Mood Swings and Insufficient Information Acquisition: A Study on Cross-Section of Stock Returns*

Jiatao Liu† Ian W. Marsh‡

First draft: April 23, 2019

Current draft: June 18, 2021

Abstract

This paper studies mood, measured by Twitter messages, which causes investors' insufficient acquisition of information about assets and the implications of asset pricing. Using a Twitter-based mood measure, we find that mood swings are negatively predictive of investors' acquisition of earnings-related information when seeking to learn about companies' performance. Therefore, we argue that this bias effect contributes to the explanation of classical (unconditional) pricing models' failures. Conducting tests on cross-sectional stock returns, we show that stocks that are more sensitive to mood earn a higher expected excess return than less mood-sensitive stocks. Sorting stocks to construct the risk factor portfolio based on mood betas as sensitivity to mood risk, we are the first to quantify the risk premium (0.56% per month) by holding stocks subject to mood risk. Our results are consistent with the theoretical prediction that investors mistakenly use mood as information rather than learning enough fundamental information about assets, thereby inducing mispricing in asset valuation.

JEL Classification: G12; G14; G41

Keywords: Information Acquisition; Mood; Mood Beta; Risk Premium; Anomalies

*We thank seminar and conference participants at Cass Business School, Market Microstructure Summer School at Stockholm Business School, INFINITI Conference, Young Finance Scholar Conference, The Academy of Behavioral Finance&Economics, The 32nd Australasian Finance and Banking Conference, Southwestern Finance Association 2021 and Eastern Finance Association 2021 for helpful comments. All errors are our own.

†Ph.D. Candidate in Finance, Cass Business School, Email address: `jiatao.liu@cass.city.ac.uk`

‡Faculty of Finance, Cass Business School, Email address: `i.marsh@city.ac.uk`

1 Introduction

A stock's mood beta is its sensitivity to variations in the mood of the public. As noted by Hirshleifer et al. (2020), mood can be viewed as a special case of investor sentiment, and as in their paper, our focus is on the effects of emotional valence - whether the public's mood is happy or sad.[1] We draw inspiration from the psychological study by Schwarz and Clore (2007), who propose that agents apply their feeling as information for decision-making judgment. Therefore, to enrich the economics study by incorporating behavioral factors as proposed by Tirole (2002), we seek to study whether affective states such as mood are a significant causal factor in variations in investors' learning of information about assets' payoff. In contrast to existing studies that mainly argue that mood biases investors' valuation on the factors in asset payoff or trading behaviors, we study the impact of mood on the average investor's decision on choice of information acquisition across different assets.

Instead of using weather as the customary proxy for mood,[2] our study is inspired by the growing body of literature in the field of textual analysis in finance and economics. Essentially, the measure of mood we used is from the Hedonometer project run by the University of Vermont Complex Systems Center. Hedonometer constructs a daily happiness index based on the analysis of the words used in messages posted on Twitter. A random sampling of approximately 50 million messages posted to the system (representing around 10% of the total number of messages posted each day) is then analyzed. The words from the English language messages are pooled, and this pool of words is assigned a happiness score based on the average happiness score of the words it contains.[3] As is immediately apparent, this index is not designed to be finance-oriented.

We use this happiness score as our proxy for the public's mood. Compared to the weather or sporting results - both exogenous shocks that are assumed to affect people's mood - the happiness score is an endogenous measure that reflects mood. A recent study by Edmans et al. (2021) uses Spotify music data in a similar manner - albeit over a shorter sample - to argue that their endogenous music measured sentiment captures information about mood swings. Nevertheless, the main discussion in the study by Edmans et al. (2021) follows the path of sentiment mispricing effect on market return in the literature, in that the theoretical foundation of the mood-biasing effect on in-

[1] Sentiment is a much broader term encompassing both affective concepts such as emotions - including mood - and non-affective concepts such as attention or heuristic beliefs.

[2] Weather as a classical proxy for mood has been comprehensively addressed in the literature on the effect of mood on financial market and investors' trading behavior. See related studies by Saunders (1993), Hirshleifer and Shumway (2003) and Chang et al. (2008).

[3] A more detailed discussion of the fundamental work on the Twitter mood index can be found in Dodds and Danforth (2010), Cody et al. (2016), Reagan et al. (2016) and Reece et al. (2017).

formation acquisition in our study is distinct from extant behavioral studies in sentiment or mood. Clearly, the appeal of Twitter differs across demographic groups. A recent survey by Wojcik and Hughes (2019) concludes that Twitter users are younger, more educated, have higher incomes and are more likely to identify as Democrats than the U.S. adult population as a whole.[4] On the other hand, Twitter is a close match to the population in terms of gender and ethnicity. As such, we acknowledge that the Twitter-based happiness score is an imperfect proxy for the public's mood. We would expect this to make it more difficult to find empirical support for the relationships we hypothesize. We discuss the Twitter mood data in more detail in section 2.

By virtue of textual analysis on Twitter messages, we are able to measure the public's mood and test the effect of changes in mood on investors' acquisition of information about assets. We follow the study by Weller (2018) and estimate a price jump ratio around earnings announcements as a firm-specific measure of information acquisition. This jump ratio is the post-announcement absolute cumulative abnormal return (ACAR) divided by the total ACAR, including the pre-announcement period. As investors acquire more (less) information about earnings before the announcement, we expect this jump ratio to be lower (higher). As an alternative and more direct measure, we also directly calculate the average number of SEC EDGAR file downloads. Finally, we estimate mood swings by calculating the average absolute change of the daily Twitter mood index or its average volatility in the most recent month before the firm's earnings announcement.

Empirically, we find evidence that mood swings predict lower levels of information acquisition by investors. When mood becomes either more positive or negative, investors decide to acquire less value-relevant information (regarding firm earnings) about assets before the information is released. As a consequence, investors inadequately learn fundamental information that ought to be incorporated into asset prices. In a seminal study, Van Nieuwerburgh and Veldkamp (2009) state that an asset's return and risk fall as investors learn about it. Therefore, an asset that the average investor understands well is expected to have a lower standard deviation of its return. In line with the importance of investors' information acquisition choice, Veldkamp (2011) proposes that conditional betas on information that the average investor knows must differ from the unconditional betas estimated by classical pricing models such as the single-factor CAPM or Fama-French multifactor models. As investors' learning about assets is *ad hoc*, another interesting question arises as to what the implications of the mood biasing effect on investors' information learning are for asset pricing.

Therefore, we argue that the effect of insufficient information learning caused by mood contributes to explain the failures of classical pricing models. When pricing assets that investors inadequately study, the beta-sensitivity to risk factors in unconditional pricing models loses effectiveness due to mood's causation of insufficient information incorporation on learning about the

[4]See the link for the article. https://www.pewresearch.org/internet/2019/04/24/sizing-up-twitter-users/.

risks contained in the assets. The economic intuition is that mood causes investors' inadequate asset information acquisition; thus, the structure of covariance between stochastic discount factor (multi-factors) and the risky asset return tends to deviate from the unconditional pricing models implied. More specifically, assets subject to mood's effect on insufficient learning contain more risk (high standard deviation). The betas (quantity of risk), which are the sensitivity of asset return to risk factors conditional on insufficient information acquisition induced by mood, should be higher than the unconditional betas, leading to a higher expected return. Using the unconditional betas underestimates the risk in assets; thus, researchers find almost a thousand risk factors in empirical asset pricing studies to invoke in situations in which CAPM and Fama-French factor models fail to explain cross-sectional stock returns. We propose the effect of mood-inducing investors' lack of study on assets provides the "opportunity" in empirical asset pricing studies to find new risk factors.

Nevertheless, we argue that not all assets are subject to the mood effect. We test the implications of the mood pricing effect, namely mood beta, for the cross-section of U.S. stock returns. We show that the returns of a significant proportion of U.S. stocks are sensitive to changes in Twitter mood and that mood is a significantly priced risk factor. As investors mistakenly incorporate mood as information and do not learn or acquire as much fundamental information as they should to price assets, mood as a behavioral factor adds additional risks which are not explained by asset fundamentals. We propose that stocks that are sensitive to vary with public mood (moody stocks) earn a higher expected return as a risk premium required by investors who hold these mood risky assets. To the best of our knowledge, we are the first to find that mood risk measured by the sensitivity of public mood is significantly priced in the cross-section stock returns.

In line with Lo et al. (2005) who propose that multi-factor asset pricing models can be enriched by considering the effect of emotional factors, we estimate the mood beta by using the method proposed by Bali et al. (2017). This involves the identification of mood-sensitive stocks by adding the Twitter mood index into the Fama-French five-factor Model in line with the momentum factor.[5] By sorting stocks according to their sensitivity to changes in the Twitter mood index, we show that portfolios of negatively (positively) mood-sensitive stocks earn excess returns of 1.66% (1.65%) per month. Portfolios of non-moody stocks earn excess returns of just 1.1% per month. We construct a mood-mimicking portfolio by taking long positions in stocks that have a large (positive or negative) sensitivity to mood and shorting stocks which are mood-insensitive. The mimicking mood portfolio has an average return of 0.56% per month. Our empirical findings are consistent with the theoretical implication that stocks investors insufficiently learn about are riskier with a higher expected return.[6]

[5] Bali et al. (2017) measure stocks' sensitivity to economic uncertainty by taking the regression coefficients on the uncertainty index in the time-series regression as the uncertainty beta.

[6] The empirical findings also strongly confirm the recent seminal studies of mood as a behavioral factor contributing

We apply standard cross-sectional asset pricing techniques to test whether these mood betas are priced in the U.S. equity market. We first sort stocks into ten portfolios according to rolling sensitivities to Twitter mood to identify which stocks are the most sensitive to either positive mood or negative mood. Stocks' sensitivity to mood is induced by its effect on investors' trading behaviors and risk sensitivity. In the *Homo economicus* paradigm, investors' valuation of assets should be rationally based on acquired or learned fundamental information. However, as we find empirically, investors incorporate less fundamental information when their mood is more volatile, in either the positive or negative directions. Therefore, when investors become moody, they mistakenly rely on their feelings as useful information with which to trade or price assets and do not acquire as much fundamental information as they should. On the one hand, when investors' mood is more positive, they are less risk-averse and tend to overprice assets (Bassi et al., 2013; Kaplanski et al., 2015), as a result of which they invest in more risky stocks, exerting buying pressure on these stocks.[7] On the other hand, when investors' mood is more negative, their pessimistic feeling causes them to be more risk-averse and perceive higher risk. Raghunathan and Pham (1999) argue that agents with a sad mood are biased in favor of high risk with high reward, on the grounds that investors seek stocks which they believe to generate high returns in negative mood days to compensate for the high risk entailed by those stocks.[8] As a consequence, positive moody stocks' returns increase as mood is more optimistic and negative moody stocks' returns increases as mood is more pessimistic. In our mood beta estimation, we find that portfolio 1 stocks with an average mood beta of -0.58 are those sensitive to negative mood, while stocks in portfolio 10 with average mood betas of 0.61 are those sensitive to positive mood. All in all, regardless of whether stocks are sensitive to either positive or negative mood, investors' risk perceptions and trading behaviors are biased by mood via the decision to acquire less information . These stocks are more likely to be affected by the irrational trigger-mood and are more risky than stocks which are less likely to be affected by mood.

First, the value-weighted excess returns of portfolios 1 and 10 are 1.66% and 1.65% per month. The average excess return of mood-insensitive stocks in portfolios 5 and 6 is around 1.1%. A high-low portfolio that takes a long position in both portfolio 1 (negative mood sensitivities) and portfolio 10 (positive mood sensitivities) and short positions in the mood-insensitive portfolios 5 and 6 generates a statistically significant average excess return of 0.52% per month.[9] Analysis

to mispricing and risks to the financial markets (Goetzmann et al., 2015; Bushee and Friedman, 2016; Hirshleifer et al., 2020).

[7] By studying investors' behavior in Finland which is considered more likely to be affected by people's mood, Kaustia and Rantapuska (2016) argue that positive mood measured by sunshine light length drives investors to buy more than they sell. Goetzmann et al. (2015) find evidence that even institutional or sophisticated investors are subject to cognitive biases such as mood, with optimism increasing buy-sell imbalances.

[8] Shu (2010) develops a model to indicate that the pessimistic mood causes investors' risk aversion and impatience to increase, as a result of which the stochastic discount factor is decreased to price the asset with a higher return.

[9] The long-short strategy takes long positions in positive and negative mood beta stocks (and short positions in zero

based on the Fama-French five-factor model suggests an alpha of 0.48% per month. This rises to 0.50% with the addition of the Carhart momentum factor, and to 0.54% with the further addition of long- and short-term reversal factors. The *t*-statistics on these alphas range from 3.25 to 4.18.

Second, we construct a mimicking mood factor portfolio to determine whether the risk premium induced by mood can be captured by benchmark pricing factors. This mood factor earns a statistically significant risk premium of 0.56% per month. However, it is positively correlated with the market, size and reversal factors and negatively correlated with profitability and momentum factors. Taking the market factor into account leaves an unexplained mean return of 0.38% per month. As successive extra factors are accounted for, the alpha increases back to 0.56%, equal to the mean return on the mood factor. We construct the orthogonalized mood factor as the component of the mimicking mood factor unexplained by all other factors. Standard factor models fail to explain the returns earned by stocks most affected by mood and the decile of stocks with the largest absolute mood beta earn a statistically significant alpha of between 0.36 and 0.41% per month, while the long-short mood strategy earns an alpha of between 0.47 and 0.55%. Adding the orthogonalized mood factor to the analysis reduces all alphas and removes all statistical significance. As the two portfolios containing stocks with the largest absolute mood betas each load significantly on the orthogonalized mood factor, we infer that it has significant pricing power.

Finally, we construct 25 portfolios based on independent sorts of market capitalization and absolute mood beta. Within each size quintile, the most mood-sensitive stocks earn higher mean returns than the least mood-sensitive stocks. This effect is economically large and statistically significant for all size quintiles. Alphas from the alternative factor models are positive and significant for the most mood-sensitive quintile of stocks in the majority of size quintiles.[10] A high minus low mood sensitivity strategy yields a positive mean alpha in all five size quintiles, largest in magnitude and statistically significant for the smallest quintile. Adding the mood factor to the analysis removes all significant alphas for the most mood-sensitive stocks and for all the high minus low mood sensitivity strategies. Indeed, incorporating the mood factor turns the slope of alphas with respect to mood sensitivity negative for the larger quintiles of stocks.

Additionally, we find that stocks which are sensitive to public mood - positively or negatively - are typically small in size, relatively young, pay lower dividends, have a large R&D ratio, are not profitable, engage in more external financing and have higher levels of idiosyncratic risk. Alternatively, these characteristics share similarities with the link between information asymmetry and stocks that are subject to the effect of mood. The theoretical channel of mood causing less information acquisition in this study, in fact, has a corresponding implication of asymmetric infor-

mood beta stocks). However, the strategy is not a zero-mood-beta smart money scheme, since combining positive and negative mood-sensitive stocks does not remove the exposure to risk from either positive or negative mood changes.

[10]They are mainly in the smallest (moody stocks in respect of size 1), mid-cap (moody stocks in respect of size 4) and large-cap size quintiles (moody stocks in respect of size 5).

mation problem in assets. Meanwhile, our findings are consistent with the study by Bushee and Friedman (2016), who argue that both noise traders and unsophisticated investors are more likely to take feelings or non-fundamental factors as useful information to price asset values and trade stocks when other information is lacking.

Although we emphasize that our study of public mood sensitivity is different from the more general concept of sentiment, the two concepts are clearly related and as such we contribute to the literature on sentiment and asset pricing. The exact meaning of sentiment is obscure but, as noted by Baker and Wurgler (2006), one possible definition is the propensity to speculate. Shifts in the propensity to speculate drive shifts in the relative demand for speculative investments and therefore have cross-sectional effects. Much of the literature draws on Baker and Wurgler (2006; 2007), who construct a sentiment index and use this to demonstrate a significant impact of investor sentiment on a cross-section of U.S. stock returns. They find that difficult-to-arbitrage stocks have a negative relation between sentiment and subsequent returns. Baker et al. (2012) confirm the power of investor sentiment in international stock markets. Other studies on the importance of investor sentiment for stock returns include those of Swaminathan (1996), Brown and Cliff (2004), Kumar and Lee (2006), Lemmon and Portniaguina (2006) and Stambaugh et al. (2012). Unlike our Twitter mood index, the sentiment indices most commonly used in the literature are constructed from proxy measures extracted from financial markets. The Baker and Wurgler (2006) sentiment measure, for example, is based upon closed-end fund discounts, market turnover, IPO numbers and first-day returns, the equity share in new issues and the dividend premium. Huang et al. (2015) use the same six proxies but a different statistical method to construct an alternative sentiment index which they show also supports the pricing power of investor sentiment.

Our study makes a unique contribution to the classical literature that favors the use of measures of mood more exogenous to financial markets. Saunders (1993) observes that NYSE stocks tend to have positive returns on sunny days and moderate returns on cloudy days. Hirshleifer and Shumway (2003) support the notion that good mood is associated with sunny weather and they find that there is a highly significant correlation between sunshine and stock returns. Kamstra et al. (2003) document the existence of an effect of seasonal affective disorder (SAD) — a psychological condition in which a daylight deficit has a detrimental impact on people's mood — on stock market returns around the world. Kamstra et al. (2000) show that stock market returns on Mondays following daylight saving clock changes are lower than returns on normal Mondays. They propose that market participants' loss of an hour of sleep may result in increased anxiety or risk aversion which adversely affects stock market returns. Edmans et al. (2007) find that sporting events affect investor mood, with soccer defeats in particular being associated with significant market decline in a country.[11] However, our study concentrates more on cross-sectional analysis instead of the ag-

[11] Goetzmann and Zhu (2005) propose that the relationship between mood or weather effects and stock returns may

gregate market. In addition, our key premise is to determine whether the significant effect of mood causing less information acquisition about assets can be seen as a risk factor on cross-sectional asset returns and whether this risk should be compensated for by investors who hold the stocks which are thought of as the ”volunteers” of irrational trading or pricing behaviors triggered by public mood.

Furthermore, our study also contributes to the growing body of studies in textual analysis in financial markets,[12] particularly using Twitter message data. For example, Bollen et al. (2011) conduct textual analysis from large-scale Twitter feeds based on a computational algorithm to measure mood. They claim, somewhat controversially, that mood calculated from Twitter feeds has predictive power with regard to the DJIA value index. Additionally, Dey (2014) shows that tweets contain useful information. He studies the polarity value of each tweet by sentiment analysis and finds a significant correlation between changes in stock price and changes in the polarity values of tweets. Both of studies use time series techniques to relate Twitter mood to stock returns. Our analysis focuses on cross-sectional analyses and seeks specifically to quantify the premium associated with mood risk in an asset pricing framework.

The study is structured as follows. Section 2 describes the data for mood and the empirical evidence of the mood effect on investors' information acquisition in line with theoretical development in asset pricing. Section 3 details mood beta estimation and portfolio. We discuss the key asset pricing tests and results of the mood factor in section 4. In section 5, we thoroughly discuss theoretical motivations and connections to our empirical findings. Section 6 briefly outlines the robustness tests we conducted. Section 7 provides our conclusions drawn from the analysis.

2 Mood Data and Theoretical Motivation

2.1 Mood Measurement from Twitter Message

We use the daily mood score from a Twitter text analysis project supported by Hedonometer.org, based at the University of Vermont Complex Systems Center. In their research, the data generate a “Twitter Happiness Score” which is explored as a function of time, space, demographics and network structure using Twitter feeds as a data source. Hedonometer samples roughly 50 million messages posted on Twitter each day. Words in messages written in English are extracted (resulting in approximately 100 million words each day) and this pool of words is assigned a happiness score based on the average happiness score of the scored words. Hedonometer considers 10,222 scored key words used to calculate mood. These are the most frequently used words in Google

be responsible for the behavior of market makers rather than individual investors.

[12]A comprehensive review and survey can be found in Tetlock (2014).

Books, New York Times articles, music lyrics and Twitter messages. Each word has been assigned a happiness score (ranging from 1=sad to 9=happy) using Amazon's Mechanical Turk service. Words scoring high on this scale include laughter (8.5), happiness (8.44), love (8.42), excellent (8.18), joy (8.16), successful (8.16) and win (8.12). Low-scoring words include terrorist (1.30), suicide (1.30), murder (1.48), death (1.54) and cancer (1.54).

The algorithm used to measure the mood score is as follows:

$$h_{avg}(T) = \frac{\sum_{i=1}^{N} h_{avg}(w_i) f_i}{\sum_{i}^{N} f_i} = \sum_{i=1}^{N} h_{avg}(w_i) p_i \tag{1}$$

In a given text T, f_i is the frequency of the $i\ th$ word w_i, and the estimation of happiness for a unique word is $h_{avg}(w_i), p_i = f_i/(\sum_{j=1}^{N} f_j)$. The given text T can be extracted flexibly based on different time intervals. For example, Twitter feeds used to calculate the mood score can be extracted as counting either minutes or days. Hedonometer.org reports that there are about 20 million tweets per day and approximately 14,000 per minute as at August 31 2011 (Dodds et al., 2011).[13] A day is considered happier than usual if happy words are used more frequently than usual, or if sad words are used less frequently than usual.

