INVESTOR SENTIMENT AND ASSET RETURNS: ACTIONS SPEAK LOUDER THAN WORDS

October 2024

Xi Dong, Dat Mai, Kuntara Pukthuanthong, Guofu Zhou

Xi Dong is an associate of finance at Baruch College in New York, NY
xi.dong@baruch.cuny.edu

Dat Mai is a quantitative researcher at MKT MediaStats LLC in Cambridge, MA
maiydat@gmail.com

Kuntara Pukthuanthong is a professor of finance at the University of Missouri-Columbia in Columbia, MO
pukthuanthongk@missouri.edu

Guofu Zhou is a professor of finance at Washington University in St. Louis in St. Louis, MO
zhou@wustl.edu

Send correspondence to Kuntara Pukthuanthong at pukthuanthongk@missouri.edu.

Abstract

The paper studies daily predictability of investor sentiment across four major asset classes and compare sentiment measures based on news and social media with those based on trade information. For the majority of assets, trade-based sentiment measures outperform their text-based equivalents for both in-sample and out-of-sample predictions. This outperformance is particularly noticeable in long-term forecasts. However, real-time mean-variance investors can only achieve economic gains using Bitcoin trade sentiment, suggesting the challenge of transforming sentiment into daily profitable trading strategies.

Keywords: sentiment, return prediction, Bitcoin, stocks, behavioral finance.

Key Findings

- Sentiment constructed from trading information (price and trading volume) outperforms sentiment constructed from media in daily return prediction across different assets.
- Trade-based sentiment is a strong predictor of daily Bitcoin return.
- It is generally difficult to develop profitable daily trading strategies from sentiment.

Sentiment has been shown to play an essential role in explaining asset returns (Baker and Wurgler 2006, 2007; Stambaugh, Yu, and Yuan 2012) and in predicting the equity risk premium (Huang et al. 2015; Jiang et al. 2019; Tetlock 2007). The finance literature has introduced a plethora of investor sentiment measures ranging from those constructed from market-based information (Baker and Wurgler 2006; Neely et al. 2014; Zhu and Zhou 2009) to those from news and social media and surveys (Calomiris and Mamaysky 2019; Garcia 2013; Greenwood and Shleifer 2014; Obaid and Pukthuanthong 2022; Tetlock 2007). Numerous research has explored these sentiment sources individually. Still, no one has compared and contrasted them to our knowledge.

In this study, we compare the return predictability of trade-based sentiment (technical indicators) vs. text-based sentiment (news and social media) across four major asset classes (Bitcoin, stocks, Treasury bonds (T-Bond), and gold). These two types of sentiment measures are readily available daily and are widely utilized by market participants, but they are fundamentally distinct. While news and social media are publicly available and reflect people's attention and beliefs (Shiller 2005), trading information such as prices and volumes indicate investors' trading activity and decisions.

By comparing these two sources of information, we can identify which type of information is most useful in predicting market movements. Moreover, the two sentiment types may be more relevant to different types of market participants. Trade sentiment may be more relevant to institutional investors with access to trade data and can execute trades based on that data. In contrast, text sentiment may be more relevant to individual investors who rely on news articles and social media for market information. By comparing these two types of sentiment, we shed light on which types of market participants are most likely to benefit from each type of sentiment data.

Notably, we apply daily sentiment data to construct our trading strategy. Daily return is more challenging to predict than monthly returns. Fama and French (1988), Campbell and Thompson (2008), and Cochrane (2008) show that stocks are more likely to be predictable over long horizons. Cochrane (2009) shows the expected asset returns are a function of risk aversion, risk (consumption volatility), and the correlation between asset returns and the stochastic discount factor. He argues that these components are more likely to alter at a long horizon than on a short horizon like daily. In other words, asset returns are more likely to be predictable over long horizons. Since asset returns are more challenging to forecast daily, it is more complicated and difficult to developing profitable trading strategies daily. Our asset allocation results support this conjecture.

For the text-based sentiment, we use the sentiment measures constructed by Refinitiv MarketPsych Indices (RMI) based on thousands of news and social media sources. Our measures are the most comprehensive text-based indexes to date. RMI provides four sentiment series for each asset class from different text sources: news articles, headlines, social media, and news and social media combined. We further consider each series's moving averages over various lag windows. In total, we examine 28 text-based sentiment indexes for each asset.

Regarding sentiment metrics based on trading activity, we create six measures utilizing prices and trading volumes. The first is price-based WRS (William's %R), created in the 1960s and widely used in technical analysis to measure overbought/oversold (high/low sentiment) conditions. The second is NHS (nearness to a recent high) introduced by Li and Yu (2012), who finds that closeness to recent high proxies for the degree of underreaction by stock traders and thus positively predicts future returns. The third component is the trading volume ratio (TVS), one of the six components of the well-known sentiment index developed by Baker and Wurgler (2006)'s, and the only one with daily data available. The following three measures are comparable to those presented by Neely et al. (2014): MAS (a moving average rule), MOM (a momentum-type measure), and OBV (a combination of both trading volumes and prices). We evaluate different specifications based on distinct lag windows for these six trade sentiment metrics. There is no theoretical guidance regarding which construction window produces the most remarkable prediction performance. Therefore, we evaluate 28 indices of trade sentiment for each asset.

Except for Bitcoin, we utilize ETFs to represent the remaining three asset classes since we need daily prices and trading volumes for each asset to create our trade sentiment measures.[1] Using ETFs ensures that real-time investors can utilize our forecasting practices. In addition, Bitcoin is chosen to represent all cryptocurrencies because it is the largest and most popular cryptocurrency and has the most extended available time series data.

We aim to compare the return predictability of 28 trade sentiment metrics to 28 text variables for each asset. Given these many predictors, we apply four prevalent dimension reduction techniques for parsimonious comparison. Utilizing information aggregation techniques decreases overfitting and improves out-of-sample return projections. The first and most straightforward technique is the simple average of all variables (AV) (Dong et al. 2022; Huang and Lee 2010). The second technique is the combination forecast mean (CF) (Rapach, Strauss, and Zhou 2010). Under this methodology, individual projections are averaged into a single forecast. The third is principal component analysis (PCA), a frequently employed technique for extracting common variations from a group of variables. Our final technique is partial least squares (PLS), which extracts signals from a set of variables most closely associated with a prediction target. PLS has gained prominence in the literature on time series prediction (Huang et al. 2015; Kelly and Pruitt 2013, 2015). Except for CF, which is only used for out-of-sample return predictions, the other three approaches are applied, both in-sample and out-of-sample.

