiment index, to give an overview of the data.

³ The Fama-French three-factor model is preferred over its five-factor counterpart for two main reasons: First, the Fama-French five-factor model has yet to be proven as an improvement compared to the previous three-factor model. Numerous recently published articles discuss the validity of the additional two factors in the Fama-French five-factor model. Second, the Fama-French five-factor model generally performs poorly with regard to small stocks. This poses problems for robustness tests.

Table 1 Summary Statistics

	Mean	Median	SD	Skew.	Kurt.	ADF	PP
S&P 500	0.0003	0.0006	0.01	-0.49	5.81	-64.11***	-64.25***
DJIA	0.0003	0.0006	0.01	-0.51	7.30	-21.14***	-62.68***
MRP_t	0.047	0.085	1.34	-0.36	5.13	-63.17***	-63.38***
SMB_t	-0.002	0.000	0.63	0.25	7.95	-57.80***	-58.02***
HML_t	-0.022	-0.035	0.78	0.29	8.43	-55.70***	-55.93***
UMD_t	0.003	0.050	1.06	-0.65	7.77	-50.00***	-49.84***
DH_t	1.794	1.794	0.01	-0.46	5.58	-4.89***	-23.62***

MRP_t represents the market risk premium on day t ; SMB_t represents the size premium (Small Minus Big) on day t ; HML_t represents the value premium (High Minus Low) on day t ; UMD_t represents the momentum factor on day t ; DH_t represents the Twitter-based sentiment index on day t ; and *, **, *** represent statistical significance at the 1%, 5%, and 10% levels, respectively.

Because investor sentiment (such as mood or happiness, expressed in Twitter messages) is unlikely to be explained by systematic factors, such as B/M ratio, firm size, or momentum, it was intuitively anticipated that the correlation between the aforementioned empirical factors and the Twitter-based sentiment index (DH) would be zero or near zero. The results in Table 2 indeed confirm this expectation and show that the Pearson correlation between Twitter-based sentiment (DH) and other explanatory variables is very low: size premium (SMB) shows the highest correlation with DH at a marginal magnitude $r=0.03$. This evidence is broadly supportive of the main hypothesis because it shows that Twitter-based sentiment is almost uncorrelated with the known empirical factors used in popular empirical asset pricing models. Therefore, if the explanatory power of DH with respect to stock returns exists, it is not likely to be captured by any known risk factors within the CAPM, Fama-French three-factor or Carhart four-factor models.

In terms of orthogonality, the value premium (HML) and momentum (UMD) show relatively high correlation at $r = -0.62$. However, it was not necessary to exclude either of these factors from the main analysis because they are empirically identified as relevant factors in accordance with prior empirical research (e.g., Carhart, 1997; Fama & French, 1993). In addition, Equations 1 and 2 naturally exclude the momentum factor (UMD), already giving a clear view of the results without the UMD.

Table 2 Pearson Correlation Matrix

	MRP	SMB	HML	UMD	DH
MRP _t	1.00				
SMB _t	0.23	1.00			
HML _t	0.34	0.22	1.00		
UMD _t	-0.30	-0.21	-0.62	1.00	
DH _t	0.02	0.03	0.03	-0.02	1.00

MRP_t represents the market risk premium on day t; SMB_t represents the size premium (Small Minus Big) on day t; HML_t represents the value premium (High Minus Low) on day t; UMD_t represents the momentum factor on day t; and DH_t represents the Twitter-based sentiment index on day t.

The primary subject of investigation involved examining the “relevance” of the DH factor in Equations (1)–(3). The results are presented in Tables 3, 4 and 5, showing the explanatory power of the Twitter-based sentiment index, in addition to the risk factors stated in the CAPM and the Fama-French three-factor and Carhart four-factor models, respectively. The results from Hierarchical Regression (HR) are also reported in Panel B of each table.