The daily Twitter score data are available from September 2008; our sample ends in December 2016. The upper panel of Figure 1 plots the Twitter mood score at a daily frequency and the lower panel plots the Baker and Wurgler sentiment index at a monthly frequency. The time range in the two plots covers the interval between September 2008 and September 2015 for which both data series are available.[14]

2.2 Mood Impact on Information Acquisition

We investigate how mood plays a role in affecting investors' economic decisions and behaviors. The key argument we want to address is that investors mistakenly incorporate mood as useful information (Schwarz and Clore, 2007). Consequently, investors rely on less fundamental information to price or trade assets when the average investor becomes moody. Therefore, the hypothesis we are going to test is:

Hypothesis 1: As mood tends to be more volatile, investors tend to learn less information about assets.

[13] The analysis of Twitter mood is from a random sample of 25% of the tweets in the database. There are about 230 million unique words. Due to the computational difficulty, Hedonometer.org analyzes the first 50,000 most frequent words without compromising estimation accuracy.

[14] The sentiment index is downloaded from Wurgler's website where the final observation is September 2015.

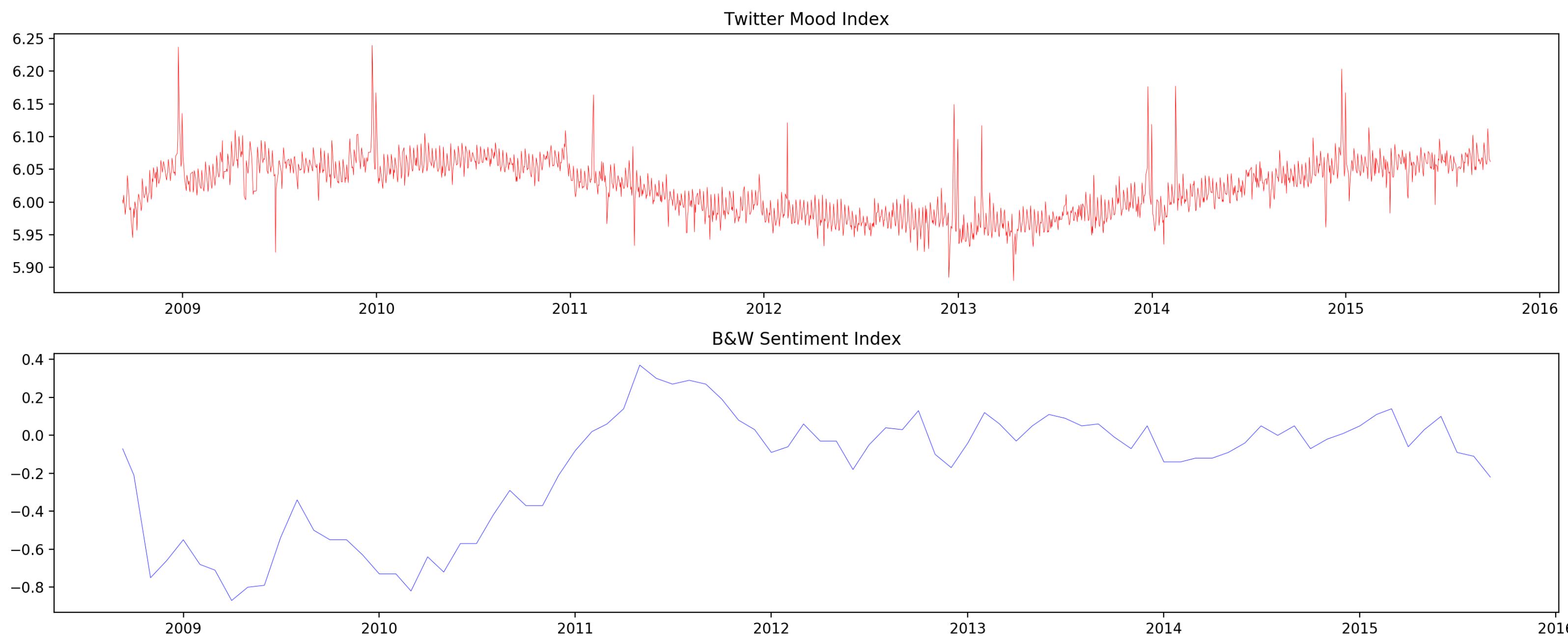

Figure 1: Twitter Mood vs. Baker & Wurgler Sentiment Index

2.2.1 Measure of Information Acquisition

Because investors' information acquisition is not directly observed, we conduct an event study of firm earnings announcements to test the inverse relationship between mood and investors' asset information acquisition.[15] First, we follow a novel study by Weller (2018) to estimate a price jump ratio as the measure of investor information acquisition. The price jump ratio is estimated by taking the post-announcement ACAR divided by the total ACAR that includes pre-announcement periods of around 21 days:

$$Jump_{i,t}^{a,b} = \frac{CAR_{i,t}^{T-1,T+b}}{CAR_{i,t}^{T-a,T+b}} \tag{2}$$

where $a = 21$ and $b = 2$ as the pre- and post-announcement window respectively. The CAR is the absolute cumulative abnormal return subject to the study window from the Fama-French five-factor model and also include the momentum factor:

$$CAR_{i,t}^{j_1,j_2} = \sum_{t=j_1}^{j_2} \left(R_{it}^{e} - \alpha_i - \sum_{m=1}^{M} \beta_{i,m} f_{m,t} \right) = \sum_{t=j_1}^{j_2} \varepsilon_{i,t} \tag{3}$$

where R_{it}^{e} is stock excess return and $f_{m,t}$ is the Fama-French and the momentum factors. We estimate the α_i and $\beta_{i,m}$ based on 252 daily observations and 90 days before the earnings announcement. We require firms that have observations on at least 63 trading days to conduct equation (3) as the estimation. To avoid the zero denominator issue in equation (2), we follow the instruction from Weller (2018) to set a threshold as $|CAR_{i,t}^{T-21,T+2}| > \sqrt{24}\hat{\sigma}_{i,t}$ where $\hat{\sigma}_{i,t}$ is daily return volatility in 24-day event window.

The rationale for the price jump ratio as a proxy of investors' information acquisition is that, as investors acquire more earnings-related information to learn about the company before the day of the announcement, due to price discovery, the price incorporates more information about earnings in the pre-announcement period. Therefore, we should observe a large denominator in equation (2). On the contrary, if investors do not acquire earnings-relevant information to learn about the company before the announcements, the stock price will jump immediately as the earnings are released at the announcement day and afterward.[16] Consequently, we should expect a large numerator relative to the denominator in equation (2). In sum, the higher the price jump ratio implies a lower firm-specific information acquisition conducted by investors to learn about the company.

In addition to the price jump ratio measure, we estimate investors' demand for learning information about companies by exploiting the SEC EDGAR logs of access to firm-specific fillings around a quarterly earnings announcement. Specifically, we calculate the average total count of

[15] Earnings announcement date and relevant data are extracted from the IBES dataset. Stock return data is from CRSP, and financial fundamentals are from Compustat. See section 2.4 and Appendix B for details.

[16] We use 2 days after the announcement as the post-announcement periods to capture the PEAD effect.

search volume for the files in the most recent month before the announcement. We then take the natural logarithm for the average of total SEC files searching volumes ($LSECV$). To some extent, the count of SEC EDGAR file searching volume is a more straightforward way to understand investors' demand for learning. As the searching volume increases before the announcement, investors are more eager to learn about the company to forecast or estimate its upcoming earnings performance.

2.2.2 Measure of Mood Swings

We use two measures to identify mood swings from the Twitter data. First, we take the average absolute change of daily mood in the most recent month before the firm earnings announcement. Therefore, regardless of the positive or negative direction of the mood change, the larger the average value of the absolute difference in the mood, the moodier the investors become. Second, we directly calculate the volatility of the daily mood in the most recent month before earnings announcements as a direct measure of mood swings. Finally, the data in the following test is from September 2008 to December 2016.

2.2.3 Hypothesis Testing

We conduct fixed effect regressions to test the *Hypothesis 1* as follows:

$$Dep_{i,t} = \beta_0 + \beta_1 MoodSwing_{t-31,t-1} + X\delta + \varepsilon_{i,t} \tag{4}$$

$Dep_{i,t}$ is either the price jump ratio from equation (2) or the direct measure of investors' information demand $LSECV$ as the proxy of the average investor's learning about firm-specific information. $MoodSwing$ is either the average absolute daily mood change or mood change volatility in the most recent month before announcements. The X is a vector of control variables (see detailed definitions in Appendix B) and the δ is a vector of coefficients. Since the investors' information acquisition or demand is significantly related to economic uncertainties (Benamar et al., 2019; Andrei et al., 2020), we add VIX and Economic Policy Uncertainty Index (EPU) by Baker et al. (2016) as additional control variables to identify the impact of mood more clearly. We are interested in testing whether β_1 has a significant inverse relationship with investors' information acquisition proxies.

Panel A in Table 1 shows both the volatilities of mood change and the absolute change of mood as a measure of investors' mood swings' significant prediction of a positive price jump ratio. This positive predictability implies that when investors become moody, they acquire less fundamental information to learn about firms' earnings before they are released. Consequently, less earnings-related information is incorporated into the price before the announcements, resulting in

Table 1: Mood Swings Impact on Investors' Information Acquisition

This table presents the results of regressions of the price jump ratio and the counts of EDGAR SEC file downloads as the proxy for investors' firm-specific information acquisition on mood swings measured by either the average absolute change of daily mood or the volatility of daily mood during the firm earnings announcement window. Columns (1)–(4) are based on the fixed-effect regression from equation (4): $Dep_{i,t} = \beta_0 + \beta_1 MoodSwing_{t-31,t-1} + X\delta + \varepsilon_{i,t}$, where $Dep_{i,t}$ is either the price jump ratio ($jump_{i,t}$ in Panel A) from equation (2) or information demand measure that is the natural logarithm for the average of total SEC files searching volumes ($LSECV_{i,t}$ in Panel B) in the most recent month before the announcements. Control variables include: economic uncertainty proxies (VIX or EPU) and the Number of Analyst Forecast is calculated as 21 days until one day before announcement. Size, Turn, Price, Return Volatility, and Institutional Ownership are calculated as 42 days up to 21 days before the announcement. Additionally, we control the day-of-the-week effect (DOW) in Panel B when we use $LSECV_{i,t}$ to measure investors' information acquisition. Detailed definition of all variables are available in Appendix B. Standard errors are clustered by both firm- and time-fixed effect in column (1)-(4). $***$, $**$, $*$ indicate statistical significance at the two-sided 1%, 5%, 10% levels, respectively.

Panel A: Information Acquisition Measured by Price Jump Ratio

	(1)	(2)	(3)	(4)
Dependent Variable	$Jump_{i,t}$	$Jump_{i,t}$	$Jump_{i,t}$	$Jump_{i,t}$
$\sigma(Mood_{t-31,t-1})$	0.558***	0.509**		
	(0.216)	(0.216)		
$\overline{\lvert\Delta Mood_{t-31,t-1}\rvert}$			22.955**	25.620**
			(9.961)	(9.967)
$VIX_{t-21,t-1}$	−0.002***		−0.002***	
	(0.000)		(0.000)	
$EPU_{t-21,t-1}$		−0.000		−0.000
		(0.000)		(0.000)
$Size_{i,t-42,t-21}$	0.005	0.010	0.005	0.010
	(0.007)	(0.007)	(0.007)	(0.007)
$Turn_{i,t-42,t-21}$	−0.004	−0.002	−0.004	−0.002
	(0.003)	(0.003)	(0.003)	(0.003)
$Price_{i,t-42,t-21}$	0.006	0.005	0.006	0.005
	(0.007)	(0.007)	(0.007)	(0.007)
$RV_{i,t-42,t-21}$	−0.010*	-0.025***	-0.009*	-0.024***
	(0.005)	(0.005)	(0.005)	(0.005)
$NUMEST_{i,t-21,t-1}$	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
$ITOW_{i,t-42,t-21}$	0.026*	0.023*	0.027**	0.024*
	(0.013)	(0.013)	(0.013)	(0.013)
FE Month	Yes	Yes	Yes	Yes
FE Firm	Yes	Yes	Yes	Yes
Observations	25,597	25,597	25,597	25,597
R-squared	0.010	0.008	0.010	0.008
Number of Firms	3,442	3,442	3,442	3,442
Clustered Standard Errors in Parentheses				
$***p < 0.01, **p < 0.05, *p < 0.1$				

Panel B: Information Acquisition Measured by $LSECV_{i,t}$

	(1)	(2)	(3)	(4)
Dependent Variable	$LSECV_{i,t}$	$LSECV_{i,t}$	$LSECV_{i,t}$	$LSECV_{i,t}$
$\sigma(Mood_{t-31,t-1})$	-1.282***	-1.511***		
	(0.321)	(0.326)		
$\overline{\lvert\Delta Mood_{t-31,t-1}\rvert}$			−57.222***	-62.151***
			(14.926)	(15.113)
$LSECV_{i,t-62,t-31}$	0.255***	0.255***	0.255***	0.255***
	(0.008)	(0.008)	(0.008)	(0.008)
$VIX_{t-21,t-1}$	0.003**		0.003**	
	(0.001)		(0.001)	
$EPU_{t-21,t-1}$		0.001***		0.001***
		(0.000)		(0.000)
$Size_{i,t-42,t-21}$	0.024**	0.025**	0.024**	0.025**
	(0.010)	(0.010)	(0.010)	(0.010)
$Turn_{i,t-42,t-21}$	0.037***	0.037***	0.037***	0.037***
	(0.005)	(0.005)	(0.005)	(0.005)
$Price_{i,t-42,t-21}$	−0.020*	-0.021**	-0.021**	-0.022**
	(0.011)	(0.011)	(0.011)	(0.011)
$RV_{i,t-42,t-21}$	0.015**	0.017**	0.014**	0.016**
	(0.007)	(0.007)	(0.007)	(0.007)
$NUMEST_{i,t-21,t-1}$	−0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$ITOW_{i,t-42,t-21}$	0.032	0.033	0.033	0.034*
	(0.020)	(0.020)	(0.020)	(0.020)
DOW	Yes	Yes	Yes	Yes
FE Year-Quarter	Yes	Yes	Yes	Yes
FE Firm	Yes	Yes	Yes	Yes
Observations	25,597	25,597	25,597	25,597
R-squared	0.841	0.841	0.841	0.841
Number of Firms	3,442	3,442	3,442	3,442

Clustered Standard Errors in Parentheses

$***p < 0.01, **p < 0.05, *p < 0.1$

a large price jump when the earnings are announced. For example, in column (3) an increase in the absolute change of mood by one standard deviation (2 basis points) is associated with a 1.15% (relative to the median jump ratio of 0.4) decrease in the proportion of earnings-related price impact that arises pre-announcement. Notably, all results across columns are robust after fixed effects and other controls that may explain investors' information acquisition. Panel B in Table 1 is the test, using the count of SEC file downloads as the proxy of investors' information acquisition. Not surprisingly, the results are consistent with those of Panel A (with the price jump measure). For instance, in column (1), an increase in one unit of the mean value of mood volatility (0.24) is associated with a 5% (relative to the median $LSECV$ value of 6.24) decrease in investors' downloads of company SEC files. As investors have mood swings, they are less willing to download the company's SEC files, showing a decrease in information demand.

Additionally, we also perform an analysis to disentangle the biasing effect from either negative or positive mood swings. We expect that the absolute change in mood, either upwards or downwards, biases investors' decision to acquire firm-specific information. First, we split the absolute change of mood by positive and negative daily percentage change. Second, we calculate the average absolute change of positive mood ($|\overline{\Delta Mood^+}|$) or negative mood ($|\overline{\Delta Mood^-}|$) independently in the most recent month extended back to 34 days before the earnings announcement.[17] Finally, we also control the proportion of positive or negative mood change days ($\%Positive, \%Negative$) in the month before the earnings announcement. Unsurprisingly, the results in Table 2 are consistent with Table 1. By splitting the mood swings into either the upward or downward, Panels A and B in Table 2 clearly show that the absolute change of positive or negative mood induces less firm-specific information acquisition proxied by pricing jump ratio and SEC file downloads respectively.

Based on the empirical evidence we find in the data, the mood has a significant impact on investors' learning about firm-specific information (here, earnings-related in the test). As stated in O'Hara (2003), classical asset pricing models (CAPM, APT, etc.) are at the root of symmetric information. However, when forming a portfolio with multiple assets, investors are subject to information choice constraints for learning about assets (Van Nieuwerburgh and Veldkamp, 2009). Therefore, investors' heterogeneous learning across different assets implies that the pricing models based on the symmetric information may not be effective as researchers expect in explaining the cross-section asset return predictability. Altogether, we conduct a normative analysis through the

[17] Because either absolute positive or negative change calculates the mood swings in this test, an issue may arise whereby the mood swings data may calculate after the initial days in the proxy of investors' information acquisition. For example, the absolute positive or negative mood changes may be calculated between $t-s$ to $t-1$. In such a case, mood swings may not fully capture the biasing effect in the period before $t-s$ if s is far less than 31. We therefore extend the data back to $t-34$ to mitigate this issue. However, the results are not sensitive to how long we extend the data to calculate the split mood swings.

Table 2: Positive or Negative Mood Swings Impact on Investors' Information Acquisition

This table presents the results of regressions of the price jump ratio and the counts of EDGAR SEC file downloads as the proxy for investors' firm-specific information acquisition on mood swings measured by either the average absolute positive change of daily mood or negative change of daily mood during the firm earnings announcement window. Columns (1)–(4) are based on the fixed-effect regression from equation (4): $Dep_{i,t} = \beta_0 + \beta_1 MoodSwing_{t-34,t-1} + X\delta + \varepsilon_{i,t}$, where $Dep_{i,t}$ is either the price jump ratio ($jump_{i,t}$ in Panel A) from equation (2) or information demand measure that is the natural logarithm for the average of total SEC files searching volumes ($LSECV_{i,t}$ in Panel B) in the most recent month before the announcements. Control variables include: economic uncertainty proxies (VIX or EPU) and the Number of Analyst Forecast is calculated as 21 days until one day before announcement. Size, Turn, Price, Return Volatility, and Institutional Ownership are calculated as 42 days up to 21 days before the announcement. We also control the proportion of positive or negative mood change days ($\%Positive, \%Negative$) in the month before the earnings announcement (see detailed definitions in Appendix B). Additionally, we control the day-of-the-week effect (DOW) in Panel B when we use $LSECV_{i,t}$ to measure investors' information acquisition. Detailed definition of all variables are available in Appendix B. Standard errors are clustered by both firm- and time-fixed effect in column (1)-(4). $***$, $**$, $*$ indicate statistical significance at the two-sided 1%, 5%, 10% levels, respectively.

Panel A: Information Acquisition Measured by Price Jump Ratio

	(1)	(2)	(3)	(4)
Dependent Variable	$Jump_{i,t}$	$Jump_{i,t}$	$Jump_{i,t}$	$Jump_{i,t}$
$\overline{\lvert\Delta Mood^{+}_{t-34,t-1}\rvert}$	10.763***	8.907***		
	(2.786)	(2.782)		
$\%Positive$	−0.046	-0.016		
	(0.032)	(0.031)		
$\overline{\lvert\Delta Mood^{-}_{t-34,t-1}\rvert}$			7.251***	5.794**
			(2.730)	(2.723)
$\%Negative$			0.128***	0.083**
			(0.037)	(0.036)
$VIX_{t-21,t-1}$	−0.002***		-0.002***	
	(0.000)		(0.000)	
$EPU_{t-21,t-1}$		−0.000		-0.000
		(0.000)		(0.000)
$Size_{i,t-42,t-21}$	0.005	0.011	0.005	0.011
	(0.007)	(0.007)	(0.007)	(0.007)
$Turn_{i,t-42,t-21}$	−0.004	-0.002	-0.004	-0.002
	(0.003)	(0.003)	(0.003)	(0.003)
$Price_{i,t-42,t-21}$	0.005	0.004	0.005	0.004
	(0.007)	(0.007)	(0.007)	(0.007)
$RV_{i,t-42,t-21}$	−0.010*	-0.026***	-0.009*	-0.025***
	(0.005)	(0.005)	(0.005)	(0.005)
$NUMEST_{i,t-21,t-1}$	0.001***	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
$ITOW_{i,t-42,t-21}$	0.026**	0.023*	0.027**	0.024*
	(0.013)	(0.013)	(0.013)	(0.013)
FE Month	Yes	Yes	Yes	Yes
FE Firm	Yes	Yes	Yes	Yes
Observations	25,592	25,592	25,592	25,592
R-squared	0.011	0.009	0.011	0.008
Number of Firms	3,442	3,442	3,442	3,442
Clustered Standard Errors in Parentheses				
$***p < 0.01, **p < 0.05, *p < 0.1$				

Panel B: Information Acquisition Measured by $LSECV_{i,t}$

	(1)	(2)	(3)	(4)
Dependent Variable	$LSECV_{i,t}$	$LSECV_{i,t}$	$LSECV_{i,t}$	$LSECV_{i,t}$
$\overline{\lvert\Delta Mood^{+}_{t-34,t-1}\rvert}$	-9.292^{**}	-10.577**		
	(4.513)	(4.532)		
$\%Positive$	-0.030	-0.033		
	(0.059)	(0.059)		
$\overline{\lvert\Delta Mood^{-}_{t-34,t-1}\rvert}$			-8.577**	-11.348***
			(3.902)	(3.985)
$\%Negative$			-0.061	-0.083
			(0.066)	(0.066)
$LSECV_{i,t-62,t-31}$	0.255***	0.255***	0.255***	0.255***
	(0.008)	(0.008)	(0.008)	(0.008)
$VIX_{t-21,t-1}$	0.003**		0.003**	
	(0.001)		(0.001)	
$EPU_{t-21,t-1}$		0.001***		0.001***
		(0.000)		(0.000)
$Size_{i,t-42,t-21}$	0.024**	0.024**	0.024**	0.024**
	(0.010)	(0.010)	(0.010)	(0.010)
$Turn_{i,t-42,t-21}$	0.037***	0.037***	0.037***	0.037***
	(0.005)	(0.005)	(0.005)	(0.005)
$Price_{i,t-42,t-21}$	-0.020*	-0.021**	-0.020*	-0.021**
	(0.011)	(0.011)	(0.011)	(0.011)
$RV_{i,t-42,t-21}$	0.015**	0.017**	0.015**	0.016**
	(0.007)	(0.007)	(0.007)	(0.007)
$NUMEST_{i,t-21,t-1}$	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
$ITOW_{i,t-42,t-21}$	0.033	0.033	0.033	0.033
	(0.020)	(0.020)	(0.020)	(0.020)
DOW	Yes	Yes	Yes	Yes
FE Year-Quarter	Yes	Yes	Yes	Yes
FE Firm	Yes	Yes	Yes	Yes
Observations	25,592	25,592	25,592	25,592
R-squared	0.841	0.841	0.841	0.841
Number of Firms	3,442	3,442	3,442	3,442
Clustered Standard Errors in Parentheses				
$***p<0.01, **p<0.05, *p<0.1$				

channel of investors' information acquisition by arguing mood as a behavioral factor mistakenly incorporated into pricing assets. In sum, using unconditional pricing models in situations in which investors' learning about assets are severely affected by mood contributes to the possibility of the exploration of hundreds of latent risk factors in empirical asset pricing studies.