In our in-sample tests, trade sentiment outperforms text sentiment across prediction approaches and forecasting horizons for the four assets: Bitcoin, stocks, T-Bond, and gold. In addition, contrary to the recent literature demonstrating the strong return predictability of news sentiment (Garcia 2013; Tetlock 2007), we find that text sentiment from news, social media, and their combination have almost no predictive power over the next day's returns for all assets. Even with the use of information aggregation, the prediction of text sentiment remains dismal. We show that the inconsistent evidence is a result of the sample period. After the 2000s, for most of our sample period, text sentiment loses its significance.

We also demonstrate that sentiment has consistent predictions across asset classes. In particular, trade sentiment positively predicts future returns across asset classes. In other words, investors underreact to trade sentiment. Text sentiment presents another positive predictor than trade sentiment, but the power is much weaker.

Our out-of-sample analysis considers whether the sentiment measures can forecast real-time asset returns. We use two commonly employed out-of-sample tests, the first of which is the out-of-sample R^2 (Campbell and Thompson 2008). As a benchmark, this metric compares the recursive return predictions generated by a predictor utilizing only information known up to each time point to the historical mean return. A positive out-of-sample R^2

indicates that the predictor variable has outperformed the historical mean benchmark. We also apply the forecast-encompassing test, a more direct comparison between trade and text sentiment (Harvey, Leybourne, and Newbold 1998). This test examines the null hypothesis that a model i forecast encompasses a competing model j forecast. A failure to reject the null implies that the model j's forecast is not more informative than the model i's.

Under these two out-of-sample tests, trade sentiment generates more accurate return forecasts for Bitcoin and T-bond. Unsurprisingly, Bitcoin sentiment produces the most stable and robust out-of-sample results. Bitcoin is a highly speculative and difficult-to-value asset, and trading sentiment has been demonstrated to be highly predictive of such assets in the stock market (Baker and Wurgler 2006; Stambaugh, Yu, and Yuan 2012). However, no sentiment metric surpasses the historical mean benchmark out-of-sample in forecasting T-bonds, stocks, and gold. Forecast-encompassing tests present similar results. Trade sentiment encompasses text sentiment clearly in both short and long terms for Bitcoin, but only for long-term predictability for T-bond. There is no clear pattern for stocks and gold.

Finally, we investigate whether the daily return estimates generated by sentiment metrics can provide economic value to real-time mean-variance investors. Strikingly, Bitcoin is the only asset whose trade sentiment can achieve remarkable economic returns and a higher Sharpe ratio than a buy-and-hold investment strategy. Both trading and text sentiment cannot surpass the historical mean benchmark or the simple buy-and-hold strategy for the remaining assets. This finding demonstrates that converting technical indicators or media opinions into profitable trading strategies for stocks, bonds, and gold is intricate. Refinitiv MarketPsych Indices (RMI) is the largest and most influential sentiment provider together with Ravenspack. Despite their seemingly wide use by investors, we find little value in using RMI's sentiment measures to create profitable daily trading strategies. This is the implication of our finding.

Our paper offers two significant contributions to the finance literature. First, we respond to Zhou (2018) by conducting the first exhaustive analysis on the comparative return predictability of trade and text sentiment. While the enormous literature on investor sentiment has studied trade and text sentiment individually, our study combines several sentiment sources and fills a gap in the literature. In contrast to the existing studies that primarily focus on equities, we analyze four distinct asset types. We find that sentiment influences on future returns vary across asset classes. However, regarding return predictability, technical indicators outperform news and social media tone. This shows that text-based sentiment appears too noisy, diminishing its ability to anticipate future returns. Our findings extend the Cheap Talk model proposed by Stein (1989), Farrell (1995), and Farrell and Rabin (1996), which is mainly applied to the Fed announcements. We demonstrate that talk is cheap; trading sentiment outperforms text sentiment in both in-sample and out-of-sample.

We also contribute to the growing literature on cryptocurrencies and Bitcoin in particular. Financial economists are still attempting to ascertain the worth of digital assets. Theoretically, as demonstrated by Dwyer (2015), Athey et al. (2016), and Pagnotta and Buraschi (2018), the value of cryptocurrencies depends on a combination of usage, degree of acceptance, scarcity, and the importance of anonymity. However, empirically assessing digital assets is extremely difficult because fundamental metrics such as cash flows to stocks are missing. Since cryptocurrencies are more difficult to value than stocks, investor sentiment can explain returns on stocks that are difficult to value and costly to arbitrage (see, for example, Baker and Wurgler (2006), and Zhou (2018) and references therein); cryptocurrencies appear more susceptible to sentiment than stocks. Recent research by Detzel et al. (2021) examines the prediction of Bitcoin using technical indicators. We extend their study by comparing trade sentiment to text sentiment and demonstrate that trading sentiment is significantly more predictive of future Bitcoin returns than text sentiment. Moreover, real-time investors can use Bitcoin trade sentiment to achieve substantial economic benefits.

DATA

Market Data

In this paper, we study the return predictability of sentiment across four asset classes. The first one is Bitcoin. We collect daily Bitcoin prices and trading volumes from bitcoinity.org.[2] Unlike stocks, Bitcoin is traded every day, and thus the prices and volumes are available seven days a week. As of October 2022, Bitcoin has a market capitalization of $391B.

Because we need both prices and trading volumes to construct our trade sentiment measures and investigate the trading strategies of all assets based on their sentiment signals, we use ETF trading data on the remaining three asset classes:

- Stocks: SPY, the world's first and largest ETF tracking the S&P 500 index since 01/22/1993 (having net asset value (NAV) of $326B as of October 2022);
- Long-term Treasury bonds: TLT, an iShares Treasury Bond ETF, tracks the investment results of an index composed of U.S. Treasury bonds with remaining maturities greater than twenty years since 07/22/2002 (having a NAV of $24B as of October 2022); and
- Gold: GLD, an SPDR gold ETF tracking the price of gold since 11/18/2004 (having a NAV of $50B as of October 2022).