Table 3 Regression Results: Additional Explanatory Power of Twitter-based Sentiment in CAPM

Panel A: OLS	S&P 500	DJIA		Panel B: Hierarchical Regression (Comparing model with DH _t vs model without DH _t)	S&P 500	DJIA
MRP _t	1.00*** (698.44)	0.93*** (237.79)				
DH _t	0.32* (1.65)	1.19** (2.23)		F Change (p-value)	5.83*** (0.01)	8.16*** (0.00)
Intercept	-0.59* (-1.69)	-2.15** (-2.24)				
Adj. R ²	0.99	0.95				
Durbin-Watson	2.12	1.92				
Breusch-Godfrey (F-stat,4 lags)	1.62	0.43				

MRP_t represents the market risk premium on day t ; DH_t represents the Twitter-based sentiment index on day t ; Intercept = the intercept of the regression; and *, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

Results presented in Table 3 reveal that investor sentiment, as expressed in daily Twitter messages, contains predictive power with respect to U.S. index returns. The coefficient of investor sentiment was found to be positive, and statistically significant at 5% and 10% levels of significance for DJIA and S&P 500, respectively. These results show that the Twitter happiness index contains relevant information about stock prices. The coefficient of the Twitter happiness index was found to be positive for both DJIA and S&P 500, suggesting that positivity in daily Twitter messages and stock returns on the same day are correlated. The findings are consistent with Da et al.'s (2015) conclusion that daily negativity (or positivity) in online posts corresponds to low (or high) market-level returns on the same day. The Durbin-Watson statistic was found to be very close to 2, indicating that no evidence of autocorrelation was detected in the cases considered. The F-statistics from serial correlation LM tests (aka the Breusch–Godfrey test) suggested no evidence of higher order serial correlation (for a maximum of 4 lags) in the sample. Finally, the results from Hierarchical Regression showed that the Twitter sentiment index (DH_t) increases the explanatory power of the model compared to the CAPM (F change p-value: 0.01).

Table 4 Additional Explanatory Power of Twitter-based Sentiment in the Fama-French Three-factor Model

Panel A: OLS	S&P 500	DJIA	Panel B: Hierarchical Regression (Comparing model with DH_t vs model without DH_t)	S&P 500	DJIA
MRP_t	1.00*** (952.76)	0.93*** (236.79)			
SMB_t	-0.13*** (-58.30)	-0.16*** (-20.29)	F Change (p-value)	20.60*** (0.00)	8.16*** (0.00)
HML_t	0.03*** (16.12)	0.07*** (10.96)			
DH_t	0.47*** (3.52)	1.31*** (2.62)			
Intercept	-0.86*** (-3.58)	-2.36*** (-2.63)			
Adj. R^2	0.99	0.96			
Durbin-Watson	1.95	1.91			
Breusch-Godfrey (F-stat,4 lags)	1.65	1.10			

MRP_t represents the market risk premium on day t ; SMB_t represents the size premium (Small Minus Big) on day t ; HML_t represents the value premium (High Minus Low) on day t ; DH_t represents the Twitter-based sentiment index on day t ; Intercept = the intercept of the regression; and *, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 4 shows the results of Fama-French regressions with Twitter sentiment as an augmented variable. It was found that the coefficients of Twitter sentiment (DH) were statistically significant at 1% for all considered cases (S&P 500 and DJIA). These results show that the Twitter happiness index contains relevant information about stock price movements. The coefficients of the Twitter happiness index reveal the following: for each increase by 1 index point, stock returns can be expected to increase by 0.47% and 1.31% for S&P 500 and DJIA, respectively. This finding suggests that investor sentiment can drive stock returns, in line with Kaplanski et al.'s (2015) prediction. The signs of the coefficients were found to be positive, suggesting that higher sentiment expressed in Twitter is associated with higher returns, and vice versa. This finding is consistent

with that of Da et al. (2015), who posited that daily negativity (or positivity) in online messages corresponds to low (high) market level returns on the same day.

The results for other factors accord with the findings of Fama and French (1993): market risk premium (MRP), size premium (SMB), and value premium (HML) were found to be significant predictors of stock returns (at 1% level of significance). Because Twitter sentiment (DH) is a statistically significant factor in describing stock returns, and because its correlation with the other explanatory variables is marginal (as shown in Table 2), the results presented in Table 4 indicate that Twitter sentiment (DH) is a predictor for stock returns which cannot be explained by the market risk premium, firm size, or B/M ratio, thus confirming *H1*. The Durbin-Watson statistics were found to be very close to 2, suggesting no evidence of autocorrelation in all considered models. The F-statistics from serial correlation LM tests (aka the Breusch–Godfrey test) suggest no evidence of higher order serial correlation (for a maximum of 4 lags) in the cases considered. Finally, the results from Hierarchical Regression indicate that the inclusion of the Twitter sentiment index (DHt) increases the explanatory power of the model compared to the Fama-French three-factor model (p-value: 0.00).