2.3 Theoretical Development

As stated in Veldkamp (2011), the asset's risk (standard deviation of return) will be subject to its payoff information as understood by the average investor. In line with this rationale, other things being equal, an asset that the average investor learns less (more) about is more (less) risky for investors to hold and requires higher (lower) expected returns. Hence assets' risk will be sensitive to investors' level of learning about the assets' payoff. We discuss the theoretical implication of investors' information learning for asset pricing through the general stochastic discount factor (SDF) pricing model, equivalent to factor pricing models. More specifically, we argue in some detail that the behavioral factor-mood can cause the failure of classical (unconditional) beta representation models through the biased effect of insufficient information learning.

A risky asset's return is R_i is priced by SDF pricing model:

$$E(R_i) = \frac{1}{E(m)} - \frac{Cov(m, R_i)}{E(m)} \tag{5}$$

Multiplying and dividing by $Var(m)$,the beta representation is:

$$\begin{aligned} E(R_i) &= \alpha + \Big(\frac{Cov(m, R_i)}{Var(m)}\Big)\Big(-\frac{Var(m)}{E(m)}\Big) \\ E(R_i) &= \alpha + \beta_{i,m}\lambda_m. \\ E(R_i) &= \alpha + \rho_{i,m}\frac{\sigma(R_i)}{\sigma(m)}\lambda_m \end{aligned} \tag{6}$$

where $\alpha \equiv 1/E(m)$.

Because the SDF m can be an affine function of factors (single-factor such as CAPM or multi-factor such as Fama-French factors etc.), factor models are equivalent to SDFs. The beta representation can be easily transformed into the factor pricing model as the SDF m is a linear combination of factors $f_1, ..., f_k$. For example:

$$m = a + b_1 f_1 + ... + b_k f_k \tag{7}$$

where a is a constant and $b_1, ..., b_k$ are the factor coefficients. Let $F = (f_1, ..., f_k)'$ and $m = a + b'F$ is an SDF for a constant a and the constant vector b. Therefore equation (5) can be written as:

$$E(R_i) = \frac{1}{E(m)} - \frac{Cov(b'F, R_i)}{E(m)} \tag{8}$$

As we normally work with the excess return $E(R_i^e) = E(R_i) - 1/E(m)$ and $E(mR_i^e) = 0$, equation (8) is equivalent to:

$$\begin{aligned} E(mR_i^e) &= E(R_i^e) + b'Cov(F, R_i^e) \\ E(R_i^e) &= -b'Cov(F, R_i^e) \\ E(R_i^e) &= -b'Var(F)Var(F)^{-1}Cov(F, R_i^e) = \lambda'\beta \end{aligned} \tag{9}$$

Therefore, the β and λ are :

$$\beta = Var(F)^{-1}Cov(F, R_i^e), \quad \lambda = -Var(F)b$$

No matter which pricing models we begin from, in a symmetric information world, we customarily assume that the estimated unconditional betas in the right hand side regression carry full pricing information in respect of the risk factors (O'Hara, 2003). However, as stated in Van Nieuwerburgh and Veldkamp (2009), investors choose to learn about assets based on their initial heterogeneous beliefs about the covariance structure of assets' payoff. Because learning will affect the conditional variance of assets' payoff, the betas that are conditional on information that the average investor knows are different from the estimated betas based on unconditional pricing models.

If investors thoroughly learn all available information about assets' payoff, using the unconditional model to price does not generate significantly positive abnormal returns (α). An asset that the average investor learns more about, this implies a lower beta conditional on this information acquisition by investors. As proposed by Veldkamp (2011), learning does not change the correlation structure. However, it reduces the standard deviation of the asset return; other things being equal, $\sigma(R_i)$ is lower; thus, the $\beta_{i,m}$ in equation (6) is lower because the asset's information is learned by the average investor.

As we find in section 2.2, mood swings in either direction induce investors to acquire less value-relevant information. Hence, assets that the average investor under-studies as a result of mood swings are relatively riskier than investors' well-researched assets. If investors' choice to learn information about assets' payoff is not comprehensive, their inadequate learning implies that using the unconditional betas from classical factor models or SDFs to price assets is severely inappropriate. More specifically, when investors do not learn or acquire enough information regarding an asset i's payoff x_i, this insufficient learning increases the asset return risk, such as $\sigma(R_i)$ in equation (6). Consequently, the actual $\sigma(R_i)$ becomes higher to investors. The use of unconditional betas that are assumed to capture full pricing information to risk factors will understate the asset's risk.

For instance, the CAPM beta measures the unconditional relationship between an asset return and the market return. The single-factor CAPM model can easily derive that the market beta conditional on investors' insufficient information caused by mood ($\beta^{nl}_{i,CAPM}$, where nl denotes insufficient learning) should be higher than the unconditional beta ($\beta_{i,CAPM}$).[18] Therefore, $\beta^{nl}_{i,CAPM}$ implies higher risk and a higher expected return $(E(R^e_i)$=$\beta^{nl}_{i,CAPM} * \lambda_{MKT})$ than $\beta_{i,CAPM}$ due to investors' insufficient learning indicating a higher $\sigma(R_i)$. By the same token, one can map the logic to multi-factor models to analyze the covariance structure $Cov(F,R^e_i)$ in equation (9) for Fama-French three or five factors and q-factor model by Hou et al. (2015). This rationale can be applied to the discussion of hundreds of "innovative" risk factors in the empirical asset pricing studies to explain traditional models' failure in asset pricing. In other words, the beta conditional on insufficient learning should be higher than the unconditional beta. Specifically, we can find that $\beta^{nl}_{i,m} > \beta_{i,m}$ due to the $\sigma^{nl}(R_i) > \sigma(R_i)$ in equation (6). Working with factor models, we can find that $\sigma(R^e_i)$ in the structure of $Cov(F,R^e_i)$ share similar characteristics based on investors' insufficient learning about the asset i ($Cov^{nl}(F,R^e) > Cov(F,R^e)$). Therefore, when pricing the assets that are marked by a severe lack of investors' learning with the factor model, the betas investors should use are β^{nl} rather than the unconditional betas β in equation (9), in fact, $\beta^{nl} > \beta$.

The classical assumption of *Homo economicus* indicates that economic agents do not make sub-optimal economic decisions or behaviors driven by psychological shortcomings. Nevertheless, as we find in the data, the empirical evidence shows when agents have mood swings, investors are less likely to acquire earnings-relevant information to learn about the company's performance than they are in a sober state. Therefore, we argue that mood has a significantly negative impact on investors' fundamental learning about assets. In other words, as investors become moody, they mistakenly rely on their feelings as part of pricing information and acquire less fundamental information to incorporate into the valuation of assets. This mood-driven learning deficit causes the asset risk ($\sigma^{nl}(R_i)$) to be higher than the scenario in which investors do not suffer the mood effect and rationally learn information about the risky asset ($\sigma(R_i)$). As we mentioned above, when using unconditional pricing models without considering this insufficient learning, the betas in classical pricing models are underestimated and do not fully capture the additional risk that is contributed by the mood in our study. Intuitively, we should expect a positive and significant abnormal return (α in equation (6)) can be found by using the underestimated beta models (CAPM, Fama-French, etc.). Furthermore, we do not expect all assets to be significantly affected by the mood effect.[19] As a result, only assets that are sensitive to this mood effect, causing investors'

[18]Note that the $\beta_{i,CAPM} = \frac{Cov(R_i,R_{MKT})}{\sigma^2(R_{MKT})} = \rho_{i,MKT}\frac{\sigma(R_i)}{\sigma(R_{MKT})}$. As investors do not learn enough information about the asset payoff, $\sigma(R_i)$ is higher conditional on the insufficient learning effect. Consequently, the $\beta^{nl}_{i,CAPM}$ is higher than the unconditional CAPM beta as the $\sigma(R_i)$ increases.

[19]It is hard to believe that all assets suffer the mood-biased effect on information acquisition. For example, large

insufficient learning, are riskier to hold in an investment portfolio. To hold these mood-sensitive assets, traders require higher expected returns to compensate for the added risk by mood. In sum, the theoretical analysis of the mood effect causing investors' insufficient learning about assets yields the following empirical prediction:

Hypothesis 2: Assets more sensitive to mood swings are riskier due to investors' failure to acquire sufficient information about them; a higher expected return is required for investors to hold these mood-sensitive assets.

2.4 Data for Empirical Tests

We conduct empirical tests in line with the theoretical hypothesis using the Twitter mood index that exhibits three key characteristics. First, there are visible outliers, most of which last a very short time. These outliers fall into two categories. Some outliers repeat and are predictable: Holidays and celebrations such as Christmas, Easter, Thanksgiving and Mother's Day naturally score very highly on the Twitter Mood Index. These are not relevant for our study, as U.S. markets are closed on these days. Other predictable holidays and celebrations - Valentine's Day, for example - do coincide with trading and we deal with this in our analysis. Other outlier days are due to unpredictable events. The saddest day in Figure 1 is the Sandy Hook Elementary School shooting on December 14th, 2012. Most, but not all, unpredictable outliers are sad days.

A second pattern not easily visible in Figure 1 is the day-of-the-week pattern. Fridays (and Saturdays) are systematically more happy than Mondays.[20] These are the day-of-the-week patterns explored in Hirshleifer et al. (2020). Since they too are predictable and seasonal we remove them from our analysis. Hirshleifer et al. (2020) also use monthly seasonalities, arguing that the months of January and March are positive mood months while Septembers and Octobers are negative mood months. This is less apparent in the Twitter series as a result of its third feature, the slow oscillation about the mean. Twitter mood peaked in early 2010 and late 2015, with troughs in late 2012/early 2013 and mid-2017. Once the effect of regular holidays is excluded, there is very little monthly seasonality in Twitter mood: January 2013 was a lot less happy than September 2015.

More importantly, Figure 1 shows that Twitter mood and Baker-Wurgler (B&W) sentiment series follow very different paths over their common interval. Indeed, key swings in the two series have often been in opposite directions.[21] Therefore, the sensitivity measure on the variation of the

firms or companies in an industry that is transparent or easy to analyze for investors. In other words, learning about these firms is, to some extent, costless.

[20] Abraham and Ikenberry (1994) study the trading behaviors of individual investors resulting in a 'weekend effect' from a relationship between the Friday stock return and the upcoming Monday return. Additionally, Birru (2018) finds that the long-short returns anomaly is related to the mood that has the day of the week patterns.

[21] The correlation between the B&W sentiment index and our Twitter mood index is around -0.5.

Twitter mood index, mood beta, is different from the measures based on the B&W sentiment index such as sentiment beta reported in studies by Glushkov (2006) and Hirshleifer et al. (2020).

We use the daily change of the Twitter mood index from September 2008 to December 2016 as a proxy for public mood changes. The raw mood data are at a daily frequency, and as already noted have a strong day-of-the-week effect. Furthermore, there are two festive days on which the market trades and which are predictably happy; Valentine's Day has an average mood score of 6.15 and Easter Monday has an average mood score 6.08, both of which are higher than the sample average mood score of 6.02. Including these two days in the calculations would lead to predictable outliers in the mood measure. Instead, the mood we wish to analyze is orthogonalized to any foreseeable patterns. Therefore, to obtain a completely exogenous factor for mood, we regress the change of daily mood from Twitter on dummies for days-of-the-week, Valentine's Day and the Easter holiday.

$$\Delta TwitterMoodScore_t = b_1 D_{Monday} + \ldots b_7 D_{Sunday} + b_8 D_{Valentine} + b_9 D_{Easter} + \varepsilon_t \tag{10}$$

$$\Delta Mood_t = \varepsilon_t$$

Table 3 shows that Monday and Sunday have a strong negative impact on the change of mood. This result is consistent with the perception that people are generally less happy on Mondays, with an associated decrease on Sunday since it precedes a Monday.[22] Not surprisingly, Thursday and Friday have strong positive significance. We take the regression residuals as a proxy for mood factor which is orthogonalized to the day-of-the-week effect and predicted holiday effects for the following cross-section stock return study.

Daily stock returns are taken from the Center for Research in Security Prices (CRSP) and financial fundamentals data from the CRSP/Compustat merged database.[23] We retain all U.S.-based common stocks with share code (SHRCD) value 10 or 11 listed on the NYSE, AMEX, and NASDAQ with exchange code (EXCHCD) 1 or 31, 2 or 32 and 3 or 33 respectively. While we expect that large sensitivities to mood are likely in smaller stocks, we do not want our results to be driven by penny stocks. Therefore we exclude stocks with prices below \$2.50 or a market capital less than the 0.5$^{\text{th}}$ percentile. There are 4962 stocks in our sample analysis.

[22] See studies by Croft and Walker (2001), Areni and Burger (2008), Ryan et al. (2010) and Stone et al. (2012).

[23] We merge CRSP returns data and the CRSP/Compustat merged database by PERMNO and LPERMNO. If there is no match between PERMNO and LPERMNO, we use the tickers and merge LPERMNO data based on tickers in the two databases. Detailed information about merging data and the definition of financial fundamentals can be found in Appendix A.

Table 3: Day of The Week and Festival Effects

This table is regressions of the daily percentage change of the Twitter mood index on dummy variables for each day of the week, weekends and predictable festivals (Valentine's day and Easter Monday). The sample period of Twitter mood index in our study is from September 2008 to December 2016. $\Delta TwitterMoodScore_t = \frac{TwitterMoodScore_t - TwitterMoodScore_{t-1}}{TwitterMoodScore_{t-1}}$, where $TwitterMoodScore_t$ is the raw measure from Twitter mood index.

	Monday	*Tuesday*	*Wednesday*	*Thursday*	*Friday*	*Saturday*	*Sunday*	*Valentines*	*Easter*	$\bar{R}^2$
Model 1										
$Coeff$	-0.0026	-0.0002	0.0003	0.0012	0.0031					0.11
t_{Coeff}	-10.70	-1.48	1.80	5.53	13.24					
Model 2										
$Coeff$	-0.0026	-0.0002	0.0003	0.0012	0.0031	0.0003	-0.0020			0.13
t_{Coeff}	-10.70	-1.48	1.80	5.53	13.24	1.25	-10.11			
Model 3										
$Coeff$	-0.0026	-0.0003	0.0003	0.0011	0.0031	0.0002	-0.0021	0.0155		0.16
t_{Coeff}	-10.96	-1.97	1.80	5.48	13.35	0.94	-10.65	12.32		
Model 4										
$Coeff$	-0.0026	-0.0003	0.0003	0.0011	0.0031	0.0002	-0.0022	0.0155	0.0079	0.17
t_{Coeff}	-10.96	-1.97	1.80	5.48	13.35	0.94	-11.64	12.42	8.33	
Average Mood Score	6.015	6.014	6.016	6.023	6.042	6.043	6.031			

3 Mood Beta Estimation and Portfolios

3.1 Estimation of Mood Betas

Because not all assets are subject to the mood effect that causes investors' irrationally insufficient learning, we first identify the sensitivity of the stock returns to mood. We use the Fama-French five-factor model (Fama and French, 2015) augmented with Carhart's (1997) momentum factor as the benchmark model in our analysis. The factors are all taken from Kenneth R. French Data Library. Table 4 shows the Pearson correlation among factors and mood variables in our sample period from September 2008 to December 2016. There is no multi-collinearity problem in our time series analysis. The correlation between $\Delta Mood_t$ and the Fama-French factors are around 0.015. The correlation between $\Delta Mood_t$ and the momentum factor is very small and negative.

Table 4: Pearson Correlations among Factors

This table reports Pearson correlations between the change of daily Twitter mood and classical pricing factors that are downloaded from Kenneth R. French Data Library. The sample period in our study is from September 2008 to December 2016. $\Delta Mood_t$ is from equation (10), where the residuals are orthogonalized to the day-of-the-week effect and predicted holiday effects.

	MKT	*SMB*	*HML*	*CMA*	*RMW*	*MOM*	$\Delta Mood_t$
MKT		0.263	0.419	-0.151	-0.457	-0.395	0.013
SMB			0.122	0.009	-0.342	-0.097	0.018
HML				0.232	-0.462	-0.548	0.024
CMA					0.039	0.127	-0.019
RMW						0.290	-0.011
MOM							-0.0003

Following the similar method applied by Bali et al. (2017), for each stock we perform the following regression:

$$R_{i,t} = \alpha_i + \sum_{m=1}^{M} \beta_{i,m} f_{m,t} + b_i \Delta Mood_t + \varepsilon_{i,t} \tag{11}$$

where $\Delta Mood_t$ is the daily percentage change of the Twitter mood score and $f_{m,t}$ is a vector of pricing factors. The regression coefficient on $\Delta Mood_t$ in equation (11) measures the time series sensitivity of each stock to Twitter mood innovations controlling for the other six benchmark factor effects.[24] This coefficient (b_i) is our mood beta. Intuitively, the larger the absolute value of b_i is, the more sensitive the stock return R_i is affected by mood.

We estimate mood betas from rolling regressions. Following Bali et al. (2016) we use windows of 200 daily observations. Since a naïve 200-observation window would ignore some IPO, merger and acquisition activities in a stock, our regression analysis follows the dynamic data-rolling method proposed by Bali et al. (2016) which takes into account all corporate actions.[25] Figure 2 plots the data distribution of the regression coefficient on $\Delta Mood_t$, which is the mood beta.

3.2 Identification of Mood Impact

We start the analysis by forming portfolios based on mood beta sorts and identifying whether the stocks that are sensitive to mood have high expected returns, as indicated in Hypothesis 2. From the end of December 2009 we sort our stocks based on historical 200-day rolling regression coefficients of $\Delta Mood_t$ into ten portfolios of ascending sensitivity to mood based on breakpoints derived from NYSE stocks, and calculate value-weighted portfolio excess returns for the next month.[26] Portfolio 1 includes stocks with the lowest regression coefficients on $\Delta Mood_t$, that is, stocks that are negatively sensitive to mood. Portfolio 10 includes stocks with the highest regression coefficients on $\Delta Mood_t$. Stocks in portfolio 10 are the most positively sensitive to mood. We then calculate the value-weighted average excess returns for each portfolio over the subsequent month.

The top panel in Table 5 reports the average mood beta for the ten mood portfolios together with their monthly value-weighted excess returns. Standard errors are Newey-West-corrected. By construction, the magnitude of mood factor loading increases monotonically from portfolio 1 to portfolio 10. Stocks in portfolio 1 have negative sensitivity to mood change and an average mood

[24] In robustness checks we also include a lagged value of $\Delta Mood$ in the regression since the Twitter mood score has a strong mean-reverting time-series pattern. The Twitter mood score is measured for all Tweets posted within 24 hours of midnight (00:00 until 23:59). It is not hard to identify mood impact on stocks which are sensitive to mood during trading hours (09:30 to 16:00). However, mood out of trading hours may also affect stock returns. Therefore, part of the mood impact on stock return at day t may come from the mood between hours 16:00 to 23:59 at day $t-1$, so we add the lagged change in mood. Our results are unaffected by this adjustment for lagged Twitter mood changes.

[25] In fact, the regression coefficient of mood is largely insensitive to how stocks are incorporated into the sample.

[26] Our results are not sensitive to the use of all stocks breakpoints.