Text-Based Sentiment Data

Our text-based sentiment measures are from Refinitiv MarketPsych Indices (RMI) managed under the umbrella of Refinitiv. The indices are constructed from news and social media content and identify specific sentiment, macroeconomic, and general buzz-related words relevant to the entity. Subsequently, the volume and tone of phrases and words are converted into measurable variables. In short, RMI sentiment measures are based on various news media and are computed from overall positive references net of negative references of each asset class. Michaelides, Milidonis, and Nishiotis (2019) and Michaelides et al. (2015) also use RMI.[3] For each asset, we have four RMI sentiment measures constructed from different content sources: whole news articles (hereafter News), only news headlines (NewsHL), social media (Social), and news and social media combined (NewsSoc).

RMI indexes are updated either hourly or daily at 3:30 p.m. EST using data from the past 24 hours as their construction period. Because Bitcoin is traded continuously, the daily data for Bitcoin reports the price at midnight UTC as its closing price. Consistent with this convention, we use the RMI indexes for Bitcoin updated at midnight UTC. Specifically, our prediction analyses use Bitcoin sentiment measures computed from midnight UTC on day t to midnight UTC on $t+1$ to forecast Bitcoin returns measured from midnight UTC on the day $t+1$ onward.

We use the default RMI sentiment measures for all other assets, updated daily at 3:30 p.m. EST. According to RMI, this timestamp is selected so investors have enough time to update their investments based on the sentiment signals before the NYSE market closes. During the weekends or national holidays when the stock market is closed, we average the sentiment scores over these days with the score on the day when the market is reopened (i.e., the forward average). For example, sentiment scores over Saturday and Sunday are combined with those on Monday to predict returns on Tuesday and beyond. We adopt this timing choice to ensure that our sentiment measures do not overlap with returns, which is especially important in out-of-sample analyses. Continuing with the above example, if sentiment scores on Saturday and Sunday are combined with sentiment scores on Friday (i.e., backward average) to predict Monday returns computed from closing prices on Monday and last Friday, sentiment and return are overlapped, resulting in information leakage in out-of-sample prediction.

It should be noted that only English-language texts were used for constructing RMI sentiment measures before February 2020, which might cause bias in countries where English is not used as an official language and articles are not primarily written with it. However, this bias is partly alleviated because even in non-English countries, the most critical business, finance, and economics-related news are often published in English. Moreover, informed traders will likely post significant news and events internationally in English. Since February 2020, RMI's real-time translation and analysis engine has included Arabic, Chinese, Japanese, and Portuguese-language news sources. In January 2021, Dutch, French, German, Indonesian, Italian, Korean, and Spanish language sources were added.

Trade-Based Sentiment Construction

Based on trading data, we construct sentiment measures analogous to the most famous investor sentiment index of Baker and Wurgler (2006, 2007). In constructing their index, Baker and Wurgler (2006) use six proxies that are unavailable daily for our assets. Consistent with their study, we define TVS, one of our six sentiment measures, based on trading volume. Mathematically, the log trading volume ratio over the past L days is computed as

$$\mathrm{TVS}_t^i(L) = \log\left(\frac{\mathrm{TV}_t^i}{\mathrm{TV}_{t-L+1}^i}\right) \tag{1}$$

where TV_t^i is the trading volume of asset i on day t. A higher trading volume indicates greater sentiment.

The second trade sentiment measure is based on the most famous over-bought and over-sold technical analysis indicator, Williams' %R (WRS). The indicator was developed by Larry Williams in the 1960s and was popularized by his book (Williams 1976). Mathematically, we define

$$\mathrm{WRS}_t^i(L) = \frac{P_{max,t}^i(L) - P_t^i}{P_{max,t}^i(L) - P_{min,t}^i(L)} \tag{2}$$

where $P^i_{max,t}(L)$ and $P^i_{min,t}(L)$ are the highest and lowest daily prices of asset i over the window from day $t-L+1$ to day t. Intuitively, if $\mathrm{WRS}^i_t(L)$ is less than 20%, asset i is regarded as over-bought (high-sentiment) as the trading price P^i_t is closest to its highest price. Likewise, when $\mathrm{WRS}^i_t(L)$ is greater than 80%, asset i is regarded as over-sold (low-sentiment). The measure is very obvious when the asset price oscillates around a certain price level.

As a low value of $\mathrm{WRS}^i_t(L)$ indicates high sentiment, it has the opposite sign compared to all other sentiment measures. To remain consistent and streamline our empirical analyses, we define a new $\mathrm{WRS}^i_t(L)$ as the negative of $\mathrm{WRS}^i_t(L)$ in (2).

Our third sentiment measure is based on only the current and max prices,

$$\mathrm{NHS}^i_t(L) = \frac{P^i_t}{P^i_{max,t}(L)} \tag{3}$$

Li and Yu (2012) interpret it as the nearness to the recent high and seem the first to use it in the stock market. It is clear that $\mathrm{NHS}^i_t(L)$ is closely related to $\mathrm{WRS}^i_t(L)$. However, they are different. $\mathrm{WRS}^i_t(L)$ depends on the minimum price and measures the incremental price move from the minimum. In contrast, $\mathrm{NHS}^i_t(L)$ is independent of the minimum price and measures the closeness of P^i_t to $P^i_{max,t}(L)$.

Note that no theories or economic intuitions guide the choice of L. As with studies in technical analysis, such as Brock, Lakonishok, and LeBaron (1992), we consider several plausible choices, with $L = 5, 10, 20, 50, 100$, respectively. While it will then be difficult to select which L works the best in the future, the data can tell whether all the sentiment measures collectively can forecast returns out-of-sample. In our empirical results, we denote the $\mathrm{WRS}^i_t(L)$ and $\mathrm{NHS}^i_t(L)$ computed with a specific value of L as WRSL and NHSL, respectively.