Table 5 Additional Explanatory Power of Twitter-based Sentiment in the Carhart Four-factor Model

Panel A: OLS	S&P 500	DJIA		Panel B: Hierarchical Regression (Comparing model with DH_t vs. model without DH_t)	S&P 500	DJIA
MRP_t	1.00*** (946.41)	0.94*** (236.55)				
SMB_t	-0.13*** (-58.12)	-0.16*** (-20.09)		F Change (p-value)	20.63*** (0.00)	8.11*** (0.00)
HML_t	0.03*** (13.13)	0.08*** (10.05)				
UMD_t	0.00 (-0.16)	0.01* (1.84)				
DH_t	0.47*** (3.52)	1.30*** (2.61)				
Intercept	-0.86*** (-3.58)	-2.35*** (-2.63)				
Adj. R^2	0.99	0.99				
Durbin-Watson	1.95	1.91				
Breusch-Godfrey (F-stat,4 lags)	1.72	1.11				

MRP_t represents the market risk premium on day t ; SMB_t represents the size premium (Small Minus Big) on day t ; HML_t represents the value premium (High Minus Low) on day t ; UMD_t represents the momentum factor on day t ; DH_t represents the Twitter-based sentiment index on day t ; Intercept = the intercept of the regression; and *, **, *** represent statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5 presents the results from the Carhart four-factor regression model, with Twitter sentiment as an augmented variable. Once again, it was found that the coefficients of Twitter sentiment (DH) were positive and statistically significant at 1% for all considered U.S. stock indexes. These results suggest that investor sentiment, as expressed in Twitter messages, has significant predictive power with respect to U.S. stock returns in addition to the four factors (market risk premium, firm size, value premium, and momentum) presented in the Carhart four-factor model,

thus confirming *H2*. Once again, the signs of the coefficients were found to be positive, supporting Da et al. (2015), who posited that daily negativity (or positivity) in online posts corresponds to low (high) market-level returns on the same day. These results show that the Twitter happiness index contains relevant information about stock prices. The coefficients of the Twitter happiness index reveal that each increase by 1 index point results in an increase in stock returns of 0.47% and 1.30% for S&P 500 and DJIA, respectively.

The Durbin-Watson statistics were found to be very close to 2, indicating that no evidence of autocorrelation was detected in any of the considered models. The F-statistics from serial correlation LM tests (aka the Breusch–Godfrey test) suggest no evidence of higher order serial correlation (for a maximum of 4 lags) in the cases considered. Once again, the results from Hierarchical Regression show that the inclusion of Twitter sentiment index (DHT) increases the explanatory power of the model compared to the Carhart four-factor model (p-value: 0.00).

Overall, the results from all of the tests conducted confirm the main hypotheses *H1* and *H2* — that investor sentiment, as expressed in Twitter daily messages, contains predictive power with respect to U.S. stock returns. These results are consistent with Kaplanski et al. (2015), and also support prior research in the field (Siganos, et al., 2014; Zhang et al., 2018) establishing that Twitter-based sentiment Granger-causes stock returns. All signs of the coefficients of Twitter sentiment in all considered cases were found to be positive and statistically significant at 1%, in line with Da et al.’s (2015) prediction. The results are consistent, too, with empirical theories that investor sentiment predicts stock returns (Baker & Wurgler, 2007; Baker et al., 2012; De Long et al., 1990; Zhang et al., 2018). The results from all considered models show that the inclusion of online search-based sentiment index significantly increases the R-squares of popular asset pricing models such as CAPM, Fama-French three-factor, and Carhart four-factor models. In terms of practical considerations, the evidence presented in this research suggests that investors and fund managers should be concerned about the new metric of investor sentiment from online sources and its effect towards stock value. This result highlights the importance of investor sentiment in short-term equity values. At the same time, the findings provide valuable practical implications that would help stock market policymakers to efficiently stabilize equity markets and prevent mispricing.

Robustness Checks

Because some professionals and academics view the S&P 500 and DJIA indexes as representative of large-cap stocks, it is possible that firm size differences could be a source of bias. To rule out any potential bias toward large-cap stocks, additional tests were performed to investigate the Wilshire 5000, S&P Midcap 400, Russell 2000, and NASDAQ composite indexes. The first three of these additional indexes were chosen as some professionals consider them representative of the total U.S. stock market, mid-cap stocks, and small-cap stocks, respectively. The results confirm the initial findings that Twitter sentiment has significant predictive power with respect to stock returns, albeit with lower significance.

Because the sample period includes observations from a few major global events (the subprime crisis in 2009 and the COVID-19 pandemic between 2019 and 2021), the trends of these subperiods were also analyzed. Overall, they showed no concrete differences from the main analysis.