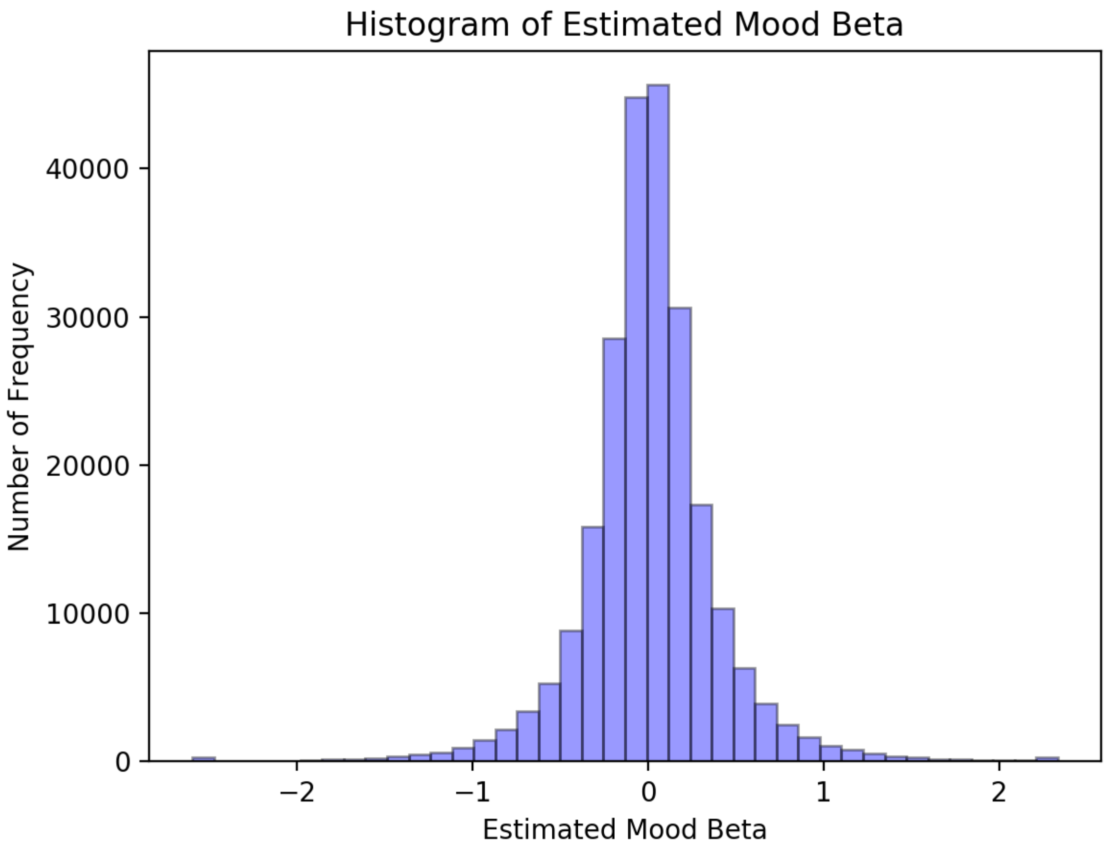

Figure 2: Estimated Mood Betas from equation (11)

beta of -0.58. The mood beta is markedly higher for portfolio 2 at just -0.23. Stocks in portfolio 10 have positive sensitivity to mood change, with a mood beta of 0.61. Again, moving from portfolio 10 to portfolio 9 sees a large decrease in mood beta to 0.26. While mood betas rise as we move from portfolio 2 to portfolio 9, the magnitudes of the changes are much smaller than when considering the extreme portfolios. Extreme sensitivity to mood (either positive or negative) is concentrated in a relatively small number of stocks that lie in portfolios 1 and 10.

The negative and positive mood betas are consistent with the study by Goetzmann et al. (2015), who argue that stock returns have co-movement patterns during optimistic and pessimistic mood days, which are proxied by weather in their study. This understanding can be also interpreted with the argument of Hirshleifer et al. (2020), who propose that the the correlation between stock returns and seasonal patterns is due to the effect of mood. In contrast to their seasonality argument of mood-congruent stock returns, we streamline the congruence of stock returns into two major categories: the tendency of positive and negative mood days measured and the change in the Twitter mood index. More specifically, the extreme sensitivity to public mood, either optimistic

Table 5: Factor Regression for Monthly Excess Returns of Mood Sorted Portfolios, NYSE Breakpoints,Value-weighted Returns (12/2009-12/2016), 84 Months

For each month from 2010 to 2016, we sort stocks into 10 portfolios based on the factor loading (NYSE breakpoints) of $\Delta Mood_t$ in model (1.11) and calculate the value-weighted monthly excess return for each portfolio. We report the average regression coefficient of $\Delta Mood_t$ on each portfolio and the average portfolio monthly excess returns. The negative slop coefficients indicate stocks which are negatively sensitive to mood change. The positive slop coefficients indicate stocks which are positively sensitive to mood change. Standard errors are subject to Newey-West correction. H/L is the high-low portfolio which is half of a portfolio to long both mood-affected stocks (portfolios 1 and 10) and to short mood-insensitive stocks (portfolios 5 and 6). Market factor (MKT), size factor (SMB), value factor (HML), investment factor (CMA), profitability factor (RMW), momentum factor (MOM), short-term reversal factor (ST) and long-term reversal factor (LT) data are downloaded from the Kenneth R. French Data Library. We report the mean monthly alphas across 10 portfolios based on the Fama-French five-factor regressions, ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA}CMA_t + \beta_{RMW}RMW_t + \varepsilon_{p,t}$), the Carhart momentum factor regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \varepsilon_{p,t}$) and the full behavioral factors regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \beta_{ST,p}ST_t + \beta_{LT,p}LT_t + \varepsilon_{p,t}$).

	1	2	3	4	5	6	7	8	9	10	H/L
β_{Mood}	-0.58	-0.23	-0.14	-0.07	-0.02	0.03	0.09	0.16	0.26	0.61	
$Mean$	1.66%	0.97%	0.86%	1.18%	1.13%	1.14%	1.05%	1.07%	1.16%	1.65%	0.52%
t_{Mean}	3.94	2.62	2.19	4.06	3.51	3.17	3.15	2.97	2.76	4.02	3.47
$Fama-French$											
α	0.48	-0.21	-0.33	0.06	0.04	0.11	-0.05	0.01	-0.10	0.47	0.40
t_α	3.25	-1.33	-1.20	0.51	0.31	1.00	-0.41	0.09	-0.54	3.73	2.74
MKT	1.09	1.07	1.04	1.01	1.02	0.96	0.99	0.94	1.12	1.05	0.08
t_{MKT}	22.32	34.93	16.17	38.46	25.76	33.50	34.34	19.35	23.40	29.20	1.68
SMB	0.27	0.10	0.06	0.03	-0.18	-0.10	-0.07	0.02	0.10	0.17	0.36
t_{SMB}	4.15	1.38	0.74	0.60	-3.51	-2.36	-0.91	0.35	1.00	2.20	4.21
HML	0.04	-0.01	0.05	0.01	0.01	-0.03	-0.07	0.04	0.25	0.05	0.06
t_{HML}	0.36	-0.09	0.41	0.11	0.10	-0.48	-0.78	0.38	1.56	0.57	0.59
CMA	-0.19	0.01	0.16	-0.05	-0.02	0.07	0.18	0.14	0.05	0.06	-0.09
t_{CMA}	-1.29	0.04	1.46	-0.59	-0.14	0.60	1.82	1.01	0.27	0.52	-0.63
RMW	-0.14	-0.07	0.15	0.28	0.08	-0.11	0.05	0.02	-0.03	0.05	-0.03
$\bar{R}^2$	0.91	0.90	0.84	0.94	0.91	0.92	0.91	0.88	0.86	0.88	0.32
$CARH$											
α	0.50	-0.21	-0.38	0.06	0.02	0.09	-0.07	0.02	-0.06	0.51	0.45
t_α	3.86	-1.41	-1.40	0.50	0.18	0.84	-0.49	0.17	-0.30	4.05	4.22
MKT	1.08	1.07	1.05	1.01	1.02	0.96	0.99	0.93	1.12	1.04	0.07
t_{MKT}	24.29	35.90	17.70	38.19	26.91	41.19	33.18	18.35	22.67	30.40	1.88
SMB	0.28	0.10	0.04	0.02	-0.19	-0.11	-0.08	0.02	0.11	0.19	0.38
t_{SMB}	4.29	1.32	0.55	0.59	-3.58	-2.68	-1.05	0.43	1.14	2.68	4.88
HML	-0.01	-0.01	0.14	0.01	0.05	0.02	-0.05	0.02	0.17	-0.03	-0.06
t_{HML}	-0.10	-0.12	0.98	0.12	0.76	0.30	-0.43	0.18	1.14	-0.40	-0.55
CMA	-0.15	0.01	0.10	-0.05	-0.04	0.04	0.17	0.15	0.10	0.11	-0.02
t_{CMA}	-1.09	0.05	0.88	-0.57	-0.32	0.33	1.73	1.03	0.55	0.89	-0.13
RMW	-0.13	-0.06	0.12	0.27	0.07	-0.12	0.05	0.02	-0.01	0.07	0.00
t_{RMW}	-1.45	-0.87	0.95	3.57	1.13	-1.52	0.50	0.23	-0.09	0.65	-0.04
MOM	-0.09	0.00	0.15	0.00	0.07	0.08	0.05	-0.03	-0.12	-0.13	-0.18
t_{MOM}	-2.20	-0.03	1.76	0.03	1.78	2.71	1.20	-0.41	-1.50	-1.78	-3.34
$\bar{R}^2$	0.91	0.90	0.85	0.93	0.91	0.92	0.91	0.88	0.87	0.89	0.41
$CARH\&ST\&LT$											
α	0.54	-0.20	-0.39	0.03	0.02	0.08	-0.02	0.05	0.02	0.55	0.50
t_α	3.82	-1.25	-1.52	0.27	0.18	0.62	-0.13	0.35	0.08	4.18	3.69
MKT	1.04	1.07	1.03	1.02	1.01	0.95	0.98	0.93	1.04	1.04	0.06
t_{MKT}	17.39	28.65	26.11	30.86	30.49	37.72	27.46	17.05	17.50	24.21	1.18
SMB	0.27	0.10	0.04	0.03	-0.19	-0.11	-0.08	0.02	0.10	0.19	0.38
t_{SMB}	4.84	1.26	0.54	0.67	-3.51	-2.87	-1.22	0.39	1.12	2.90	5.51
HML	-0.06	-0.01	0.14	0.04	0.05	0.03	-0.10	-0.01	0.07	-0.08	-0.11
t_{HML}	-0.48	-0.17	1.32	0.52	0.62	0.41	-1.22	-0.05	0.61	-0.98	-1.19
CMA	-0.16	0.01	0.15	-0.01	-0.03	0.08	0.08	0.10	0.04	0.03	-0.09
t_{CMA}	-1.36	0.06	1.06	-0.05	-0.19	0.74	0.64	0.58	0.29	0.19	-0.61
RMW	-0.07	-0.06	0.12	0.23	0.07	-0.13	0.11	0.06	0.11	0.12	0.06
t_{RMW}	-0.65	-0.78	0.72	2.43	1.16	-1.42	1.13	0.58	0.72	0.93	0.49
MOM	-0.06	0.00	0.17	0.00	0.07	0.09	0.04	-0.04	-0.09	-0.15	-0.19
t_{MOM}	-1.41	0.01	1.68	-0.01	1.72	3.00	0.96	-0.50	-1.11	-1.98	-3.32
ST	0.11	0.01	0.11	0.01	0.04	0.09	-0.09	-0.06	0.13	-0.10	-0.06
t_{ST}	1.83	0.27	0.95	0.27	0.59	2.40	-1.76	-0.94	2.06	-1.35	-0.94
LT	0.09	0.01	-0.04	-0.09	-0.01	-0.05	0.17	0.09	0.22	0.15	0.15
t_{LT}	0.78	0.07	-0.35	-1.13	-0.10	-0.72	1.54	0.90	1.82	1.07	1.22
$\bar{R}^2$	0.92	0.90	0.84	0.93	0.91	0.92	0.92	0.88	0.87	0.89	0.41

or pessimistic, arises from the congruence of investors' trading behaviors which are affected by the unconscious incorporation of mood as information rather than learning sufficient fundamental information about assets into their decision-making. On the one hand, for example, the positive mood sensitivity of stocks is caused by investors with optimistic mood bias tending to overprice stocks and place more irrational buying orders (Kliger and Levy, 2003; Ifcher and Zarghamee, 2011; Goetzmann et al., 2015; Kaustia and Rantapuska, 2016). On the other hand, for instance, negative mood sensitivity is caused by investors with pessimistic mood bias perceiving more risk and becoming more risk-averse. This leads to lower expected firm earnings and gives rise to a requirement for higher returns as compensation for the excess perceived risk biased by the mood (Shu, 2010; Jiang et al., 2019).

The ten portfolio returns in the top panel of Table 5 have an approximate U-shape (see Figure 3) as portfolio 1 and portfolio 10 — the most mood-sensitive portfolios — generate higher monthly excess returns (1.66% and 1.65% respectively) than less mood-sensitive portfolios. In fact, from portfolio 5 to portfolio 9, the monthly excess returns are all around 1.1%. We form a high-low portfolio by taking a long position in both portfolio 1 (the negative-mood portfolio) and portfolio 10 (the positive-mood portfolio) at each investment period, and short positions in the mood-insensitive portfolios 5 and 6. This high-low strategy generates an excess return of 0.52% per month in-sample which is both statistically and economically significant (with a t-stat of 3.47). The mood portfolio results illustrate that stocks that are sensitive to mood in the financial market earn a higher monthly return than stocks which are less sensitive or are insensitive to mood. The mood beta from equation (11) already controls for the Fama-French five factors, and the Carhart momentum factor. Therefore, we propose that returns generated by our mood investment strategy can be considered as a possible new factor which cannot be explained by the most commonly used factors. These results are consistent with the study by Hirshleifer et al. (2020), who propose that the bias from mood contributes to the factor that is liable to be affected. Nevertheless, we focus on testing the argument that investors demand a risk premium (in the form of higher expected returns) as compensation for holding mood-sensitive assets about which they acquire less information instead of arguing for a particular misvaluation on factors within a risky asset payoff as in Hirshleifer et al. (2020).

The second panel in Table 5 reports the results of a time series regression of the monthly value-weighted returns on each of the ten portfolios sorted by mood beta and of the high-low strategy on Fama-French's five factors. Portfolio 1 generates a positive alpha of 0.48% per month while Portfolio 10 generates a positive alpha of 0.47% per month. Both are more than three standard errors from zero. Conversely, portfolios 2-9 generate smaller and statistically insignificant alphas of between -0.3 - +0.11% per month. Even though the adjusted-R^2 figures are around 90%, the significant positive alphas in the most mood-sensitive portfolios illustrate that the Fama-French

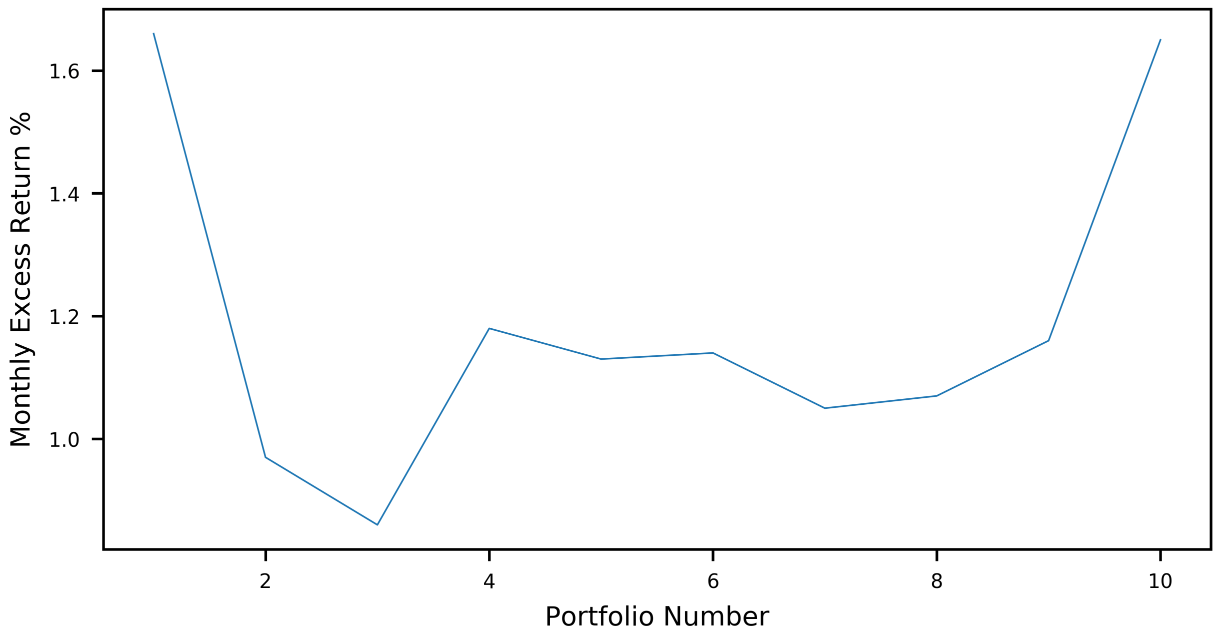

Figure 3: Portfolio Return Sorted by Mood Betas

five-factor model cannot capture the pricing information based on the mood effect. More importantly, the high-low strategy yields a statistically significant alpha of 0.40% per month.

In the third panel we add the momentum factor. The results are quite similar to the Fama-French five-factor model analysis. The momentum factor is significantly negative for portfolio 1 and for the high-low strategy. The sign of the loading on momentum for the high-low portfolio is unexpected, but the large residual alphas for the most mood-sensitive portfolios and long-short strategy suggests that the momentum factor cannot capture the mood effect on cross-section stock returns.[27]

The mood effect investigated in this study is clearly a behavioral factor, and we believe it is not captured by other documented behavioral patterns such as short- and long-run reversal effects in the stock market. The bottom panel of Table 5 adds the short- and long-term reversal factors to the regression. These reversal factors only serve to increase the magnitudes of the significant positive alphas in portfolios 1 and 10, and of the long-short strategy. Our key mood results are robust to controlling for the main behavioral factors in the literature.

More importantly, the empirical results are consistent with our theoretical results. The positive abnormal return in moody stocks (portfolios 1 and 10) is due to the unconditional betas from clas-

[27]The negative momentum factor exposures for the high-low portfolio can be explained as a consequence of psychological bias leading to irrational decisions of investment strategy and security selection. See the discussions by Shefrin and Statman (1985), Bikhchandani et al. (1992),Barberis et al. (1998), Daniel and Titman (2006) and Frazzini (2006).

sical asset pricing understating the risk of stocks sensitive to the mood effect inducing insufficient fundamental information acquisition.

3.3 Financial Characteristics

The empirical results concerning the excess returns of mood-sensitive portfolios conditional on the Fama-French and Carhart factors demonstrate that mood does indeed have a significant impact on cross-sectional stock returns. Clearly, not all stocks are affected by mood and it is interesting to characterize the kinds of firms that are more likely to be mood-sensitive.

The firm-level data are drawn from the merged CRSP-Compustat database. Size is measured by the market value of equity. The book-to-market ratio is calculated using the method of Fama and French (1992). To identify whether mood-sensitive stocks pay less in dividends we calculate both the dividend yield (D/Y) and the percentage of companies that are paying dividends. For the measurement of profitability, we consider operating cash flow (OCF), earnings per share (EPS), return on assets (ROA), EBITA/Assets and sales revenue. We measure the age of firms by calculating the total number of years for which data are available in the CRSP database back to 1926. To measure financing activity, we consider book leverage and external financing (EF) scaled by asset growth. We also consider the tangible asset ratio (PPE/Asset) and research and development (R&D) expenditure in the analysis. Finally, idiosyncratic risk is measured by taking the RSE of residuals from the Carhart model. Detailed information about the definition of financial fundamentals is available in Appendix A.

Table 6 reports the financial characteristics of 10 value-weighted portfolios formed by sorting mood factor loadings. While mood is a specific and distinct type of sentiment, we expect there to be commonality between our findings and those of, for example, Baker and Wurgler (2006) and Glushkov (2006) and to a large extent this is indeed the case. In summary, we find that mood-sensitive stocks, regardless of whether their mood-sensitivity is negative or positive, are small in size, relatively young, pay less in dividends, have more expenditure on R&D, are not profitable, engage in more external financing and have higher levels of idiosyncratic risk. This is in line with previous work identifying more sentiment-sensitive stocks as ”hard-to-value and difficult-to-arbitrage.” More importantly, the evidence of financial characteristics is consistent with the argument made by Bushee and Friedman (2016), who state that stocks with lack of disclosure are more likely to be invested in by noise traders or unsophisticated investors who are in turn more likely to be affected by non-fundamental factors such as mood. Moreover, mood-sensitive stocks, to some extent, shed light on the study by De Long et al. (1990), who argue that noise traders add risk and contribute to mispricing induced by sentiment. Therefore, mood can be thought of as another trigger that induces noise trading. In contrast to expectations and previous findings, however, there is no clear evidence that mood-sensitive stocks have different book-to-market or tangible asset ratios or different asset growth rates than less mood-sensitive stocks. Of course, given that our results show that mood-sensitive stocks to offer higher average returns than mood-insensitive stocks, while Glushkov (2006) finds the complete opposite for sentiment-sensitive stocks, our approach to the selection of stocks is different to those adopted by previous analyses. We now briefly discuss the defining characteristics of mood-sensitive stocks.