Our subsequent three sentiment measures are those used in Neely et al. (2014). The first is a moving average (MA) rule that generates a buy or sell signal ($\mathrm{MAS}^i_t = 1$ or $\mathrm{MAS}^i_t = 0$, respectively) for asset i on day t by comparing two moving averages:

$$\mathrm{MAS}^i_t(s,l) = \begin{cases} 1 & \text{if} \quad \mathrm{MA}^i_{s,t} \geq \mathrm{MA}^i_{l,t}, \\ 0 & \text{if} \quad \mathrm{MA}^i_{s,t} < \mathrm{MA}^i_{l,t} \end{cases} \tag{4}$$

where

$$\mathrm{MA}^i_{j,t} = \frac{1}{j} \sum_{k=0}^{j-1} P^i_{t-k} \quad \text{for} \quad j = s, l;$$

P^i_t is the price index of asset i on day t, and $s(l)$ is the length of the short (long) MA ($s < l$). Intuitively, the MA rule detects changes in price trends because the short MA will be more sensitive to recent price movements than the long MA. In our study, we consider daily MA rules with $s = 10, 20$ and $l = 50, 100$ so that we have four measures of MAS for each asset i. Our results denote a MAS constructed with a specific value of s and l as MASs_l.

Our fifth measure is a momentum-based strategy. A simple momentum rule generates the following signal:

$$\mathrm{MOM}^i_t(L) = \begin{cases} 1 & \text{if} \quad P_t \geq P_{t-L+1}, \\ 0 & \text{if} \quad P_t \geq P_{t-L+1}. \end{cases} \tag{5}$$

Intuitively, a current price higher than its level L days ago indicates "positive" momentum, generating a buy signal. In this study, for each asset i, we compute MOM_t^i for $L = 5, 10, 20, 50, 100$ and denote them as MOML.

Our final trade sentiment measure combines prices with trading volumes. More specifically, we first define

$$V_t^i = \sum_{k=1}^{t} \mathrm{VOL}_k^i \mathrm{D}_k^i, \tag{6}$$

where VOL_k^i is the trading volume of asset i on day k and D_k^i is an indicator equal to 1 if $P_k - P_{k-1} \geq 0$ and -1 otherwise. We then form a trading signal from V_t^i as follows

$$\mathrm{OBV}_t^i(s,l) = \begin{cases} 1 & \text{if} \quad \mathrm{MA}_{s,t}^{V,i} \geq \mathrm{MA}_{l,t}^{V,i}, \\ 0 & \text{if} \quad \mathrm{MA}_{s,t}^{V,i} < \mathrm{MA}_{l,t}^{V,i} \end{cases} \tag{7}$$

where

$$\mathrm{MA}_{j,t}^{V,i} = \frac{1}{j} \sum_{k=0}^{j-1} V_{t-k}^i \quad \text{for} \quad j = s, l;$$

Intuitively, relatively high volume and recent price increases indicate a strongly positive market trend and generate a buy signal. Similar to MAS, we consider $s = 10, 20$ and $l = 50, 100$ for OBV and denote them as OBVs_l.

Text-Based Sentiment Construction

Because we have 28 trade sentiment measures in total, we increase the number of RMI sentiment measures by computing various moving averages for each sentiment measure to provide a fair playground between trade and text sentiment. Specifically,

$$x_t^i(L) = \frac{1}{L} \sum_{k=0}^{L-1} x_{t-k}^i,$$

where x_t^i is one of News, NewsHL, Social, and NewsSoc sentiment measures for asset i on the day t. Similar to trade sentiment measures, we consider $L = 5, 7, 10, 20, 50, 100$ and denote them as News$L$, NewsHLL, Social$L$, and NewsSoc$L$ for a specific choice of L in our empirical results. Thus, we have 28 text-based sentiment measures for each asset i.

Considering the availability of both trade and text sentiment measures and using the first 100 days to construct various sentiment measures ($L = 100$ in our setups), our final sample period for stocks, bonds, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively, to September 2022.[4]

Sentiment Orthogonalization

Since our goal is to compare the true return predictability of sentiment measured from textual and trading information, we want to make sure other market forces do not impact the return predictability. To accomplish this goal, we follow the literature (Garcia 2013; Huang et al. 2015) to orthogonalize our sentiment measures before using them in the forecasting exercise.

Specifically, following Garcia (2013), we use the residuals in the following regression as a measure of our orthogonalized sentiment:

$$\begin{aligned} x_t^i = &\big(1 - D_t\big)\big(\boldsymbol{\beta_1 L_{0-5}}(r_t^i) + \boldsymbol{\gamma_1 L_{0-5}}(r_t^{2,i}) + \boldsymbol{\delta_1 L_{1-5}}(x_t^i)\big) \\ &D_t\big(\boldsymbol{\beta_2 L_{0-5}}(r_t^i) + \boldsymbol{\gamma_2 L_{0-5}}(r_t^{2,i}) + \boldsymbol{\delta_2 L_{1-5}}(x_t^i)\big) + \boldsymbol{\eta} Z_t + \epsilon_t^i \end{aligned} \tag{8}$$

where x_t^i is a sentiment of asset i on day t, $\boldsymbol{L_{0-5}}(r_t^i)$ are returns of asset i on days t to $t-5$, $\boldsymbol{L_{0-5}}(r_t^{2,i})$ are squared returns of asset i on days t to $t-5$, $\boldsymbol{L_{1-5}}(x_t^i)$ are five lags of sentiment x_t^i, D_t is a dummy variable equal to 1 (0) if day t belongs to a recession (expansion) as defined by NBER, and Z_t is a list of control variables including a constant and weekday indicators.

Except in Exhibit 2 where we report the contemporaneous correlations between the raw sentiment measures and asset returns,[5] we use the orthogonalized version in the remaining empirical analyses. In our out-of-sample tests, we recursively orthogonalize sentiment variables using data up to the estimation time to avoid the look-ahead bias. Since an estimation window may not contain any recessionary days, we exclude the recession dummy variable from (8) in the out-of-sample analyses.[6]

METHOD

Predictive Regression

Following most studies, we analyze the return predictability of sentiment via the standard predictive regression model,

$$r_{t+1\to t+h}^i = \alpha_j^i + \beta_j^i \times x_{j,t}^i + \epsilon_{t+1\to t+h}^i, \tag{9}$$

where $r_{t+1\to t+h}^i$ is the normal return of asset i over days from $t+1$ to $t+h$, $x_{j,t}^i$ is a sentiment j for asset i on day t, and $\epsilon_{t+1\to t+h}^i$ is the residual term. β_j^i measures the strength of predictability. In our empirical results, we report its significance based on the Newey and West (1987)'s standard error with h lags (if $h < 5$, we set lag to 5).