This study also experimented with newer asset pricing models such as the Fama-French five-factor model (Fama & French, 2015). The results showed no material difference from the main analysis, with similar significance. More importantly, to address concerns that the main models used in this research did not correspond to the usual approach in the literature regarding time-series data, additional GRS tests (Gibbons et al., 1989) were conducted on 2x3 and 5x5 portfolios formed on size and B/M ratios in order to compare the performance of the models including a Twitter sentiment index as an augmented variable with the performance of its traditional model counterparts. It was found that models including a Twitter sentiment index as an augmented variable show lower absolute alphas compared to those of the original models, for all cases (the Fama-French three-factor, Carhart four-factor, and Fama-French five-factor models on 2x3 and 5x5 portfolios). This finding suggests that the Twitter sentiment index has explanatory power for stock returns, which is consistent with the initial findings. Finally, Hierarchical Linear Regressions were conducted to investigate whether the sentiment add explanatory power of the regression. The results indicated that the increased R^2 is statistically significant for all considered models.

In summary, the robustness tests supported the initial findings from the main analysis that Twitter-based sentiment has additional explanatory power for U.S. stock returns. This finding is robust to changes in the asset pricing models used in the study (the Fama-French three-factor,

Carhart four-factor, and Fama-French five-factor models), as well as changes in type of stocks (large, medium or small market-capitalization portfolios).

Conclusions and Implications

According to classical finance theory, investor sentiment does not play any role in stock prices, expected returns, or realized returns. Based on the behavioral framework documented in prior research, this paper provides evidence that contradicts that view. This study used a basic and straightforward model to show that Twitter sentiment index is able to explain deviations of U.S. stock returns from the “rigorous” model’s prediction. It was found that a direct survey measure of investor sentiment, as expressed in Daily Twitter messages, predicts stock returns on the same day, and that this measure has the ability to explain deviations from intrinsic values as predicted by popular asset pricing models. In all cases studied, the significance of the sentiment index was found to be significant and robust to changes in asset pricing models (the Fama-French three-factor, Carhart four-factor, or Fama-French five-factor models) and significant for all considered equity portfolios. In addition, the Twitter sentiment index was almost uncorrelated with popular risk factors, suggesting that the predictive power of Twitter sentiment index is unlikely to be captured by any known risk factors, such as market risk premium, firm size, B/M ratio, or momentum.

The results consistently suggested that the employed sentiment variable is relevant to daily market returns and helps to explain deviations from popular asset pricing models, which is consistent with prior research (e.g., Baker & Wurgler, 2007; Baker et al., 2012; Da et al., 2015; De Long et al., 1990; Kaplanski et al., 2015; Siganos et al., 2014; Zhang et al., 2018). This finding has several important implications. First, the results support behavioral theories that predict that the irrational sentiments of investors do in fact affect asset prices. Second, the findings suggest that empirical asset pricing models should acknowledge the possible role of investor sentiment.

In terms of practical implications, the results presented in this study imply that a sudden change in sentiment could translate into a large wealth shock with the potential to depress the stock market. Market regulators and government officials should be aware of the potential for market biases or “irrationalities” caused by investor sentiment from social media. Ultimately, individual investors and fund managers should be aware of the impact social media sentiment can have on both their own portfolios and fund managers’ investment strategies.

Research Limitations and Possible Future Research

One of the limitations of this study is that this research focuses primarily on U.S. stocks. This limitation arises from the lack of online sentiment data and social media user demographics on these stocks, mainly due to the limitations of natural language processing techniques for non-English languages. For example, current algorithms in artificial intelligence cannot correctly understand ambiguous words in Japanese and Chinese; hence, sentiment observations from non-anglophone investors are largely ignored due to this technological constraint. This research systematically focused primarily on the U.S. stock market because, according to Twitter usage statistics (Kemp, 2020), by far the largest number of English Twitter users are from the United States. Accordingly, the exclusive focus on stock markets in the anglophone world, and the lack of attention to emerging markets, are recognized as limitations of this study and identified as promising areas for future research should the data become available.

Another minor topic worth mentioning is that the results from the present study do not distinguish between the effects on stock prices caused by volatility and those caused by investor sentiment. There are two main reasons for this. First, prior research has documented that investor sentiment is more relevant to returns than to risks. For example, Kaplanski et al. (2015) argued that investor sentiment affects expected returns more intensely than does expected risk. Second, this position is also consistent with Da et al.'s (2015) position and is broadly in line with recent research by Ding, Mazouz & Wang (2019), who showed that the effect of sentiment on returns is not related to systematic risk.