Both negative (portfolio 1) and positive (portfolio 10) mood stocks tend to have lower market values. As the sensitivity to mood decreases, firm size increases. Portfolios 5 and 6 (mood-insensitive firms) have the largest market capital, and are on average about 3.5 times the size of firms in either portfolio 1 or 10. Consistent with these findings, Lee et al. (1991) argue that individual investors who are more responsive to sentiment shifts have significant impact on smaller stocks. The size characteristic in mood-sensitive stocks found in our sample is also consistent with the studies of Baker and Wurgler (2006) and Glushkov (2006).

The B/M ratios are slightly higher for mood-sensitive stocks; however, the differences are not statistically significant. This contrasts with Baker and Wurgler's (2006) study which finds that firms with extreme values for B/M are more subject to the impact of investor sentiment but is in line with our findings in Table 5 that the value factor is not relevant to the explanation of mood-beta sorted portfolio returns.

The dividend yield is around 0.2 per share in portfolios 1 and 10, much less than the 0.5 per share found in portfolios 5 and 6. Similarly, only 33% of mood-sensitive firms pay dividends, much fewer than the 57% of mood-insensitive firms paying dividends. Again, this is consistent with the findings of Baker and Wurgler (2006) and Glushkov (2006), who show that non-dividend-paying stocks are more subject to investor-sentiment change.[28]

Operating cash flow, earnings per share, return on assets, company age and sales revenue each present an inverted U-shape, showing that mood-sensitive stocks are less profitable. These differences can be large: mood-insensitive stocks generate more than three-times the operating cash flow of mood-sensitive stocks and return on assets is around 3% across portfolios in low mood sensitivity-portfolios but highly negative in portfolios 1 and 10. While age is related to mood-sensitivity, moody stocks are past their early teenage years. The average moody stock is 21-22 years old, younger than the the mood-insensitive stocks, which are on average 27 years old, but they can hardly be characterized as "new" firms.

Three variables display clear U-shaped relationships with mood betas. Research and development expenditure (R&D), external financing (EF) and idiosyncratic risk are all high for mood-sensitive (moody) stocks, and low for mood-insensitive (sober) stocks.

[28]Chung et al. (2012) find that sentiment impact is more significant on non-dividend-paying stocks during economic expansion states.

Table 6: Mean of Statistics on Financial Characteristics of 10 Mood Portfolios

We calculate the average value of financial variables for each portfolio across our sample period. Mkt. Cap is the total market value of equity. B/M is the book value of equity over market value of equity. D/Y is dividend paid per share. Div. is the probability of firms paying a dividend. OCF is operating cash flow. EPS is earnings per share. ROA is return on assets calculated as net income over total assets. EBITDA/Assets is calculated as earnings before interest over total assets. Lever. is total debt over the book value of total assets. PPE/Assets is the value of property, plant and equipment divided by the book value of total assets. R&D is research and development expenditure divided by book value of total assets. Revenue is total sales revenue. Asset growth is the percentage change of total assets between two fiscal years. EF is external financing calculated as difference between asset growth and percentage change of retained earnings. RE is the level of retained earnings between year t to year $t+1$. Age is measured as the date from which data is first available in the database up to December 2016. Risk is idiosyncratic risk measured as the RSE of residuals from the Carhart pricing model for each stock (see detailed information in Appendix A).

Port.	Mkt.Cap	B/M	D/Y	Div.	OCF	EPS	ROA	EBITDA/Assets	Lever.	PPE/Assets	R&D	Revenue	Asset growth	EF	RE	Age	Δ RE	Risk
1	2239.18	0.71	0.22	0.33	205.71	0.39	-0.05	0.03	0.17	0.40	0.06	1905.99	0.65	0.74	303.84	21.33	-0.08	0.029
2	3673.01	0.67	0.32	0.46	356.53	1.04	0.01	0.08	0.17	0.40	0.04	3356.67	0.72	0.73	776.54	24.02	-0.06	0.021
3	6816.13	0.64	0.44	0.54	624.73	1.39	0.02	0.09	0.16	0.40	0.03	5231.95	0.77	0.75	2019.37	26.02	-0.07	0.019
4	8280.46	0.63	0.50	0.57	754.20	1.54	0.03	0.10	0.16	0.40	0.03	6122.44	0.61	0.57	2550.52	26.97	-0.01	0.018
5	8546.49	0.63	0.52	0.57	806.64	1.61	0.03	0.10	0.17	0.41	0.03	6305.89	0.69	0.61	2341.86	27.32	-0.02	0.017
6	8203.45	0.64	0.51	0.57	811.12	1.62	0.03	0.10	0.17	0.42	0.03	6191.50	0.65	0.58	2470.35	27.37	0.04	0.017
7	6799.51	0.65	0.48	0.54	700.14	1.52	0.03	0.10	0.16	0.43	0.03	5511.20	0.60	0.53	2054.69	26.81	0.04	0.018
8	5486.10	0.65	0.41	0.50	576.66	1.35	0.02	0.09	0.17	0.43	0.03	4639.71	0.64	0.56	1545.79	26.07	0.06	0.019
9	3520.94	0.69	0.29	0.40	367.00	0.91	0.00	0.08	0.17	0.43	0.04	3048.98	0.75	0.70	849.98	23.55	0.05	0.023
10	2117.83	0.74	0.22	0.33	241.61	0.36	-0.04	0.04	0.17	0.44	0.06	1981.35	0.71	0.73	364.40	21.83	-0.01	0.030

In general, these U-shaped patterns are consistent with the findings of the sentiment-sensitivity literature. For example, the relation between mood and idiosyncratic risk is noted by Glushkov (2006); however, the intuition behind this specific empirical finding differs. Merton et al. (1987) proposed that when investors do not diversify their portfolio, expected return and idiosyncratic risk have a positive relation. More specifically, De Long et al. (1990) and Lee et al. (1991) state that if noise traders do not trade randomly across assets, the risk created by noise traders cannot be mitigated with diversification. In equilibrium, risk from stochastic investor sentiment will be priced accordingly. The high idiosyncratic risk in our moody stocks illustrates that unsophisticated investors are trading stocks subject to mood swings and that these stocks entail higher firm-specific risk which is not priced in by the traditional factor asset pricing model. While it is outside the scope of this study to consider the causal link between idiosyncratic risk and investor mood, we believe there is the possibility that part of the idiosyncratic risk derives from the trading activities of noise or mood traders (traders who are more likely to be affected by their mood). Hence, a more comprehensive study on this topic is an opportunity for future research.

4 Mood Factor Construction and Pricing Power

As stated in section 2.3, when mood causes investors to acquire less information about risky assets, the betas should be conditional on this effect. Using unconditional pricing models will lead to positive abnormal returns, which can be considered risk premium to compensate for holding moody stocks. Once the additional risk induced by mood is controlled in the model, we should expect a significant reduction in the abnormal returns as the moody stocks' information about investors' insufficient learning triggered by mood is incorporated into the pricing. Therefore, we follow classical empirical asset pricing studies to conduct risk factor testing to verify our argument about the mood effect.

Before constructing the mood-mimicking portfolio, we re-conduct the factor analysis above based on absolute mood betas; that is, we use NYSE breakpoints of the absolute values of mood beta from equation (11) to populate 10 mood portfolios. Portfolio 10 contains the stocks most sensitive to public mood, regardless of whether the sensitivity is negative or positive. Portfolio 1 contains the stocks least sensitive to changes in mood. The high-low portfolio returns represent high mood stocks (portfolio 10) minus low mood stocks (portfolio 1). The top panel in Table 7 reports value-weighted portfolio excess returns and average regression coefficients of the mood variable for each portfolio. The high mood portfolio earns the highest excess return: about 1.52% per month with high statistical significance. Figure 4 shows that there is an increasing pattern of monthly excess returns from the low mood sensitivity portfolio 1 to the high mood sensitivity portfolio 10. The high-low mood portfolio generates significantly positive excess returns of 0.53%

Table 7: Factor Regression for Monthly Excess Returns of Absolute Mood Beta Sorted Portfolios, NYSE Breakpoints,Value-weighted Returns (12/2009-12/2016), 84 Months

For each month from 2010 to 2016, we sort stocks into 10 portfolios based on the absolute factor loading (NYSE breakpoints) of $\Delta Mood_t$ in model (1.11) and calculate the value-weighted monthly excess return for each portfolio. Standard errors are subject to Newey-West correction. H/L is the high-low portfolio which is to long mood-affected stocks (portfolio 10) and to short mood-insensitive stocks (portfolio 1). Market factor (MKT), size factor (SMB), value factor (HML), investment factor (CMA), profitability factor (RMW), momentum factor (MOM), short-term reversal factor (ST) and long-term reversal factor (LT) are downloaded from the Kenneth R. French Data Library. We report the mean monthly alphas across 10 portfolios based on the Fama-French five-factor regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \varepsilon_{p,t}$), the Carhart momentum factor regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \varepsilon_{p,t}$) and the full behavorial factors regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \beta_{ST,p}ST_t + \beta_{LT,p}LT_t + \varepsilon_{p,t}$).

	1	2	3	4	5	6	7	8	9	10	H/L
$\beta_{\|Mood\|}$	0.01	0.04	0.07	0.10	0.13	0.17	0.21	0.28	0.38	0.68	
$Mean$	0.99%	1.04%	0.96%	1.17%	1.12%	0.82%	1.08%	0.78%	1.19%	1.52%	0.53%
t_{Mean}	2.95	3.22	2.79	3.54	3.06	2.29	2.71	1.65	2.44	4.02	3.16
$Fama-French$											
α	-0.11	-0.06	-0.18	0.10	-0.01	-0.25	-0.13	-0.54	0.03	0.36	0.47
t_α	-0.79	-0.60	-1.16	0.80	-0.08	-1.04	-0.60	-3.28	0.18	3.19	2.49
MKT	1.02	1.00	1.03	0.97	1.01	0.96	1.09	1.18	1.10	1.04	0.02
t_{MKT}	23.57	30.40	23.03	28.97	31.24	14.45	24.71	22.05	26.15	25.74	0.30
SMB	-0.16	-0.07	-0.19	-0.02	0.00	0.01	0.01	0.19	0.31	0.16	0.32
t_{SMB}	-3.09	-0.83	-1.76	-0.27	-0.01	0.14	0.16	2.37	2.43	2.49	3.81
HML	0.04	-0.08	0.02	-0.08	-0.01	0.05	0.17	0.05	0.14	-0.04	-0.08
t_{HML}	0.65	-1.25	0.22	-0.76	-0.10	0.31	1.32	0.42	1.47	-0.49	-0.88
CMA	0.03	0.10	0.13	0.14	0.10	0.15	0.02	0.13	-0.39	0.05	0.02
t_{CMA}	0.27	0.64	0.70	1.02	0.82	1.13	0.13	0.68	-2.55	0.53	0.15
RMW	0.01	0.08	0.09	0.06	0.13	0.01	0.05	-0.20	-0.20	0.03	0.02
t_{RMW}	0.15	0.71	0.76	0.69	1.92	0.04	0.44	-1.28	-1.94	0.33	0.19
$\bar{R}^2$	0.92	0.91	0.87	0.89	0.93	0.82	0.87	0.87	0.88	0.91	0.11
$CARH$											
α	-0.15	-0.07	-0.19	0.11	-0.02	-0.27	-0.12	-0.51	0.08	0.38	0.53
t_α	-1.27	-0.72	-1.17	0.84	-0.18	-1.04	-0.50	-3.14	0.48	3.76	3.69
MKT	1.02	1.00	1.03	0.96	1.01	0.96	1.09	1.17	1.09	1.03	0.01
t_{MKT}	29.54	29.71	22.95	26.64	33.51	14.14	22.61	23.06	23.23	28.40	0.15
SMB	-0.17	-0.07	-0.19	-0.02	0.00	0.00	0.02	0.20	0.34	0.17	0.35
t_{SMB}	-3.77	-0.87	-1.85	-0.21	-0.08	0.05	0.24	2.39	2.99	2.66	4.24
HML	0.12	-0.06	0.05	-0.10	0.01	0.08	0.14	0.00	0.03	-0.08	-0.20
t_{HML}	1.93	-0.84	0.40	-1.03	0.21	0.41	0.96	0.00	0.38	-0.87	-2.12
CMA	-0.02	0.09	0.12	0.16	0.09	0.13	0.04	0.16	-0.32	0.07	0.10
t_{CMA}	-0.21	0.56	0.62	1.14	0.73	0.91	0.25	0.85	-2.20	0.75	0.59
RMW	-0.01	0.07	0.08	0.07	0.12	0.00	0.06	-0.19	-0.18	0.04	0.05
t_{RMW}	-0.14	0.67	0.75	0.76	1.90	-0.01	0.50	-1.20	-1.67	0.44	0.56
MOM	0.13	0.03	0.04	-0.04	0.03	0.05	-0.05	-0.08	-0.18	-0.07	-0.19
t_{MOM}	4.32	0.76	0.92	-0.52	1.56	0.59	-0.76	-1.51	-2.95	-1.49	-3.68
$\bar{R}^2$	0.93	0.91	0.87	0.89	0.93	0.82	0.87	0.87	0.89	0.91	0.20
$CARH\&ST\&LT$											
α	-0.14	-0.14	-0.14	0.10	-0.07	-0.21	-0.07	-0.49	0.13	0.41	0.55
t_α	-1.15	-1.44	-0.75	0.70	-0.63	-0.87	-0.30	-2.96	0.68	3.61	3.21
MKT	0.99	1.03	1.00	0.99	1.04	0.90	1.05	1.12	1.03	1.03	0.04
t_{MKT}	34.41	34.74	16.91	24.58	27.70	19.42	20.64	21.51	20.44	20.51	0.59
SMB	-0.18	-0.07	-0.20	-0.01	0.00	-0.01	0.01	0.19	0.33	0.17	0.35
t_{SMB}	-4.34	-0.83	-2.07	-0.17	0.03	-0.13	0.15	2.32	2.96	2.82	4.90
HML	0.09	0.02	-0.02	-0.08	0.08	0.00	0.08	-0.03	-0.03	-0.12	-0.21
t_{HML}	1.89	0.25	-0.17	-0.84	1.56	-0.01	0.65	-0.24	-0.33	-1.20	-2.03
CMA	0.00	0.20	0.05	0.14	0.16	0.09	-0.01	0.20	-0.36	0.02	0.02
t_{CMA}	0.00	1.20	0.23	1.01	1.60	0.55	-0.05	0.93	-2.46	0.14	0.08
RMW	0.02	-0.03	0.16	0.04	0.04	0.10	0.13	-0.15	-0.10	0.08	0.06
t_{RMW}	0.33	-0.23	1.20	0.42	0.67	0.55	1.13	-0.76	-0.77	0.74	0.49
MOM	0.15	0.04	0.04	-0.05	0.03	0.07	-0.04	-0.05	-0.16	-0.07	-0.22
t_{MOM}	4.01	0.87	1.23	-0.76	1.29	0.74	-0.52	-0.97	-2.74	-1.45	-3.50
ST	0.11	0.08	0.00	-0.09	0.01	0.10	0.04	0.16	0.09	-0.06	-0.18
t_{ST}	1.99	1.51	0.03	-1.26	0.19	0.84	0.68	3.39	1.92	-0.95	-1.94
LT	0.02	-0.24	0.16	-0.02	-0.18	0.18	0.15	0.03	0.14	0.11	0.09
t_{LT}	0.35	-3.57	1.31	-0.19	-3.16	1.36	1.23	0.27	1.01	1.01	0.66
$\bar{R}^2$	0.93	0.91	0.87	0.89	0.93	0.82	0.87	0.88	0.89	0.91	0.23

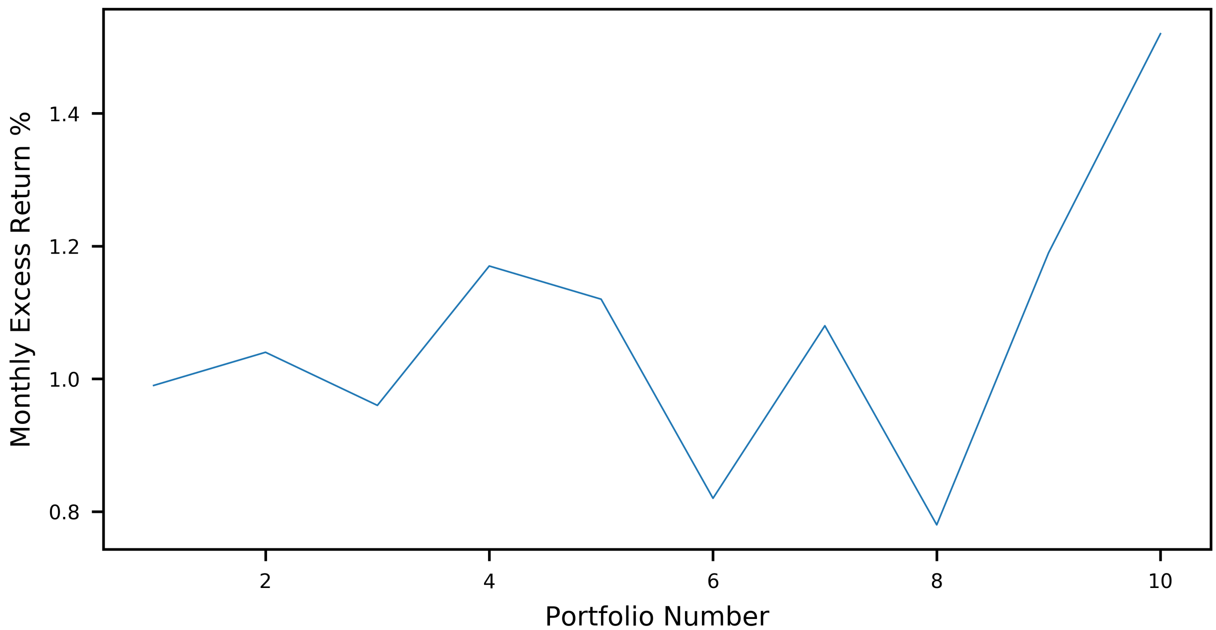

Figure 4: Portfolio Return Sorted by Absolute Mood Betas

per month.

The lower panels in Table 7 report the factor analysis of the monthly excess returns of the ten portfolios sorted by absolute mood betas. The results are entirely comparable to those discussed above based on the raw mood betas. Each of the factor models returns large and statistically significant alphas for the most mood-sensitive portfolio and for the long-short strategy. As before, increasing the complexity of the factor model only serves to increase the mood exposure-driven alphas.

4.1 Mood Factor Portfolio Return

We construct the mimicking mood factor portfolio following the standard method used in empirical asset pricing studies. At the end of each month, we first use the NYSE breakpoints of market capitalization to split stocks into two size portfolios - small and big. Independently, we use the NYSE breakpoints of the absolute value of mood betas estimated from equation (11) to rank stocks into three mood portfolios: low 30%, middle 40% and high 30%. Stocks within the lowest 30$^{\text{th}}$ percentile are the most insensitive to mood; stocks within the highest 30$^{\text{th}}$ percentile are the most sensitive to mood either negatively or positively; and the stocks within the middle 40% have neutral mood sensitivity. We thus form six interacted value-weighted portfolios in respect of size and mood effect: $L/S, N/S, H/S$, $L/B, N/B, H/B$ sorting by size and the absolute value of mood betas independently. The zero-cost mood factor portfolio is constructed by taking the average of long positions in the two mood-sensitive portfolios high 30% ($H/S, H/B$) and the average of short positions in the two mood-insensitive portfolios low 30% ($L/S, L/B$) each month.

Panel A of Table 8 gives the Pearson correlations between our mood factor and Fama-French five factors, Carhart's momentum factor, and both short- and long-term reversal factors. The mood factor is positively correlated with the market, size and reversal factors, and negatively correlated with profitability and momentum. These correlations are comparable in magnitude to those between the other previously identified factors. On average, the risk premium of the mood factor is 0.56% per month and is highly statistically significant, with a t-statistic of 4.7. We run time series

Table 8: Factor Regression Analysis on Mood Factor

At the end of each month, we use NYSE breakpoints of market capitalization to split stocks into two size portfolios - small and big. Independently, we use NYSE breakpoints of the absolute value of mood betas estimated from (1.11) to rank stocks into three mood portfolios: low 30%, middle 40%, and high 30%. Stocks within the lowest 30$^{\text{th}}$ percentile are the most insensitive to mood; stocks within the highest 30$^{\text{th}}$ percentile are the most sensitive to mood either negatively or positively, and the stocks within the middle 40% are neutral to mood-sensitivity. We thus form six interacted value-weighted portfolios respecting size and mood effect: $L/S, N/S, H/S, L/B, N/B, H/B$ sorting on the size and the absolute value of mood betas independently. The zero-cost mood factor portfolio is constructed by taking the average of long positions in in the two mood-sensitive portfolios-high 30% ($H/S, H/B$) and the average of short positions in the two mood-insensitive portfolios-low 30% ($L/S, L/B$) each month. Panel A reports Pearson correlation between the mood portfolio return factor and other factors. Panel B is the regression analysis of mood factor portfolio returns on Fama-French five factors, momentum and short- and long-term reversal factors. Robust t statistics are in parentheses.