We generate out-of-sample forecasts of asset returns using an expanding estimation window. Given an initial estimation window length W, the out-of-sample forecasts of $r_{t+1\to t+h}^i$ is given by

$$\hat{r}_{t+1\to t+h}^i = \hat{\alpha}_{j,t}^i + \hat{\beta}_{j,t}^i \times x_{j,t}^i, \tag{10}$$

where $\hat{\alpha}_{j,t}^i$ and $\hat{\beta}_{j,t}^i$ are estimates of α_j^i and β_j^i in equation (9) respectively by using data up to time t for $t = W, W+1, \ldots, T-h$ with T being the sample size. We repeat this procedure until the end of the out-of-sample period. Our empirical analyses use an initial estimation window of 750 days for all assets, i.e., $W = 750$.

Information Aggregation

As discussed above, this paper compares the return predictability of 28 trade sentiment variables against 28 text sentiment measures for each asset. We employ several commonly used dimension reduction methods to provide a parsimonious comparison between these two groups of sentiment variables. Dimension reduction also helps reduce overfitting, resulting in better out-of-sample prediction performance.

The first and most simple dimension reduction method is the average variable (AV). This AV technique is used in Dong et al. (2022) to combine a large set of returns on long-short anomalies into a strong aggregate market predictor. Under this method, at each time point, we compute the cross-sectional average of all standardized variables within a group of predictors, i.e.

$$\mathrm{AV}_{g,t}^{i} = \frac{1}{n_g} \sum_{j=1}^{n_g} x_{j,t}^{i}, \tag{11}$$

where $n_g = 28$ and $x_{j,t}^{i}$ is one of the trade or text sentiment variables for asset i and sentiment group j.

In our out-of-sample forecasts, we also use the combination forecast (CF) method as in Rapach, Strauss, and Zhou (2010), which shows that this is a powerful filter to reduce noises in forecasts, leading to more accurate predictions. Under this method, at each point in the out-of-sample window, we compute the average of all return forecasts, i.e.

$$\bar{r}_{g,t+1\to t+h}^{i} = \frac{1}{n_g} \sum_{j=1}^{n_g} \hat{r}_{j,t+1\to t+h}^{i}, \tag{12}$$

where $\hat{r}_{j,t+1\to t+h}^{i}$ is a return forecast for asset i made by predictor j as in equation (10).

Our third method is principal component analysis (PCA), a widely used tool for dimension reduction, reducing the set of factors to a smaller set of components that explains most of those factors' variance. Suppose we want to extract k principal components from p factors. We want to reduce the number of explanatory variables from p to k, $k \leq p$. The k^{th} principal component PC_k is a normalized linear combination of the p factors: $PC_k \equiv Xv_k$ where X is the $T \times p$ centered matrix of factors and the $p \times 1$ vector v_k solves

$$\begin{aligned} \max_{v} \quad & \mathrm{Var}(Xv) \\ \text{s.t.} \quad & \|v\| = 1 \quad \text{and} \\ & (Xv)'(Xv_l) = 0, \quad l = 1, \ldots, k-1. \end{aligned} \tag{13}$$

From the objective function, it is easy to see that the first principal component PC_1 indicates the direction of the most significant sample variance in the column space of X; the second principal component PC_2 suggests the direction of the second largest sample variance among all normalized linear combination of the p factors; so on and so forth. The second constraint $(Xv)'(Xv_l) = 0$ ensures the k^{th} principal component is uncorrelated to all previous $k-1$ principal components. An advantage when applying PCA is that it does not require any prior information about the rate of returns to be explained. Following the standard practice, we standardize all predictors to zero mean and unit variance before implementing PCA.

Our final method is partial least squares (PLS), which has gained popularity in the return prediction literature (Huang, Li, and Wang 2020; Huang et al. 2015; Kelly and Pruitt 2013, 2015). Like PCA, PLS is also a dimension reduction method by constructing a (smaller) set of linear combinations of original factors. But unlike PCA, which focuses only on the components' high variance, PLS transforms the factors to capture both the high variance of the extracted components and the high correlation with asset returns to be predicted.

Mathematically, the k^{th} PLS component is $PLS_k \equiv X\theta_k$ where θ_k solves

$$\begin{aligned} \max_{\theta} \quad & \text{Corr}^2(y, X\theta)Var(X\theta) \\ \text{s.t.} \quad & \|\theta\| = 1 \quad \text{and} \\ & (X\theta)'(X\theta_l) = 0, \quad l = 1, \ldots, k-1, \end{aligned} \tag{14}$$

where y is the $T \times 1$ vector of returns to be predicted. When the variance term in the objective function dominates, PLS will behave much like PCA. However, when the correlation between X and y is strong, we expect PLS to perform better because it incorporates such information in dimension reduction. At the same time, PCA ignores the correlation and might return components with a high variance but little covariance with y.

Out-of-Sample R^2

We use two measures to evaluate the performance of out-of-sample forecasts: the out-of-sample R^2 (Campbell and Thompson 2008) and the realized utility gains for a risk-averse investor in the mean-variance world (Campbell and Thompson 2008; Rapach, Strauss, and Zhou 2010).

The out-of-sample R^2, R^2_{OS}, evaluates the performance of out-of-sample forecasts of asset return $\hat{r}^i_{t+1\to t+h}$ against its historical average $\bar{r}^i_{t+1\to t+h}$ within the estimation window, with $\bar{r}^i_{t+1\to t+h} = \frac{1}{t-h}\sum_{s=1}^{s=t-h} r^i_{s+1\to s+h}$. More specifically, the R^2_{OS} statistics is given by

$$R^{i,2}_{OS}(h) = 1 - \frac{\sum_{t=W}^{t=T-h}(r^i_{t+1\to t+h} - \hat{r}^i_{t+1\to t+h})^2}{\sum_{t=W}^{t=T-h}(r^i_{t+1\to t+h} - \bar{r}^i_{t+1\to t+h})^2}, \tag{15}$$

If $R^2_{OS} > 0$, the constructed forecast $\hat{r}^i_{t+1\to t+h}$ beats the historical average forecast since $\hat{r}^i_{t+1\to t+h}$ has smaller mean squared prediction error (MSPE) than the latter does. We examine the statistical significance of the R^2_{OS} with the Clark and West (2007) adjusted MSPE test.