References

- Baker, M., & Stein, J. C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets*, 7(3), 271-299. doi: 10.1016/j.finmar.2003.11.005
- Baker, M., & Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *Journal of Finance*, 61(4), 1645-1680. doi: 10.1111/j.1540-6261.2006.00885.x
- Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, 21(2), 129-152.
- Baker, M., Wurgler, J., & Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of Financial Economics*, 104(2), 272-287. doi: 10.1016/j.jfineco.2011.11.002

- Black, F. (1986). Noise. *Journal of Finance*, 41(3), 529-543.
- Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57-82.
- Da, Z., Engelberg, J., & Gao, P. (2015). The sum of all FEARS investor sentiment and asset prices. *Review of Financial Studies*, 28(1), 1-32. doi: 10.1093/rfs/hhu072
- Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under- and overreactions. *Journal of Finance*, 53(6), 1839-1885.
- De Long, J. B., Shleifer, A., Summers, L. H., & Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of Political Economy*, 98(4), 703-738.
- Dergiades, T. (2012). Do investors' sentiment dynamics affect stock returns? Evidence from the US economy. *Economic Letters*, 116(3), 404-407. doi: 10.1016/j.econlet.2012.04.018
- Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366), 427-431.
- Ding, W., Mazouz, K., & Wang, Q. (2019). Investor sentiment and the cross-section of stock returns: New theory and evidence. *Review of Quantitative Finance and Accounting*, 53(2), 493-525.
- Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., & Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and Twitter. *PLoS One*, 6(12), e26752. doi: 10.1371/journal.pone.0026752
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3-56. doi: 10.1016/0304-405X(93)90023-5
- Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, 116(1), 1-22. doi: 10.1016/j.jfineco.2014.10.010
- Fama, E. F., & French, K. R. (2020). Comparing cross-section and time-series factor models. *Review of Financial Studies*, 33(5), 1891-1926. doi: 10.1093/rfs/hhz089
- French, K. R. (2021). French data library. Retrieved January 20, 2021, from https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

- Gibbons, M. R., Ross, S., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica*, 57(5), 1121-52.
- Kaplanski, G., Levy, H., Veld, C., & Veld-Merkoulova, Y. (2015). Do happy people make optimistic investors? *Journal of Financial and Quantitative Analysis*, 50(1/2), 145-168.
- Kemp, S. (2020). Global report 2020. Retrieved February 3, 2021, from <https://www.hootsuite.com/pages/digital-2020>
- Kim S., & Kim, D. (2014). Investor sentiment from internet message postings and the predictability of stock returns. *Journal of Economic Behavior & Organization*, 107(Pt. B), 708-729. doi: 10.1016/j.jebo.2014.04.015
- Lee, C. M. C., Shleifer, A., & Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *Journal of Finance*, 46(1), 75-109.
- Lemmon, M., & Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. *Review of Financial Studies*, 19(4), 1499-1529.
- Naeem, M. A., Farid, S., Balli, F., & Shahzad, S. J. H. (2020). Can happiness predict future volatility in stock markets? *Research in International Business and Finance*, 54, 101298. doi: 10.1016/j.ribaf.2020.101298
- Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75(2), 335-346.
- Shleifer, A., & Summers, L. H. (1990). The noise trader approach to finance. *Journal of Economic Perspectives*, 4(2), 19-33.
- Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *Journal of Finance*, 52(1), 35-55.
- Siganos, A., Vagenas-Nanos, E., & Verwijmeren, P. (2014). Facebook's daily sentiment and international stock markets. *Journal of Economic Behavior & Organization*, 107(Pt. B), 730-743. doi: 10.1016/j.jebo.2014.06.004

- Zhang, W., Li, X., Shen, D., & Tegli, A. (2016). Daily happiness and stock returns: Some international evidence. *Physica A: Statistical Mechanics and its Applications*, 460, 201-209. doi: 10.1016/j.physa.2016.05.026
- Zhang, W., Wang, P., Li, X., & Shen, D. (2018). Twitter's daily happiness sentiment and international stock returns: Evidence from linear and nonlinear causality tests. *Journal of Behavioral and Experimental Finance*, 18, 50-53. doi: 10.1016/j.jbef.2018.01.005
- Zhao, R. (2020). Quantifying the cross sectional relation of daily happiness sentiment and stock return: Evidence from US. *Physica A: Statistical Mechanics and its Applications*, 538, 122629. doi: 10.1016/j.physa.2019.122629