Panel A : Pearson Correlation Matrix

	Mood	*MKT*	*SMB*	*HML*	*CMA*	*RMW*	*MOM*	*ST*	*LT*
Mood	1.00	0.47	0.44	0.01	-0.12	-0.38	-0.30	0.33	0.39
MKT		1.00	0.42	0.18	0.09	-0.38	-0.13	0.47	0.53
SMB			1.00	0.23	0.13	-0.42	-0.05	0.24	0.41
HML				1.00	0.62	-0.16	-0.34	0.18	0.62
CMA					1.00	0.07	-0.09	-0.02	0.46
RMW						1.00	0.15	-0.32	-0.49
MOM							1.00	-0.31	-0.28
ST								1.00	0.39
LT									1.00

Panel B: Factor Regression of Mood factor Portfolio Return

	Mean	α	*MKT*	*SMB*	*HML*	*CMA*	*RMW*	*MOM*	*ST*	*LT*	$\bar{R}^2$
r_{Mood}	0.56	0.38	0.16								0.21
	(4.70)	(3.26)	(4.67)								
		0.46	0.11	0.16	-0.02	-0.16	-0.11				0.30
		(4.97)	(3.29)	(3.00)	(-0.20)	(-1.34)	(-1.55)				
		0.50	0.10	0.17	-0.09	-0.11	-0.09	-0.13			0.37
		(6.49)	(3.38)	(3.46)	(-1.02)	(-1.00)	(-1.28)	(-2.25)			
		0.50	0.10	0.17	-0.09	-0.11	-0.09	-0.12	0.01		0.37
		(6.52)	(2.83)	(3.48)	(-1.03)	(-0.96)	(-1.25)	(-2.28)	(0.31)		
		0.56	0.07	0.17	-0.16	-0.18	-0.01	-0.12	0.00	0.19	0.40
		(7.19)	(1.60)	(4.12)	(-2.43)	(-1.57)	(-0.11)	(-2.11)	(-0.10)	(1.91)	

regressions of the mood factor on subsets of the other factors. The most general regression is of the form:

$$r_{Mood_t} = \alpha + \sum_{m=1}^{M} \beta_{i,m} f_{m,t} + \varepsilon_t \tag{12}$$

where $f_{m,t}$ is the vector of pricing factors. Panel B of Table 8 reports the various regression results.

Market, size, momentum and short/long-term reversal factors display consistent explanatory power across alternative specifications, but none of the factor models can fully explain the mood factor. Our mood factor consistently earns highly significant positive alphas, and again alphas increase with the complexity of the factor model. With the most basic CAPM regression, the mood factor has an alpha of 0.38% per month. Once orthogonalized to all factors, the mood factor alpha increases to 0.56% per month, exactly equal to its mean return.

In the rest of the study we test the pricing power of the mood factor orthogonalized to all the other factors.[29] We define the mood factor orthogonalized to the other factors as the alpha from equation (12) plus the regression residuals:

$$r^{\perp}_{Mood_t} = \hat{\alpha} + \varepsilon_t \tag{13}$$

We first regress our absolute mood sensitivity-based portfolio excess returns on Fama-French factors and the orthogonalized mood factor:

$$r_{p,t} - r_f = \alpha_p + \sum_{m=1}^{M} \beta_{i,m} f_{m,t} + \beta_{Mood^{\perp},p} r^{\perp}_{Mood_t} + \varepsilon_{p,t} \tag{14}$$

where $f_{m,t}$ includes Fama-French five factors. We subsequently augment the regression with momentum and with reversal factors. The second panel in Table 9 reports the regression results for equation (14). With the orthogonalized mood factor included in the model, portfolio 10, which contains the most moody stocks, has a loading on the mood factor of 0.79, which is highly significant ($t = 8.12$), and a statistically insignificant alpha of -0.08% per month, down from the 0.36% per month reported in Table 7. Similarly, the high-low portfolio alpha decreases from 0.47% to a statistically insignificant -0.10% per month. The adjusted-R^2 figure for portfolio 10 is slightly improved by adding the mood factor to the model, increasing from 0.91 to 0.96, but the improvement is most noticeable for the long-short strategy, increasing from 0.11 to 0.42. Portfolio 9 loads positively and significantly on the mood factor, while portfolio 1 loads negatively. The adjusted-R^2 for portfolio 1 is barely affected, and the mood factor, while statistically significant, does not appear to add much explanatory power but pushes the alpha back closer to zero.

The other panels in Table 9 report the regression results from more complete factor models, including momentum and reversal factors. Results differ only slightly from those in the second panel. In short, the mood factor is highly significant for portfolio 10 and the high-low portfolio, which earn insignificant alphas once the mood factor is included in the regression. The significant beta on the mood factor captures the additional risk from the mood effect on returns.

[29] In fact, our results are essentially unchanged if we instead use the original mood factor, but given the weak correlations with other factors we use the orthogonalized version to provide the strictest test of its explanatory power.

Table 9: Orthogonalized Mood Factor Pricing of Portfolios Sorted by Absolute Value of Mood Betas

For each month from 2010 to 2016, we sort stocks into 10 portfolios based on the absolute factor loading(NYSE breakpoints) of $\Delta Mood_t$ in model (1.11) and calculate the value-weighted monthly excess return for each portfolio. Standard errors are subject to Newey-West correction. H/L is the high-low portfolio which to long both mood-affected stocks (portfolios 10) and to short mood-insensitive stocks (portfolio 1). Market factor (MKT), size factor (SMB), value factor (HML), investment factor (CMA), profitability factor (RMW), momentum factor (MOM), short-term reversal factor (ST) and long-term reversal factor (LT) are downloaded from the Kenneth R. French Data Library. We report the mean monthly alphas across 10 portfolios based on the Fama-French five-factor regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{Mood^{\perp},p}r^{\perp}_{Mood,t} + \varepsilon_{p,t}$), the Carhart momentum factor regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \beta_{Mood^{\perp},p}r^{\perp}_{Mood,t} + \varepsilon_{p,t}$) and the full behavorial factors regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \beta_{ST,p}ST_t + \beta_{LT,p}LT_t + \beta_{Mood^{\perp},p}r^{\perp}_{Mood,t} + \varepsilon_{p,t}$).

	1	2	3	4	5	6	7	8	9	10	H/L
$\beta_{\lvert Mood \rvert}$	0.01	0.04	0.07	0.10	0.13	0.17	0.21	0.28	0.38	0.68	
$Mean$	0.99%	1.04%	0.96%	1.17%	1.12%	0.82%	1.08%	0.78%	1.19%	1.52%	0.53%
t_{Mean}	2.95	3.22	2.79	3.54	3.06	2.29	2.71	1.65	2.44	4.02	3.16
$Fama-French$											
α	0.02	0.11	0.08	0.12	0.09	-0.14	-0.03	-0.74	-0.25	-0.08	-0.10
t_{α}	0.14	0.88	0.54	0.90	0.74	-0.64	-0.13	-5.39	-1.34	-0.73	-0.46
MKT	1.02	1.00	1.03	0.97	1.01	0.96	1.09	1.18	1.10	1.04	0.02
t_{MKT}	24.71	30.49	25.15	28.64	31.64	14.32	24.74	22.25	27.78	34.80	0.35
SMB	-0.16	-0.07	-0.19	-0.02	0.00	0.01	0.01	0.19	0.31	0.16	0.32
t_{SMB}	-3.00	-0.88	-1.77	-0.27	-0.01	0.14	0.16	2.49	2.38	2.85	3.99
HML	0.04	-0.08	0.02	-0.08	-0.01	0.05	0.17	0.05	0.14	-0.04	-0.08
t_{HML}	0.62	-1.38	0.21	-0.77	-0.10	0.29	1.26	0.41	1.57	-0.58	-0.98
CMA	0.03	0.10	0.13	0.14	0.10	0.15	0.02	0.13	-0.39	0.05	0.02
t_{CMA}	0.30	0.65	0.71	1.03	0.86	1.11	0.12	0.65	-2.65	0.55	0.19
RMW	0.01	0.08	0.09	0.06	0.13	0.01	0.05	-0.20	-0.20	0.03	0.02
t_{RMW}	0.14	0.78	0.75	0.68	2.06	0.04	0.43	-1.27	-1.83	0.42	0.18
$r_{Mood^{\perp}}$	-0.24	-0.30	-0.46	-0.05	-0.17	-0.21	-0.18	0.36	0.51	0.79	1.02
$t_{Mood^{\perp}}$	-2.28	-1.88	-3.33	-0.42	-1.38	-1.66	-1.25	1.26	2.66	8.12	6.69
$\bar{R}^2$	0.92	0.91	0.88	0.89	0.93	0.82	0.87	0.88	0.89	0.94	0.42
$CARH$											
α	-0.02	0.10	0.06	0.13	0.08	-0.15	-0.02	-0.71	-0.20	-0.06	-0.04
t_{α}	-0.13	0.78	0.44	0.91	0.67	-0.67	-0.06	-5.71	-1.06	-0.63	-0.28
MKT	1.02	1.00	1.03	0.96	1.01	0.96	1.09	1.17	1.09	1.03	0.01
t_{MKT}	31.19	30.23	25.40	26.31	33.69	13.96	22.44	23.44	24.57	39.83	0.19
SMB	-0.17	-0.07	-0.19	-0.02	0.00	0.00	0.02	0.20	0.34	0.17	0.35
t_{SMB}	-3.61	-0.93	-1.91	-0.20	-0.08	0.04	0.24	2.49	2.91	2.91	4.27
HML	0.12	-0.06	0.05	-0.10	0.01	0.08	0.14	0.00	0.03	-0.08	-0.20
t_{HML}	1.87	-0.93	0.36	-1.04	0.22	0.40	0.92	0.00	0.44	-1.05	-2.60
CMA	-0.02	0.09	0.12	0.16	0.09	0.13	0.04	0.16	-0.32	0.07	0.10
t_{CMA}	-0.23	0.56	0.61	1.14	0.75	0.90	0.24	0.82	-2.24	0.75	0.68
RMW	-0.01	0.07	0.08	0.07	0.12	0.00	0.06	-0.19	-0.18	0.04	0.05
t_{RMW}	-0.13	0.73	0.76	0.75	2.06	-0.01	0.49	-1.20	-1.63	0.55	0.62
MOM	0.13	0.03	0.04	-0.04	0.03	0.05	-0.05	-0.08	-0.18	-0.07	-0.19
t_{MOM}	4.28	0.88	0.78	-0.53	1.40	0.57	-0.77	-1.87	-3.37	-2.57	-5.70
$r_{Mood^{\perp}}$	-0.24	-0.30	-0.46	-0.05	-0.17	-0.21	-0.18	0.36	0.51	0.79	1.02
$t_{Mood^{\perp}}$	-2.72	-1.86	-3.49	-0.39	-1.38	-1.57	-1.16	1.32	2.96	9.36	9.44
$\bar{R}^2$	0.93	0.91	0.88	0.89	0.93	0.82	0.87	0.88	0.90	0.94	0.51
$CARH\&ST\&LT$											
α	-0.01	0.03	0.12	0.12	0.02	-0.09	0.03	-0.70	-0.15	-0.02	-0.02
t_{α}	-0.04	0.22	0.66	0.78	0.18	-0.43	0.13	-4.90	-0.75	-0.25	-0.12
MKT	0.99	1.03	1.00	0.99	1.04	0.90	1.05	1.12	1.03	1.03	0.04
t_{MKT}	38.51	34.02	19.66	24.78	29.48	19.21	20.49	23.60	23.49	30.40	0.86
SMB	-0.18	-0.07	-0.20	-0.01	0.00	-0.01	0.01	0.19	0.33	0.17	0.35
t_{SMB}	-4.30	-0.88	-2.19	-0.17	0.03	-0.13	0.16	2.37	2.85	3.05	5.18
HML	0.09	0.02	-0.02	-0.08	0.08	0.00	0.08	-0.03	-0.03	-0.12	-0.21
t_{HML}	1.80	0.23	-0.17	-0.85	1.58	-0.01	0.62	-0.23	-0.34	-1.60	-2.63
CMA	0.00	0.20	0.05	0.14	0.16	0.09	-0.01	0.20	-0.36	0.02	0.02
t_{CMA}	0.00	1.25	0.22	1.02	1.64	0.56	-0.05	0.88	-2.56	0.16	0.11
RMW	0.02	-0.03	0.16	0.04	0.04	0.10	0.13	-0.15	-0.10	0.08	0.06
t_{RMW}	0.31	-0.26	1.17	0.42	0.73	0.54	1.09	-0.76	-0.84	0.90	0.53
MOM	0.15	0.04	0.04	-0.05	0.03	0.07	-0.04	-0.05	-0.16	-0.07	-0.22
t_{MOM}	4.15	1.08	1.08	-0.76	1.30	0.73	-0.53	-1.14	-2.92	-2.71	-6.15
ST	0.11	0.08	0.00	-0.09	0.01	0.10	0.04	0.16	0.09	-0.06	-0.18
t_{ST}	2.16	1.54	0.03	-1.26	0.19	0.84	0.70	3.28	1.75	-1.13	-2.65
LT	0.02	-0.24	0.16	-0.02	-0.18	0.18	0.15	0.03	0.14	0.11	0.09
t_{LT}	0.40	-3.97	1.25	-0.19	-3.37	1.37	1.25	0.26	1.10	1.46	1.19
$r_{Mood^{\perp}}$	-0.24	-0.30	-0.46	-0.05	-0.17	-0.21	-0.18	0.36	0.51	0.79	1.02
$t_{Mood^{\perp}}$	-2.89	-2.05	-3.76	-0.39	-1.53	-1.61	-1.17	1.33	3.04	9.61	9.58
$\bar{R}^2$	0.93	0.92	0.88	0.89	0.93	0.83	0.87	0.88	0.90	0.94	0.55

4.2 Pricing Analysis for 25 Size-Mood Portfolios

Beginning with December 2009, we use the NYSE breakpoints to split stocks into quintiles of market capitalization at the end of each month. Independently, we use the NYSE breakpoints to split stocks into quintiles based on absolute mood beta estimated from equation (11). We then form 25 size and mood portfolios by taking intersections.[30] Monthly value-weighted portfolio returns are calculated from the end of month t to the end of month $t+1$, and portfolios are re-balanced at the end of each month $t+1$.

Table 10 reports average excess returns and alphas from the Fama-French five-factor, Carhart six-factor, and full eight-factor models, including reversals. There is clear evidence that moody stocks earn high excess returns per month. From the upper left panel of Table 10, we see that the smallest stocks with high sensitivity to mood earn 1.71% per month. Average returns for moody stocks are reasonably constant for successively larger size quintiles. Even the largest moody stocks earn returns of 1.52% per month. In fact, within all five mood quintiles, size does not appear to be an important differentiating factor in terms of average returns. However average returns clearly drop as we examine successively less moody stocks within each size category. The most sober (least moody) stocks typically earn no more than half the average return of the most moody stocks. The high-low portfolios defined as taking a long position in moody stocks and a short position in sober stocks within each size quintile earn economically and statistically significant positive returns.

The lower panels in Table 10 report alphas from progressively more complex factor models. Considering first the estimates based on the Fama-French five-factor model, alphas are typically positive and significant for the most moody stocks, and negative and significant for less moody stocks (including, in several places, portfolios comprising stocks with intermediate levels of mood sensitivity). The high-low strategy alphas are always positive, often significantly so. Again, consistent with earlier results, both the magnitude and statistical significance of alphas increase as the factor models become more complex. In the bottom panel reporting the results of the Fama-French five-factor model augmented with momentum and short- and long-term reversal factors, high-low strategy alphas are very similar to the average excess returns (even though individual portfolio alphas are very different). It is evident that conventional factor models cannot adequately price size-mood beta-sorted portfolios.

We now test the pricing power of our mood factor constructed in the section above on the same 25 portfolios. The top panel of Table 11 reports the key results from pricing these portfolios

[30] Here we perform independent sorts on size and mood beta. As noted, stocks that are sensitive to mood tend to be smaller and so the portfolios are unbalanced in terms of numbers of stocks. Each month, the portfolio of large, moody stocks contains much fewer stocks than the portfolio of small moody stocks. Nevertheless, there are sufficient stocks in even the smallest portfolio to perform the analysis. In the robustness tests we perform conditional sorts, first on size then on mood betas, in order to balance the portfolios better. Our findings are robust to this approach.

Table 10: 25 Size and Mood Value-Weighted Portfolio Analysis

Starting from December 2009, we use the NYSE breakpoints to split the NYSE, Amex and NASDAQ stocks into quintiles on the market equity at the end of each month. Independently, we use the NYSE breakpoints to split the NYSE, Amex and NASDAQ stocks into quintiles on absolute value of mood beta calculated from model (1.11). We form 25 size and mood portfolios by taking intersections. Monthly value-weighted portfolio returns are calculated from the end of month t to the end of month $t+1$, and portfolios are rebalanced at the end of each month $t+1$. H/L is the high-low portfolio to long mood-affected stocks (portfolio 5 ranked on mood) and to short mood-insensitive stocks (portfolio 1 ranked on mood) respecting each size portfolio. Market factor (MKT), size factor (SMB), value factor (HML), investment factor (CMA), profitability factor (RMW), momentum factor (MOM), short-term reversal factor (ST) and long-term reversal factor (LT) are downloaded from the Kenneth R. French Data Library. We report the mean monthly alphas across 25 portfolios based on Fama-French five-factor regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \varepsilon_{p,t}$), the Carhart momentum factor regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \varepsilon_{p,t}$) and the full behavioral factors regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \beta_{ST,p}ST_t + \beta_{LT,p}LT_t + \varepsilon_{p,t}$).

	Sober	2	3	4	Moody	High-Low		Sober	2	3	4	Moody	High-Low
				$Mean$							t_{Mean}		
Small	0.77	1.46	0.87	0.94	1.71	0.93	Small	1.30	2.66	1.70	1.68	2.87	4.34
2	0.77	1.34	1.06	1.12	1.28	0.51	2	1.74	2.64	2.22	2.14	2.28	2.30
3	0.63	0.91	1.10	1.25	1.37	0.75	3	1.36	1.93	2.07	2.53	2.11	2.54
4	0.82	1.13	1.15	0.79	1.79	0.96	4	2.07	2.41	2.66	1.75	3.30	3.94
Large	0.74	0.99	0.67	0.85	1.52	0.78	Large	2.23	3.09	1.83	2.27	3.42	3.15

	Sober	2	3	4	Moody	High-Low		Sober	2	3	4	Moody	High-Low
				α^{FF}							$t_{\alpha^{FF}}$		
Small	-0.41	0.20	-0.37	-0.36	0.43	0.84	Small	-1.77	0.82	-2.21	-2.42	2.53	3.06
2	-0.43	0.25	-0.19	-0.19	-0.16	0.27	2	-4.57	1.35	-1.54	-1.70	-0.95	1.48
3	-0.54	-0.32	-0.20	0.01	-0.08	0.46	3	-2.93	-2.12	-1.15	0.08	-0.30	1.71
4	-0.30	-0.02	0.02	-0.42	0.40	0.70	4	-1.95	-0.13	0.14	-1.69	2.08	3.29
Large	-0.38	-0.13	-0.45	-0.42	0.39	0.77	Large	-2.82	-0.98	-2.08	-2.72	2.52	3.46

	Sober	2	3	4	Moody	High-Low		Sober	2	3	4	Moody	High-Low
				α^{CARH}							$t_{\alpha^{CARH}}$		
Small	-0.40	0.27	-0.35	-0.30	0.54	0.93	Small	-1.75	1.22	-2.08	-2.19	2.93	3.26
2	-0.39	0.30	-0.14	-0.13	-0.06	0.33	2	-3.94	1.72	-1.17	-1.23	-0.37	1.88
3	-0.54	-0.30	-0.20	0.05	-0.02	0.52	3	-2.90	-2.03	-1.15	0.26	-0.07	2.16
4	-0.29	0.00	0.03	-0.37	0.47	0.76	4	-1.94	0.00	0.23	-1.70	2.66	3.62
Large	-0.39	-0.13	-0.47	-0.39	0.45	0.84	Large	-3.30	-0.91	-2.02	-2.21	2.89	4.64

	Sober	2	3	4	Moody	High-Low		Sober	2	3	4	Moody	High-Low
				$\alpha^{CARH\&ST\&LT}$							$t_{\alpha^{CARH\&ST\&LT}}$		
Small	-0.37	0.31	-0.29	-0.23	0.58	0.95	Small	-1.61	1.26	-1.85	-1.68	3.20	3.40
2	-0.41	0.22	-0.14	-0.09	0.04	0.45	2	-3.57	1.11	-1.16	-0.85	0.22	2.22
3	-0.57	-0.30	-0.22	0.07	0.11	0.67	3	-3.12	-1.77	-1.12	0.35	0.44	2.77
4	-0.31	0.02	0.04	-0.34	0.61	0.92	4	-2.06	0.14	0.25	-1.47	4.00	5.08
Large	-0.41	-0.08	-0.48	-0.36	0.47	0.89	Large	-3.31	-0.52	-2.10	-1.98	2.66	3.82

Table 11: 25 Size and Mood Value-Weighted Portfolio Analysis with Mood Factor

Starting from December 2009, we use the NYSE breakpoints to split the NYSE, Amex and NASDAQ stocks into quintiles on the market equity at the end of each month. Independently, we use the NYSE breakpoints to split the NYSE, Amex and NASDAQ stocks into quintiles on absolute value of mood beta calculated from model (1.11). We form 25 size and mood portfolios by taking intersections. Monthly value-weighted portfolio returns are calculated from the end of month t to the end of month $t+1$, and portfolios are rebalanced at the end of each month $t+1$. H/L is the high-low portfolio to long mood-affected stocks (portfolio 5 ranked on mood) and to short mood-insensitive stocks (portfolio 1 ranked on mood) respecting each size portfolio. Market factor (MKT), size factor (SMB), value factor (HML), investment factor (CMA), profitability factor (RMW), momentum factor (MOM), short-term reversal factor (ST) and long-term reversal factor (LT) are downloaded from the Kenneth R. French Data Library. We report the mean monthly alphas of the orthogonalized mood factor pricing across 25 portfolios based on Fama-French five-factor regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{Mood^{\perp},p}r^{\perp}_{Mood,t} + \varepsilon_{p,t}$), the Carhart momentum factor regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \beta_{Mood^{\perp},p}r^{\perp}_{Mood,t} + \varepsilon_{p,t}$) and the full behavorial factors regressions ($r_{p,t} - r_f = \alpha_p + \beta_{MKT,p}MKT_t + \beta_{SMB,p}SMB_t + \beta_{HML,p}HML_t + \beta_{CMA,p}CMA_t + \beta_{RMW,p}RMW_t + \beta_{MOM,p}MOM_t + \beta_{ST,p}ST_t + \beta_{LT,p}LT_t + \beta_{Mood^{\perp},p}r^{\perp}_{Mood,t} + \varepsilon_{p,t}$).