Utility Gains

Even though R^2_{OS} is widely used in forecast evaluation, it has its limitation: it does not consider investors' risk preferences. It, therefore, does not directly speak to the economic significance of the forecasts. To address this, we also calculate the average utility gain of a risk-averse investor in a mean-variance world (Campbell and Thompson 2008). After forming a forecast of the next period's return on risky asset i, at the end of period t, a mean-variance investor can decide how to allocate the total wealth between a risk-free asset and a risky asset. The weight on asset i, $\hat{w}^i_{t+1}$, is decided according to the following formula

$$\hat{w}^i_{t+1} = \frac{1}{\gamma}\left(\frac{\hat{r}^i_{t+1} - r^f_{t+1}}{\hat{\sigma}^{2,i}_{t+1}}\right), \tag{16}$$

where r^f_{t+1} is the risk-free rate on the day $t+1$, $\hat{r}^i_{t+1}$ is the forecast of the return on asset i on the day $t+1$, $\hat{\sigma}^{2,i}_{t+1}$ is a 20-day rolling window estimate of the variance of the excess return, $r^i_{t+1} - r^f_{t+1}$, and γ is the coefficient of risk aversion. Our empirical results consider γ

equal to 3 and 5. Following Campbell and Thompson (2008), we constrain $\hat{w}^i_{t+1}$ between 0 and 1.5 to exclude short-selling and limit the portfolio leverage to only 50%.

The realized rate of return of this portfolio on the day $t+1$ is

$$r^{i,p}_{t+1} = \hat{w}_{t+1} r^i_{t+1} + (1 - \hat{w}^i_{t+1}) r^f_{t+1}. \tag{17}$$

The realized average utility of this risk-averse investor exploiting such a portfolio strategy is calculated as follows.

$$\hat{u}^i = \hat{\mu}^i - \frac{1}{2}\gamma \hat{\sigma}^{2,i}, \tag{18}$$

where $\hat{\mu}^i$ and $\hat{\sigma}^{2,i}$ are the sample mean and sample variance of the realized return $r^{i,p}_{t+1}$, respectively, of the constructed portfolio over the out-of-sample period.

The evaluation benchmark is still the historical average forecast. If using the historical average forecast, this same investor will allocate $\bar{w}^i_{t+1}$ portion of total wealth to asset i according to the following formula

$$\bar{w}^i_{t+1} = \frac{1}{\gamma}\left(\frac{\bar{r}^i_{t+1} - r^f_{t+1}}{\hat{\sigma}^{2,i}_{t+1}}\right), \tag{19}$$

where $\bar{r}^i_{t+1}$ is the historical mean forecast of return. The realized rate of return of this portfolio is

$$r^{i,p}_{t+1} = \bar{w}_{t+1} r^i_{t+1} + (1 - \bar{w}^i_{t+1}) r^f_{t+1}. \tag{20}$$

The realized average utility of this risk-averse investor using a historical average forecast is given by

$$\bar{u}^i = \bar{\mu}^i - \frac{1}{2}\gamma \bar{\sigma}^{2,i}. \tag{21}$$

The utility gain is measured as the difference between $\hat{u}^i$ and $\bar{u}^i$

$$\text{Utility Gain} = 25000 \times (\hat{u}^i - \bar{u}^i) \tag{22}$$

We multiply this difference by 25000 to report it in the average annualized percentage return, assuming 250 trading days in a year. In addition to the utility gain, we present the annualized Sharpe ratio of the realized portfolio returns $r^{i,p}_{t+1}$. We examine the statistical significance of the utility gains and investigate whether the Sharpe ratio of the portfolio using the predictive model is greater than that of the portfolio using the historical mean forecasts with the tests in DeMiguel, Garlappi, and Uppal (2009).

EMPIRICAL RESULTS

Contemporaneous Correlations

Before presenting the empirical prediction results, we evaluate the contemporaneous correlations between our trade sentiment and text sentiment variables and asset returns. As sentiment measurements, we anticipate a favorable correlation between them and returns.

[Insert Exhibit 2 here.]

Panel A of Exhibit 2 displays the results for measures of trade sentiment. Notable is that while aberrant trading volume (TVS) is connected with higher returns for Bitcoin, it is inversely correlated with returns on the remaining assets, with stocks exhibiting the most substantial negative associations. Only Bitcoin and T-bond returns are favorably connected with sentiment indicators based on moving averages of prices (MAS) and a combination of prices and trading volumes (OBV). Three price-based sentiment groups (WRS, NHS, and MOM) exhibit a substantial positive association with the returns of all assets.

Panel B presents the concurrent connections between text sentiment and returns. The link between sentiment from news headlines and asset returns is the greatest among the four sources of text sentiment. Gold has the strongest positive correlation with returns, at 18%, and is the asset most affected by text sentiment. On the other hand, Treasury bonds exhibit only positive relationships between news headline sentiment and their returns.

As expected, Exhibit 2 shows that trade and text sentiment measures positively correlate with asset returns.

In-Sample Predictions with Individual Sentiment Measures

Next, we examine the in-sample return predictability of sentiment variables across asset classes. We analyze the returns forecast for the following day ($h = 1$) and 20 days ($h = 20$).

[Insert Exhibit 3 here.]

Panel A of Exhibit 3 displays the results for measures of trade sentiment. As Bitcoin is a highly speculative asset, it is unsurprising that measures of trade sentiment, including WRS and MOM, can positively predict Bitcoin returns over the next 1 to 20 days. While OBV and MAS yield good predictions for the long term, NHS and TVS variables display the weakest results. Our findings indicate a momentum pattern in Bitcoin return forecasts: high trading sentiment today portends high returns over the following 20 days.

T-bonds exhibit a momentum-like relationship between trade sentiment and long-term returns. While the prediction coefficient for next-day returns is negative for TVS, WRS, and NHS, it is positive for MAS and OBV.

The relationship between trade sentiment and future returns in the stock market is similar: high trade sentiment is generally associated with higher returns over the following 1 to 20 days. WRS provides strong stock prediction results over the next 1 to 20 days. MOM also provides strong short-term predictive power. There are some exceptions such as TVS5, NHS10, and OBV20_50 which show negative coefficients. Lastly, gold is almost unpredictable based on indicators of trade sentiment, except for WRS and MO, which can modestly positively forecast gold returns over a 20-day horizon.

Panel B displays the results of return prediction with text sentiment. Surprisingly, news and social media tone measurements fail to predict Bitcoin and gold returns. We observe the prediction results for Treasury bonds and stocks, but they are not strong.

At first glance, it appears odd that news and social media tone cannot forecast market returns, given the literature's overwhelming evidence of predictability (Garcia 2013; Tetlock 2007). We use the news sentiment measure developed by Garcia (2013) and published on the author's website to resolve this discrepancy.[7] We apply the predictive regression specification from Tetlock (2007) and Garcia (2013) to three sample periods: 1905 to 2005, 1984 to 1999, and 2000 to 2005. Following Tetlock (2007) and Garcia (2013), we utilize five lags of sentiment and control for five lags of returns, five lags of squared returns, and weekday indicators for this regression and show the results in Exhibit 8. Accordingly, we can replicate Tetlock (2007) and Garcia (2013) findings that negative (positive) sentiment is a substantial negative (positive) predictor of the next day's return. Nonetheless, over the sample 2000-2005, the predictability of news sentiment nearly vanishes, corresponding with the findings shown in Exhibit 3. This finding suggests that news tone has been fully integrated into stock prices on the same day during the past two decades, rendering its future predictability null and void.

Overall, individual trade sentiment measures are significantly more accurate in predicting future returns on all assets than individual text sentiment measures. Indicators of trade sentiment exhibit momentum for all assets. Meanwhile, text sentiment is a positive/negative predictor for short-term returns on T-bonds/stocks.

In-Sample Predictions with Aggregate Sentiment Measures

Examining univariate regressions with individual sentiment variables, the preceding section demonstrates that trade sentiment measures can predict returns better than text sentiment measures. In this section, we formally compare trade and text sentiment measurements. We begin by developing indexes that summarize the differences between these two categories of sentiment indicators. As stated in the Method section, our first and most straightforward technique is to compute the cross-sectional means of trade and text sentiment variables separately (AV). Our more advanced methods involve creating the PCA and PLS indexes.

[Insert Exhibit 1 here.]

Exhibit 1 illustrates each asset's PCA and PLS weights. For Bitcoin, WRS and MOM measures dominate the PLS weight. This is expected given the results in Exhibit 3 because PLS weights are scaled correlations between sentiment measures and next one-day returns. PCA weights are spread throughout all trading sentiments. For text sentiment, PCA weights are distributed through all sentiments, while NewsHL dominates the PLS weights.

Except for TVS, PCA weights are distributed equally across all trade sentiment techniques for T-bonds. Only MAS, OBV, and MOM have positive PLS weights. For text, PLS weights are concentrated in News, NewsHL, and NewsSoc. Only TVS measures have positive weights for stocks, yet they are low. While PCA weights are distributed equally across measures for text sentiment, News and NewsHL dominate the PLS weights.

For gold, PLS weights are dominated by TVS and MOM. Interestingly, all trade sentiments have negative PCA weight. For text, PCA weights are spread across text sentiments, whereas PLS weights are negative for all except NewsHL.

After creating our AV, PCA, and PLS indexes from trade and text sentiment measurements for each asset, we employ them in univariate predictive regressions (9) and bivariate predictive regressions using one trade sentiment with one text index. As there is no restriction on the sign of each component in PCA and PLS, we reverse the sign of the aggregate PCA and PLS indexes when appropriate to ensure that the trade (text) PCA and PLS indexes are always positively associated with the trade (text) average-variable index. Since a high value of each sentiment measure indicates a high level of sentiment, their averages (i.e., AV indexes) are still sentiment. Thus, the PCA and PLS indices can still be regarded as a measure of sentiment when inverted.

[Insert Exhibit 4 here.]

For Bitcoin in Panel A of Exhibit 4, all three measures of aggregate trade sentiment predict Bitcoin returns over the following 1 to 20 days, whether examined alone or in conjunction with the text sentiment variables. In contrast, the text sentiment indexes fail to forecast future returns when used alone or with trade sentiment, as seen in Exhibit 3.

The results for Treasury bonds in Panel B are comparable to those for Bitcoin in that trade sentiment indicators are positive return predictors, especially for the next five days. Text sentiment measurements predict future returns only for the next day under AV and PCA. Under PLS, trade sentiment can predict returns better than text.

Panel C presents the stock forecasting results. Trade sentiment indicators can positively predict future returns, with shorter periods producing more substantial impacts. When employed alone, text sentiment indexes such as AV and PCA cannot forecast future returns. Still, they become positive predictors (significant at 10%) and increase the statistical significance of trade indices when the two groups are compared. When predicting future stock returns, trading sentiment outperforms text sentiment, exhibiting a reversal prediction pattern for AV and PCA but the positive pattern for PLS.

Under PLS, trade sentiment is a positive predictor of future returns for gold in Panel D. When compared, trade sentiment predicts the following one to twenty days with more accuracy than text sentiment.

In conclusion, trading sentiment surpasses text sentiment in predicting future returns for all four assets. Trade sentiment exhibits strong and persistent predictive power for Bitcoin. Text sentiment works only for stocks, yet the result is weak.

Out-of-Sample R^2

In the previous section, we discover that trade sentiment measures predict returns more accurately than text sentiment measures across all four asset classes. We expand our empirical analysis to an OOS setting by examining the common metric R^2_{OS}. Out-of-sample prediction is more complicated than in-sample predictability, and most economic forecasters have failed to predict out-of-sample returns (Goyal and Welch 2008). In addition to the three dimension reduction strategies employed in the in-sample tests, we consider the combination forecast (CF), wherein we cross-sectionally average the forecasts generated by individual trade and text sentiment variables.

[Insert Exhibit 5 here.]