FF & Mood

	Sober	2	3	4	Moody	High-Low
			$\alpha^{Mood^{\perp}}$			
Small	-0.39	0.32	-0.38	-0.23	0.23	0.62
2	-0.13	0.38	-0.05	-0.03	-0.09	0.05
3	-0.28	0.04	-0.06	0.04	-0.25	0.03
4	-0.13	0.05	0.03	-0.25	0.15	0.28
Large	-0.15	0.02	-0.31	-0.55	-0.24	-0.09

	Sober	2	3	4	Moody	High-Low
			$t^{\alpha^{Mood^{\perp}}}$			
Small	-1.89	1.19	-1.80	-1.30	1.40	2.62
2	-1.20	2.03	-0.28	-0.23	-0.43	0.21
3	-1.46	0.20	-0.32	0.22	-0.86	0.10
4	-0.86	0.23	0.19	-1.03	0.56	1.20
Large	-1.09	0.13	-1.68	-3.15	-1.10	-0.31

	Sober	2	3	4	Moody	High-Low
			$\beta_{Mood^{\perp}}$			
Small	-0.04	-0.22	0.02	-0.22	0.36	0.40
2	-0.54	-0.25	-0.26	-0.29	-0.12	0.41
3	-0.46	-0.65	-0.24	-0.04	0.31	0.78
4	-0.32	-0.12	-0.02	-0.31	0.44	0.76
Large	-0.42	-0.27	-0.26	0.23	1.14	1.55

	Sober	2	3	4	Moody	High-Low
			$t_{\beta_{Mood^{\perp}}}$			
Small	-0.23	-0.93	0.11	-1.06	2.04	1.71
2	-3.43	-1.63	-1.47	-1.50	-0.81	2.10
3	-3.32	-3.23	-1.61	-0.25	1.39	3.14
4	-1.73	-0.62	-0.12	-1.31	1.88	3.23
Large	-2.46	-2.45	-1.60	1.40	4.49	5.62

CARH & Mood

	Sober	2	3	4	Moody	High-Low
			$\alpha^{Mood^{\perp}}$			
Small	-0.38	0.39	-0.36	-0.18	0.33	0.71
2	-0.09	0.44	0.00	0.03	0.01	0.10
3	-0.28	0.06	-0.06	0.07	-0.19	0.09
4	-0.11	0.07	0.05	-0.20	0.23	0.34
Large	-0.16	0.02	-0.32	-0.52	-0.18	-0.02

	Sober	2	3	4	Moody	High-Low
			$t_{\alpha^{Mood^{\perp}}}$			
Small	-1.83	1.72	-1.72	-1.24	1.73	2.73
2	-0.92	2.47	0.03	0.27	0.07	0.48
3	-1.45	0.31	-0.32	0.37	-0.78	0.36
4	-0.83	0.37	0.25	-0.93	0.99	1.64
Large	-1.34	0.14	-1.63	-2.60	-1.02	-0.09

	Sober	2	3	4	Moody	High-Low
			$\beta_{Mood^{\perp}}$			
Small	-0.04	-0.22	0.02	-0.22	0.36	0.40
2.00	-0.54	-0.25	-0.26	-0.29	-0.12	0.41
3.00	-0.46	-0.65	-0.24	-0.04	0.31	0.78
4.00	-0.32	-0.12	-0.02	-0.31	0.44	0.76
Large	-0.42	-0.27	-0.26	0.23	1.14	1.55

	Sober	2	3	4	Moody	High-Low
			$t_{\beta_{Mood^{\perp}}}$			
Small	-0.22	-1.03	0.11	-1.26	2.66	2.57
2.00	-4.04	-1.39	-2.13	-1.97	-0.83	1.89
3.00	-3.32	-3.25	-1.60	-0.26	1.27	3.05
4.00	-1.72	-0.63	-0.11	-1.08	2.32	3.92
Large	-2.46	-2.40	-1.54	1.49	5.88	7.50

CARH&ST&LT&Mood

	Sober	2	3	4	Moody	High-Low
			$\alpha^{Mood^{\perp}}$			
Small	-0.34	0.44	-0.30	-0.11	0.38	0.72
2.00	-0.11	0.35	0.01	0.07	0.10	0.22
3.00	-0.31	0.06	-0.08	0.09	-0.07	0.24
4.00	-0.14	0.09	0.05	-0.17	0.36	0.50
Large	-0.18	0.07	-0.34	-0.49	-0.16	0.03

	Sober	2	3	4	Moody	High-Low
			$t_{\alpha^{Mood^{\perp}}}$			
Small	-1.69	1.70	-1.45	-0.73	2.00	2.82
2.00	-1.15	1.68	0.04	0.59	0.55	1.01
3.00	-1.59	0.31	-0.39	0.45	-0.31	1.09
4.00	-0.92	0.51	0.26	-0.73	1.68	2.78
Large	-1.47	0.45	-1.72	-2.37	-0.81	0.11

	Sober	2	3	4	Moody	High-Low
			$\beta_{Mood^{\perp}}$			
Small	-0.04	-0.22	0.02	-0.22	0.36	0.40
2.00	-0.54	-0.25	-0.26	-0.29	-0.12	0.41
3.00	-0.46	-0.65	-0.24	-0.04	0.31	0.78
4.00	-0.32	-0.12	-0.02	-0.31	0.44	0.76
Large	-0.42	-0.27	-0.26	0.23	1.14	1.55

	Sober	2	3	4	Moody	High-Low
			$t_{\beta_{Mood^{\perp}}}$			
Small	-0.23	-1.03	0.10	-1.24	2.71	2.46
2.00	-4.22	-1.35	-2.22	-1.96	-0.87	2.12
3.00	-3.19	-3.30	-1.57	-0.27	1.54	3.60
4.00	-1.69	-0.64	-0.11	-1.11	2.25	3.93
Large	-2.54	-2.76	-1.58	1.42	6.03	7.46

using the Fama-French five-factor model augmented with the orthogonalized mood factor. The first block reports alphas. In most cases, the alphas are much smaller than reported in Table 10, and statistical significance is lost. The main exceptions are for the smallest quintile of stocks where the most sober stocks report a marginally significant and negative alpha which, when paired with the positive but insignificant alpha from the most moody stocks gives, a positive long-short alpha. This is one-third smaller than the equivalent alpha reported in Table 10, but it remains statistically significant. The second block reports the loadings on the mood factor in the regression. The pattern of these loadings is as expected: positive and significant for the most moody stocks and for the long-short strategy, negative and significant for the less moody portfolios. Subsequent panels report results for more complex factor models but the inferences are quite similar to those from the Fama-French model. Loadings on the mood factor are often very significant and follow the expected patterns in both sign and magnitudes, while the magnitudes of alphas for the 25 portfolios are typically so reduced as to lose statistical significance. High-low portfolios, however, still offer positive alphas that are sometimes statistically significant.

We conclude that the orthogonalized mood factor has important additional pricing power beyond that offered by the benchmark Fama-French five-factor model, even when augmented with previously identified behavioral factors. The mood factor does a reasonable but not perfect job of explaining the “mispricing” caused by incorporating mood as pricing information rather than acquiring fundamental information to learn about assets across these 25 portfolios. In subsequent drafts, we will test its ability to explain returns on portfolios sorted on the basis of other known anomalies.

4.3 Zero Mood Beta Test

The high excess monthly returns on mood-sensitive portfolios (portfolio 1 and portfolio 10) indicate that moody stocks earn higher expected returns and are consistent with the theoretical result in section 2. In fact, the excess returns of moody stocks are due to the irrational information acquisition decision induced by mood contributes an additional risk that can not be captured by traditional pricing factors or the unconditional betas. Naively, looking at the sign of the sensitivity of mood risk, the mood beta in equation (11), it could be questioned whether the mood risk can be canceled out by holding both negative and positive sensitive mood stocks. Therefore, in this section we test whether the exposure to mood risk is hedged by taking a long position in both negative mood portfolio (P1) and positive mood portfolio (P10).

To test whether this mood risk hedging strategy is feasible, we separate the mood beta for each stock into upward mood betas and downward mood betas. The upward mood beta measures the exposure of a stocks to increases in the public mood, while the downward mood beta measures

Table 12: Zero Mood Beta Test

Panel A: The upward mood beta and downward mood beta are calculated based on the regression model : $r_{i,t} = \alpha_i + \beta_{MKT,i}MKT_t + \beta_{SMB,i}SMB_t + \beta_{HML,i}HML_t + \beta_{CMA,i}CMA_t + \beta_{RMW,i}RMW_t \beta_{MOM,i}MOM_t + \beta_{Mood^+,i}D * \Delta Mood_t^+ + \beta_{Mood^-,i}(1-D) * \Delta Mood_t^- + \varepsilon_{i,t}$. D is a dummy variable to identify if a day has an upward mood change. We conduct time series regression for each stock to isolate risk exposure to upward and downward mood change. The upward mood beta for negative mood sensitive stocks is the sensitivity of portfolio 1 return decreased on the days with upward mood change. The downward mood beta for positive mood sensitive stocks is the sensitivity of portfolio 10 return decreased on the days with downward mood change. Panel B: For each regression period, we hold stocks from portfolio 1 and 10 and calculate the cross-sectional mean of upward and downward mood betas. The upward mood hedging is to add upward mood beta from portfolio 1 and portfolio 10 at each period. The downward mood hedging is to add downward mood beta from portfolio 1 and 10.

Panel A : Upside and Downside Mood Beta

Portfolio	β^{Mood^+}	β^{Mood^-}
1	-0.52	-0.64
2	-0.23	0.24
3	-0.14	-0.14
4	-0.08	-0.07
5	-0.03	-0.02
6	0.02	0.04
7	0.08	0.10
8	0.16	0.16
9	0.26	0.26
10	0.65	0.59

Panel B: Mood Beta Hedging Test

Mood Hedging	Upside Mood	Downside Mood
$Coef$	0.061	-0.035
$t_{statistics}$	6.42	-3.01

the exposure to mood when public mood decreases. For each stock we conduct the following regression:

$$r_{i,t} = \alpha_i + \sum_{m=1}^{M} \beta_{i,m} f_{m,t} + \beta_{Mood^+,i} D * \Delta Mood_t^+ + \beta_{Mood^-,i}(1-D) * \Delta Mood_t^- + \varepsilon_{i,t} \tag{15}$$

where $f_{m,t}$ includes Fama-French five factors and the momentum factor. β_{Mood^+} captures the mood exposure for each stock using days when the public mood improves and β_{Mood^-} captures the mood exposure days of worsening mood. D is a dummy variable to identify whether a day has an upward mood change. Panel A in Table 12 gives the time series average of cross-sectional upward and downward mood betas. Portfolio 10 has an upward mood beta of 0.65 and a downward mood beta of 0.59. Returns on this portfolio increase more on days when mood improves than they fall on days when mood worsens by an equivalent amount. Conversely, portfolio 1 has a -0.52 beta to upward mood and a -0.64 beta to downward mood. Its returns on bad mood days are larger than its losses on equivalent good mood days. There is clear (and statistically significant) asymmetry between upward and downward mood betas in the most moody stock portfolios. It is noticeable that such asymmetry is only found for the two extreme portfolios (P1 and P10). Less moody portfolios have symmetric mood betas.

To assess the performance of the hedging strategy further, we analyze the upward and downward mood betas of the portfolio formed by combining portfolios 1 and 10. Panel B in Table 12 gives the time series averages of each beta. The downward mood beta is -0.035, more than 3 standard errors below zero. The average upward mood beta is 0.061, with an even larger t-statistic. These findings imply that taking long positions in both positive and negative beta stocks does not eliminate exposure to mood. On good mood days, the gains from the positive mood beta stocks outweigh the losses on negative mood beta stocks - with the opposite holding true on bad mood days - because mood betas are asymmetric. Positive (negative) mood beta stocks gain more on good (bad) mood days than they lose on bad (good) mood days.

In summary, combining positive and negative mood beta stocks does not hedge mood risk exposure. Taking a long position on portfolio 1 and portfolio 10 will always involve an exposure to mood risk on either upward mood days or downward mood days. As argued in the literature, mood can irrationalize investors with biased decision making and asset valuation through different channels such as deflecting risk version, tolerance, biasing common pricing factors etc. We contribute to the literature by exploring the mood effect through the biasing channel on investors' information acquisition in respect of assets. More importantly, the additional risk originating from the biased information acquisition triggered by mood implies a higher return as required by investors to hold the mood-sensitive assets. In the next section, we conduct a detailed discussion about the theoretical motivation in our study through our interpretation of how mood gives rise to bias and becomes an indispensable risk factor that should be recognized and compensated for by investors.

5 Discussion

The empirical results support our key argument that stocks that are sensitive to mood as a bias factor in investors' decision to acquire information about assets' payoff and valuation earn higher expected returns. These can be thought of as risk premia to compensate investors who want to hold these moody stocks. In fact, our study is closely connected to studies that discuss theoretical perspectives of how mood creates bias in financial markets.

On the one hand, some studies argue that there is a negative relationship between people's risk aversion and mood (Kamstra et al., 2003; Kramer and Weber, 2012; Bassi et al., 2013; Kaplanski et al., 2015). More specifically, investors in a positive mood tend to be less risk-averse or more risk-tolerant and vice versa. In fact, the effect of risk aversion deflected by mood causes mispricing into two ways: either through trading behavior from misvaluation of asset payoff or through incorrect perception of the stochastic discount factor.

First, positive mood causes investors to feel less risk-averse and more likely to perceive lower risk in stocks, or to believe that stocks are more likely to be underpriced (Goetzmann et al., 2015). Therefore, investors choose either to invest in more risky assets or to conduct more buying rather than selling as they are not consciously aware of their positive mood bias. Negative mood makes investors feel more risk-averse and more likely to perceive stocks as overpriced. As stated by

Nagel (2005), the short sale constraint acts as an indispensable condition in the market, especially for investors who are more likely to be biased by mood factors (Goetzmann et al., 2015). As a consequence, investors in a pessimistic mood choose stocks more carefully, with high expected returns as rewards for taking risk (Raghunathan and Pham, 1999). Eventually, as found by Goetzmann et al. (2015), stock returns that are liable to be affected by mood generate comovement during positive and negative mood days respectively. This is consistent with the evidence found in our mood beta estimation, in which mood-sensitive stocks have positive and negative mood betas. In other words, positive moody stock returns comove with positive mood days; by the same token, negative moody stock returns move together with negative mood days.

Second, the lower risk aversion from the positive mood effect could also bias investors to increase the stochastic discount factor (SDF) (Shu, 2010). As investors use this subjective discount factor to price assets, they anticipate a lower expected return as lower risk aversion implies higher risk tolerance. Therefore, they overprice the stocks and induce more buying instead of selling. Of course, higher risk aversion from a negative mood effect decreases the SDF. According to the short sale constraint, moody investors seek to invest in stocks with higher expected returns as risk compensation, as investors over-perceive the risk on negative mood days.

Overall, stocks invested in by moody investors are risky on average, as shown in our findings on financial characteristics (small in size, high idiosyncratic risk etc.) for these stocks subject to mood risk. More importantly, mood as ”feeling” information is added to investors' valuation of risky assets, which contributes to risks in addition to the risk factors contained in fundamental information. Therefore, in our study, we seek to answer the question of how much the risk induced by the mood through investors' insufficient information acquisition in respect of risky assets should be compensated for with high expected returns as a risk premium by introducing the existence of mood risk in a proportion of stocks.

On the the other hand, the recent seminal study by Hirshleifer et al. (2020) proposes a multi-factor asset pricing model. They argue that the hard-to-value factor is liable to be biased by investors' mood in subsequent trading periods. In the mean time, the aggregate market factor is biased by the public mood as well. They too use the term 'mood beta', although in their application this is constructed by regressing a stock's returns on equal-weighted market returns during periods conjectured to be associated with investor mood swings. Stocks with a higher mood beta earn higher (lower) returns during future positive (negative) mood seasons. These results are consistent with several recent papers identifying seasonality in a cross-section of security returns (Heston and Sadka, 2008; Keloharju et al., 2016; Birru, 2018). The focus of Hirshleifer, Jiang, and DiGiovanni's (2020) paper is on seasonality and as such they naturally use seasonal patterns in mood to identify mood sensitivities.

In contrast, Bali et al. (2017) measure stocks' sensitivity to economic uncertainty by running time-series regressions of a multi-factor pricing model including an uncertainty index as another proposed factor. In line with this, we directly estimate our mood betas as sensitivities to an exogenous measure of mood in the factor pricing model. For this reason, our mood betas are probably very different to those of Hirshleifer et al. (2020), who argue that a particular factor in asset payoff is liable to be biased by mood. In particular, while their analysis draws on the conjecture that investors are in a happier mood in January and March compared with the supposedly sadder months of September and October, we note that our Twitter mood index of happiness follows a low-frequency cycle. This cycle is not the business cycle (Twitter mood peaked in early 2010 and late 2015, and hit troughs in late 2012 and mid-2017). It also means that once the effect of regular holidays is excluded there is very little monthly seasonality in Twitter mood.

It might be tempting to think that the factor biased by investor mood in the study by Hirshleifer et al. (2020) takes a view similar to our proposal regarding mood sensitivity. However, we conduct our study in a more parsimonious or generalized setting and do not specify how the bias contributes specific parameters (risk aversion, time preference or factors etc.) in asset pricing models. The bias might be affecting any of the key factors in a pricing model, or the mood itself might serve as a factor mistakenly incorporated by investors into pricing models. In actual fact, those particular biasing channels to investors' mispricing can be summarized by the the problem of information incorporation in asset valuation. If investors' decisions on the acquisition of fundamental information about assets are not as rational expected under *Homo economicus*, investors eventually incorporate insufficient fundamental information when they price assets. Therefore, we take the behavioral finance perspective to argue that the mood swings can result in insufficient acquisition of information that should be incorporated in valuation. Finally, this comes back to our key premise, which is to find which stocks are the "volunteers" to the mood risk regardless of how stocks or investors are affected by mood as its existence of bias on particular factors in decision-making and asset pricing has been broadly addressed. More importantly, our interest is to elucidate whether the mood is priced as a risk premium which investors need to consider in holding the stocks which are liable to be biased by the mood effect.

In sum, we propose that not all stocks in financial markets are liable to be biased by investor mood; that there is a subset of stocks which are traded by investors who are more likely to be affected by mood with consequent insufficient information acquisition; that regardless of the specific learning factors through which mood bias causes mispricing, the mood effect acts as a risk factor in addition to fundamental factors in asset pricing models; and that stocks which are subject to mood risk earn a higher expected return (risk premium) as a compensation for investors to hold them.

6 Robustness Tests

Our analysis entailed several conscious choices. In this section we test whether our results depend on any of them. Specifically, we replicate our results after making the following changes:

1. Break points:

In previous sections we allocated stocks to portfolios according to break points determined by NYSE stocks. We repeat our analysis using break points determined by poling stocks from all three main venues-NYSE, Nasdaq and Amex. Our results are not affected by this change in any material way.

2. Portfolio weights:

We used value weighs to construct portfolios. We repeat our analysis using equal weights. Naturally, this does change the magnitude of some of our estimates, but less than expected. We suspect that this is because sorting by mood sensitivities is positively correlated with sorting by size. The decile of most positively mood-sensitive stocks are all relatively small, as are the stocks in the decile of most negatively mood-sensitive stocks. The large stocks typically - but not always - fall into the mood-insensitive deciles. Equal weighting within each decile is then not very different from value weighting. Since value weighting is the norm in this field, we report results based on this approach in the study.

3. Factor models:

Earlier drafts of this study used the Fama-French three-factor model as the starting point of all analyses. Mood betas were calculated after conditioning on these factors in equation (11), and subsequent analysis began with the Fama-French three-factor model before augmenting this with momentum and reversal factors. Over the last two decades, there have been studies that clearly point out the inability of the Fama-French three-factor model to explain a range of anomalies in cross-sectional stock returns.[31] For this reason, we adopt the Fama and French (2015) five-factor asset pricing model, which adds investment and profitability factors to the original three Fama-French factors. We could equally have taken the q-factor model of Hou et al. (2015) as our starting point.

Overall, our findings are all robust to the use of the Fama and French three-factor (1992), the five-factor (2015) or the q-factor models as a starting point. We obtain very similar results if we compute mood betas conditioned on a simple CAPM or Fama-French three-factor model (or, indeed, allow for no conditioning in equation (11)), even if we apply more complex models in subsequent stages of the analysis.

The traditional fundamental financial factors—market, size, value, profitability and invest-

[31] The studies of, for instance, Ikenberry et al. (1995), Loughran and Ritter (1995), Spiess and Affleck-Graves (1995), Chan et al. (1996), Dichev (1998), Griffin and Lemmon (2002), Ang et al. (2006), Daniel and Titman (2006), Campbell et al. (2008) and Hou et al. (2015) all question the performance of the Fama-French model.

ment (however defined)—cannot capture the effects of sensitivity to mood, even when augmented by behavioral factors such as Carhart's (1997) momentum and reversal factors.

4. Mood factor construction:

We orthogonalized the mood factor to our most general factor model by taking only the unexplained components of equation (12). This assigns all explanatory power common to mood and another factor to that other factor, thereby giving the mood factor the least possible credit for any explanatory power it might have. Nevertheless, we find that it has considerable explanatory power over and above the previously identified factors. Repeating our analysis based on the raw mood factor changes our results in the expected way, increasing its power slightly, but it does not affect inference.

In summary, we have tried to follow the standard empirical path in testing for a new factor. Deviations from this path would not have materially affected our conclusions, and our results are robust.

7 Conclusions

The neoclassical finance paradigm answers questions about financial markets by applying models in which the economic agents are rational. However, it is becoming increasingly apparent that this framework struggles to elucidate the essential facts with respect to the aggregate stock market, the cross-section of average returns and individuals' trading behavior (Barberis and Thaler, 2003). The effects of mood on the stock market have been addressed in the literature by attempting to connect either investors' shifting risk tolerance or directly biased pricing factors. As a consequence, mood irrationalizes investors' decision-making and trading behaviors. In line with existing studies on the mood effect in behavioral finance, we explore an innovative argument that mood can bias investors' decision to acquire information pertaining to an asset. As investors' mood swings, they tend to acquire less earnings-related information to learn about companies' performance.

Based on the empirical evidence in the data showing that mood significantly affects investors' information acquisition, we study this effect by understanding the asset risk and discussing the failures of classical pricing models with investors' inadequate learning about assets induced by mood. We test theoretical predictions implied by the mood effect on a cross-section of U.S. equity returns using a high-frequency Twitter-based mood index and applying a traditional asset-pricing approach. First, we sort stocks into portfolios on the basis of their sensitivity to changes in Twitter mood. Stocks that are either highly negatively or positively sensitive to changes in Twitter mood earn higher monthly excess returns than stocks with low or no mood sensitivity. Our empirical results offer strong evidence that risk created by agents who are more likely to be affected by psychological feelings such as mood, particularly on information acquisition decisions, earns greater expected returns, and that mood risk is priced. Second, we examine the financial characteristics of mood-sensitive stocks. In particular, stocks that are small in size, relatively young, pay less in dividends, are non-profitable, engage more in external financing, and have higher levels of idiosyncratic risk are more sensitive to changes in public mood.

To identify the quantity of risk affected by public mood, we construct mimicking portfolios by taking long positions in mood-sensitive stocks and short positions in mood-insensitive stocks. The mood risk factor earns an average return of 0.56% per month, which is not captured by traditional asset-pricing models such as the Fama-French five-factor model even if augmented with more behavioral factors such as Carhart's momentum factor. When we include our mood factor in the pricing regressions, the alphas of mood sensitivity-sorted portfolios are significantly reduced, usually to levels of insignificance. As we document in this study, in addition to fundamental risks, the mood effect adds more risk to the valuation of assets through its inducement of investors' insufficient information acquisition. Investors require risk premia as compensations to hold stocks which are more likely to be affected by mood through bias channels.

References

Abraham, A. and Ikenberry, D. L. (1994). The individual investor and the weekend effect. *Journal of Financial and Quantitative Analysis*, 29(2):263–277.

Andrei, D., Friedman, H. L., and Ozel, N. B. (2020). Economic uncertainty and investor attention. *Available at SSRN 3128673*.

Ang, A., Hodrick, R. J., Xing, Y., and Zhang, X. (2006). The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299.

Areni, C. S. and Burger, M. (2008). Memories of "bad" days are more biased than memories of "good" days: past saturdays vary, but past mondays are always blue. *Journal of Applied Social Psychology*, 38(6):1395–1415.

Baker, M. and Wurgler, J. (2006). Investor sentiment and the cross-section of stock returns. *The journal of Finance*, 61(4):1645–1680.

Baker, M. and Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of economic perspectives*, 21(2):129–152.

Baker, M., Wurgler, J., and Yuan, Y. (2012). Global, local, and contagious investor sentiment. *Journal of financial economics*, 104(2):272–287.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4):1593–1636.

Bali, T. G., Brown, S. J., and Tang, Y. (2017). Is economic uncertainty priced in the cross-section of stock returns? *Journal of Financial Economics*, 126(3):471–489.

Bali, T. G., Engle, R. F., and Murray, S. (2016). *Empirical asset pricing: The cross section of stock returns*. John Wiley & Sons.

Barberis, N., Shleifer, A., and Vishny, R. (1998). A model of investor sentiment. *Journal of financial economics*, 49(3):307–343.

Barberis, N. and Thaler, R. (2003). A survey of behavioral finance. *Handbook of the Economics of Finance*, 1:1053–1128.

Bassi, A., Colacito, R., and Fulghieri, P. (2013). 'o sole mio: An experimental analysis of weather and risk attitudes in financial decisions. *The Review of Financial Studies*, 26(7):1824–1852.

Benamar, H., Foucault, T., and Vega, C. (2019). Demand for information, uncertainty, and the response of us treasury securities to news. *The Review of Financial Studies*.

Bikhchandani, S., Hirshleifer, D., and Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5):992–1026.

Birru, J. (2018). Day of the week and the cross-section of returns. *Journal of financial economics*, 130(1):182–214.

Bollen, J., Mao, H., and Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of computational science*, 2(1):1–8.

Brown, G. W. and Cliff, M. T. (2004). Investor sentiment and the near-term stock market. *Journal of empirical finance*, 11(1):1–27.

Bushee, B. J. and Friedman, H. L. (2016). Disclosure standards and the sensitivity of returns to mood. *The Review of Financial Studies*, 29(3):787–822.

Campbell, J. Y., Hilscher, J., and Szilagyi, J. (2008). In search of distress risk. *The Journal of Finance*, 63(6):2899–2939.

Carhart, M. M. (1997). On persistence in mutual fund performance. *The Journal of finance*, 52(1):57–82.

Chan, L. K., Jegadeesh, N., and Lakonishok, J. (1996). Momentum strategies. *The Journal of Finance*, 51(5):1681–1713.

Chang, S.-C., Chen, S.-S., Chou, R. K., and Lin, Y.-H. (2008). Weather and intraday patterns in stock returns and trading activity. *Journal of Banking & Finance*, 32(9):1754–1766.

Chung, S.-L., Hung, C.-H., and Yeh, C.-Y. (2012). When does investor sentiment predict stock returns? *Journal of Empirical Finance*, 19(2):217–240.

Cody, E. M., Reagan, A. J., Dodds, P. S., and Danforth, C. M. (2016). Public opinion polling with twitter. *arXiv preprint arXiv:1608.02024*.

Croft, G. P. and Walker, A. E. (2001). Are the monday blues ail in the mind? the role of expectancy in the subjective experience of mood 1. *Journal of Applied Social Psychology*, 31(6):1133–1145.

Daniel, K. and Titman, S. (2006). Market reactions to tangible and intangible information. *The Journal of Finance*, 61(4):1605–1643.

De Long, J. B., Shleifer, A., Summers, L. H., and Waldmann, R. J. (1990). Noise trader risk in financial markets. *Journal of political Economy*, 98(4):703–738.

Dey, S. (2014). Stock market prediction using twitter mood. *International Journal of Scientific & Engineering Research*, 5(5):44–47.

Dichev, I. D. (1998). Is the risk of bankruptcy a systematic risk? *the Journal of Finance*, 53(3):1131–1147.

Dodds, P. S. and Danforth, C. M. (2010). Measuring the happiness of large-scale written expression: Songs, blogs, and presidents. *Journal of happiness studies*, 11(4):441–456.

Dodds, P. S., Harris, K. D., Kloumann, I. M., Bliss, C. A., and Danforth, C. M. (2011). Temporal patterns of happiness and information in a global social network: Hedonometrics and twitter. *PloS one*, 6(12):e26752.

Edmans, A., Fernandez-Perez, A., Garel, A., and Indriawan, I. (2021). Music sentiment and stock returns around the world.

Edmans, A., Garcia, D., and Norli, Ø. (2007). Sports sentiment and stock returns. *The Journal of Finance*, 62(4):1967–1998.

Fama, E. F. and French, K. R. (1992). The cross-section of expected stock returns. *the Journal of Finance*, 47(2):427–465.

Fama, E. F. and French, K. R. (2015). A five-factor asset pricing model. *Journal of financial economics*, 116(1):1–22.

Frazzini, A. (2006). The disposition effect and underreaction to news. *The Journal of Finance*, 61(4):2017–2046.

Glushkov, D. (2006). Sentiment beta. *Available at SSRN 862444*.

Goetzmann, W. N., Kim, D., Kumar, A., and Wang, Q. (2015). Weather-induced mood, institutional investors, and stock returns. *The Review of Financial Studies*, 28(1):73–111.

Goetzmann, W. N. and Zhu, N. (2005). Rain or shine: where is the weather effect? *European Financial Management*, 11(5):559–578.

Griffin, J. M. and Lemmon, M. L. (2002). Book-to-market equity, distress risk, and stock returns. *The Journal of Finance*, 57(5):2317–2336.

Heston, S. L. and Sadka, R. (2008). Seasonality in the cross-section of stock returns. *Journal of Financial Economics*, 87(2):418–445.

Hirshleifer, D., Jiang, D., and DiGiovanni, Y. M. (2020). Mood beta and seasonalities in stock returns. *Journal of Financial Economics*.

Hirshleifer, D. and Shumway, T. (2003). Good day sunshine: Stock returns and the weather. *The Journal of Finance*, 58(3):1009–1032.

Hou, K., Xue, C., and Zhang, L. (2015). Digesting anomalies: An investment approach. *The Review of Financial Studies*, 28(3):650–705.

Huang, D., Jiang, F., Tu, J., and Zhou, G. (2015). Investor sentiment aligned: A powerful predictor of stock returns. *The Review of Financial Studies*, 28(3):791–837.

Ifcher, J. and Zarghamee, H. (2011). Happiness and time preference: The effect of positive affect in a random-assignment experiment. *American Economic Review*, 101(7):3109–29.

Ikenberry, D., Lakonishok, J., and Vermaelen, T. (1995). Market underreaction to open market share repurchases. *Journal of financial economics*, 39(2-3):181–208.

Jiang, D., Norris, D., and Sun, L. (2019). Weather, institutional investors, and earnings news. *Available at SSRN 3142635*.

Kamstra, M. J., Kramer, L. A., and Levi, M. D. (2000). Losing sleep at the market: The daylight saving anomaly. *American Economic Review*, 90(4):1005–1011.

Kamstra, M. J., Kramer, L. A., and Levi, M. D. (2003). Winter blues: A sad stock market cycle. *American Economic Review*, 93(1):324–343.

Kaplanski, G., Levy, H., Veld, C., and Veld-Merkoulova, Y. (2015). Do happy people make optimistic investors? *Journal of Financial and Quantitative Analysis*, pages 145–168.

Kaustia, M. and Rantapuska, E. (2016). Does mood affect trading behavior? *Journal of Financial Markets*, 29:1–26.

Keloharju, M., Linnainmaa, J. T., and Nyberg, P. (2016). Return seasonalities. *The Journal of Finance*, 71(4):1557–1590.

Kliger, D. and Levy, O. (2003). Mood-induced variation in risk preferences. *Journal of economic behavior & organization*, 52(4):573–584.

Kramer, L. A. and Weber, J. M. (2012). This is your portfolio on winter: Seasonal affective disorder and risk aversion in financial decision making. *Social Psychological and Personality Science*, 3(2):193–199.

Kumar, A. and Lee, C. M. (2006). Retail investor sentiment and return comovements. *The Journal of Finance*, 61(5):2451–2486.

Lee, C. M., Shleifer, A., and Thaler, R. H. (1991). Investor sentiment and the closed-end fund puzzle. *The journal of finance*, 46(1):75–109.

Lemmon, M. and Portniaguina, E. (2006). Consumer confidence and asset prices: Some empirical evidence. *The Review of Financial Studies*, 19(4):1499–1529.

Lo, A. W., Repin, D. V., and Steenbarger, B. N. (2005). Fear and greed in financial markets: A clinical study of day-traders. *American Economic Review*, 95(2):352–359.

Loughran, T. and Ritter, J. R. (1995). The new issues puzzle. *The Journal of finance*, 50(1):23–51.

Merton, R. C. et al. (1987). A simple model of capital market equilibrium with incomplete information.

Nagel, S. (2005). Short sales, institutional investors and the cross-section of stock returns. *Journal of financial economics*, 78(2):277–309.

O'Hara, M. (2003). Presidential address: Liquidity and price discovery. *The Journal of Finance*, 58(4):1335–1354.

Raghunathan, R. and Pham, M. T. (1999). All negative moods are not equal: Motivational influences of anxiety and sadness on decision making. *Organizational behavior and human decision processes*, 79(1):56–77.

Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., and Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5(1):1–12.

Reece, A. G., Reagan, A. J., Lix, K. L., Dodds, P. S., Danforth, C. M., and Langer, E. J. (2017). Forecasting the onset and course of mental illness with twitter data. *Scientific reports*, 7(1):1–11.

Ryan, R. M., Bernstein, J. H., and Brown, K. W. (2010). Weekends, work, and well-being: Psychological need satisfactions and day of the week effects on mood, vitality, and physical symptoms. *Journal of social and clinical psychology*, 29(1):95–122.

Saunders, E. M. (1993). Stock prices and wall street weather. *The American Economic Review*, 83(5):1337–1345.

Schwarz, N. and Clore, G. L. (2007). Feelings and phenomenal experiences.

Shefrin, H. and Statman, M. (1985). The disposition to sell winners too early and ride losers too long: Theory and evidence. *The Journal of finance*, 40(3):777–790.

Shu, H.-C. (2010). Investor mood and financial markets. *Journal of Economic Behavior & Organization*, 76(2):267–282.

Spiess, D. K. and Affleck-Graves, J. (1995). Underperformance in long-run stock returns following seasoned equity offerings. *Journal of financial economics*, 38(3):243–267.

Stambaugh, R. F., Yu, J., and Yuan, Y. (2012). The short of it: Investor sentiment and anomalies. *Journal of Financial Economics*, 104(2):288–302.

Stone, A. A., Schneider, S., and Harter, J. K. (2012). Day-of-week mood patterns in the united states: On the existence of 'blue monday','thank god it's friday'and weekend effects. *The Journal of Positive Psychology*, 7(4):306–314.

Swaminathan, B. (1996). Time-varying expected small firm returns and closed-end fund discounts. *The Review of Financial Studies*, 9(3):845–887.

Tetlock, P. C. (2014). Information transmission in finance. *Annu. Rev. Financ. Econ.*, 6(1):365–384.

Tirole, J. (2002). Rational irrationality: Some economics of self-management. *European Economic Review*, 46(4-5):633–655.

Van Nieuwerburgh, S. and Veldkamp, L. (2009). Information immobility and the home bias puzzle. *The Journal of Finance*, 64(3):1187–1215.

Veldkamp, L. L. (2011). *Information choice in macroeconomics and finance*. Princeton University Press.

Weller, B. M. (2018). Does algorithmic trading reduce information acquisition? *The Review of Financial Studies*, 31(6):2184–2226.

Wojcik, S. and Hughes, A. (2019). Sizing up twitter users. *PEW research center*, 24.

Appendix

A Definitions of Financial Characteristics

Financial data comes from the CRSP/Compustat merged database. We merge stock data from CRSP and financial data by linking PERMNO (CRSP) and LPERMNO (CRSP/Compustat). If data from PERMNO or LPERMNO is missing or not matched correctly, we fetch data by checking tickers in two databases. The financial data from fiscal year-ends $t-1$ for stock i in month t returns from June of year y to May of year $y+1$. All variables are winsorized at 99.5 and 0.5%.

Market Capitalization = Total Market Value (MKTVALT) at the end of fiscal year. If market value data is not available, we take closing price × common shares outstanding at the end of the fiscal year.
Book value of equity = SEQ + TXDB + ITCB -BVPS
SEQ is the book value of shareholders' equity. TXDB is deferred taxes. ITCB is investment tax credit. BVPS is book value of preferred stock, taken from PSTKRV (redemption value), PSTKL (liquidating value) or PSTK (par value) depending availability in the database. If there is no available data for preferred stock, BVPS is set to zero. We delete data which is missing either SEQ or TXDB.
B/M = Book Value of Equity / Total Market Value
Dividend Yield is dividends per share on the end of fiscal year (DVPSX_F).
If data is missing, we fill the data in as zero. Dividend Paid is the probability of a company paying a dividend. We calculate the total dividend paid as equal to dividend paid per share × common share outstanding, and then set a dummy variable as equal to 1. Otherwise, total dividend is equal to 0 if there is no dividend paid.
Operating Cash Flow is the operating activity net cash flow (OANCF).
EPS is earnings per share (Basic), excluding extraordinary items (EPSPX).
ROA = Net Income/Loss (NI) / Total Asset (AT)
EBITDA/ASSET = Earnings Before Interest (EBITDA) / Total Asset (AT)
For realistic investment purposes, we delete data missing key profitability information such as net income, operating cash flow or EBITDA.
Book Leverage = Total Debt Including Current (DT) / Total Asset (AT)
PPE/ASSET = Property, Plant and Equipment (PPEGT) / Total Asset (AT)
R&D= (XRD)/Total Asset (AT)
We calculate the research and development Expense by scaling the total assets.
If there is data missing on total debt, R&D and PPE, we fill the data in as zero.

Revenue is total revenue (REVT).

Asset Growth = $TotalAsset_t$- $TotalAsset_t - 1/TotalAsset_{t-1}$

External Financing is the difference between the percentage change of asset from year $t-1$ to year t and the percentage change of retained earnings from $t-1$ to t. However, retained earnings can be 0 or negative, and we define the calculation regarding the change of retained earnings as follows: $Retained\ earnings_{t-1}$ (re_{t-1}) is not equal to 0: $re_t - re_{t-1}/re_{t-1}$. If re_{t-1} is equal to 0 and $re_t > 0$, we set the change of retained earnings as equal to 1. If re_{t-1} is equal to 0 and $re_t < 0$, we set the change of retained earnings as equal to -1. Otherwise, the change of retained earning equals 0.

Age is calculated as the first date from which the company's data is available in the database up to December 2016.

Idiosyncratic risk is measured by taking RSE of residuals from the Carhart pricing model as below:

$$R_{i,t} = \alpha_i + \beta_{MKT,i}MKT_t + \beta_{SMB,i}SMB_t + \beta_{HML,i}HML_t + \beta_{CMA,i}CMA_t + \beta_{RMW,i}RMW_t + \beta_{MOM,i}MOM_t + \varepsilon_{i,t}$$

B Control Variable Definition

The earnings data used to test the mood biasing effect on information acquisition in section 2.2 is from Institutional Brokers Estimate System (I/B/E/S). The sample period is from 2008 to 2016, subject to Twitter data availability. We also conduct the test by extending the data to 2018, and the results are not changed a lot.

VIX: Daily closing value of VIX. Source: Wharton Research Data Services-CBOE Indexes.

EPU: Daily news-based Economic Policy Uncertainty Index. Source: BBM.

$Size$: Natural log of market value of equity. Source: Compustat.

RV: Return volatility is measured as standard deviation of daily return at each month. Source: CRSP.

$Institutional\ Ownership$ ($ITOW$): This is the institutional ownership percentage from Thomson Reuters Institutional (13f) Holdings data file.

$\overline{IVOL}$: Moving average stock idiosyncratic volatility is calculated based on the window between day $t-24$ and $t-4$. Sources: CRSP and Kenneth R. French Data Library.

$Price$: Average daily closing price from day $t-42$ to $t-21$ before a quarterly earnings announcement. Source: CRSP.

$NUMEST$: Number of analyst's earnings forecasts in the most recent month before a quarterly earnings announcement. Source: Institutional Brokers Estimate System (I/B/E/S).

$Turn$: Turnover is total number of shares traded over a period divided by total outstanding shares. Source: CRSP.

%Positive: Counts of days with positive daily mood change divided the total days in the most recent month before the firm earnings announcements.
%Negative: Counts of days with negative daily mood change divided the total days in the most recent month before the firm earnings announcements.