Panel A of Exhibit 5 demonstrates that trade sentiment measurements under all four reduction procedures produce strong R^2_{OS} with a one-day Bitcoin return. For instance, the one-day R^2_{OS} ranges between 0.18% and 1.57% for trade PCA and AV, respectively. Except for trade PCA, the significant outcomes continue to the 20-day horizon. In contrast, none of the text-reduction techniques can accurately predict Bitcoin returns out-of-sample. These results confirm the in-sample findings that only trade sentiment metrics apply to Bitcoin.

For Treasury bonds in Panel B, trade sentiment gives better out-of-sample forecasts than text sentiment, consistent with the in-sample results. Trade sentiment generates significantly favorable OOS R^2 values for the subsequent five-day prediction under CF and PCA. Only the text sentiment PCA indexes generate a positive OOS R^2.

For stocks in Panel C and gold in Panel D, neither trading nor text sentiment generates a substantial OOS R^2 over any horizon, regardless of the time horizon. As reported in Goyal and Welch (2008) and Goyal, Zafirov, and Welch (2022), this exemplifies the difficulty of accurately predicting asset returns out-of-sample.

Overall, out-of-sample results still favor Bitcoin and Treasury bond trade sentiment measures. No sentiment metric exceeds the historical mean benchmark at the 5% significant level or greater for out-of-sample stock and gold predictions.

Forecast Encompassing Tests

While the previous subsection provides empirical support for trade sentiment measures over text sentiment measures in out-of-sample forecasts for Bitcoin and Treasury bonds, such evidence is indirect because when using the R^2_{OS} metric, we compare the forecasts made by trade sentiment and text sentiment against the standard historical mean benchmark. In this part, we directly compare the informational content of trade forecasts with text sentiment forecasts.

Following Rapach, Strauss, and Zhou (2010), we employ the modified version of the test statistic created by Harvey, Leybourne, and Newbold (1998) (MHLN). This MHLN test tests the null hypothesis that the model i forecast encompasses the model j forecast to the alternative hypothesis that the model i forecast does not encompass the model j prediction. We have a total of eight return projections for each asset (four dimension reduction methods applied to trade sentiment and four methods applied to text sentiment). Thus, we examine if a prediction method's forecast incorporates those of the remaining seven prediction methods. We conduct this test for the 1-day and 20-day forecasts.

[Insert Exhibit 6 here.]

Exhibit 6 presents the p-values for testing the null hypothesis that the model's forecast in a column encompasses the model's forecast in the corresponding row. The forecast-encompassing test (Chong and Hendry 1986) implies that the predictor that encompasses another offers information regarding potential future movements of the testing asset that is absent from the latter. We emphasize the entries where we fail to reject the null hypothesis.

In Panel A, except trade PLS, all of Bitcoin trade sentiment measures encompass text sentiment measures over a one-day horizon ($h = 1$). With the 20-day return projection, except trade PCA and PLS, the other two trade sentiment measures encompass text sentiment measures again. This result is consistent with the R^2_{OS} mentioned in the preceding section. Text sentiment encompasses trade only for the PLS index.

The results for Treasury bonds in Panel B indicate that text sentiment encompasses trade sentiment over the short term, but the relation reverts for the twenty-day horizon.

For stocks in Panel C, trade and text sentiment predictions do not encompass each other. Regarding gold in Panel D, trade and text sentiment estimates overlap under all approaches except PLS for the following one-day horizon. In contrast, trade and text sentiment estimates do not overlap over the long term. The results indicate that trade and text sentiment forecasts include valuable information regarding future returns.

In short, we discover that, for Bitcoin the projections from trade sentiment measures encompass those from text sentiment measures for both short and long horizons. For Treasury bonds, the same pattern appears only for the long term. These results are consistent with prior findings in that trade sentiment measurements produce more prominent (and significant) R^2_{OS}'s for Bitcoin and government bonds. For future stock returns, combination forecasts based on trade sentiment provide the most valuable information, whereas, for gold, the combination helps for a long-term forecast.

Utility Gains

This final empirical investigation examines whether sentiment measures may produce effective trading strategies. We estimate the utility gains and Sharpe ratio for a mean-variance investor that allocates a portfolio between a risky and a risk-free asset based on the prediction method's *one-day return forecast* for each asset. The portfolio based on the historical mean forecast is the benchmark for utility gains. For comparison, we also include the Sharpe ratio for a buy-and-hold strategy.

[Insert Exhibit 7 here.]

The most striking result from Exhibit 7 is that, except for Bitcoin (Panel A), it is not easy to construct daily profitable trading strategies using sentiment measures, both trade and text. Trade sentiment measures under all four prediction models deliver impressive economic gains for Bitcoin. For example, with the risk coefficient of three ($\gamma = 3$), trade PLS delivers a superb annualized utility gain of 50% with an annualized Sharpe ratio of 1.73, much higher than the buy-and-hold Sharpe ratio of 0.76. On the other hand, text sentiment measures under all prediction models do not generate any economic gains for a mean-variance investor.

No sentiment measure can generate economic values over the historical mean for Treasury bonds, stocks, and gold in Panels B, C, and D and for bonds and gold, none of them produces a Sharpe ratio exceeding the buy-and-hold strategy. This corresponds to the OOS R^2 results provided in Exhibit 5. Nevertheless, for stocks, even though the generated portfolios fail to surpass the historical mean benchmark, they have more excellent Sharpe ratios than the buy-and-hold strategy.

In short, Bitcoin is the sole asset whose trade sentiment metrics can be leveraged to provide exceptional economic values for real-time investors. In contrast, trading based on sentiment signals does not outperform the historical mean benchmark for stocks and gold.

CONCLUSION

This article examines the return predictability of investor sentiment across four asset classes. We contrast two types of investor sentiment: those based on trading data and those based on news and social media. For each asset, we derive 28 trade sentiment measures from its time series of prices and trading volumes and 28 text sentiment measures from news articles, headlines, social media, and news and social media combined.

Then, we employ four dimension reduction techniques, namely average variable (AV), mean combination forecast (CF), principal component analysis (PCA), and partial least squares (PLS), to facilitate a parsimonious setting in which trading and text sentiment can be easily compared. These approaches of information aggregation also improve out-of-sample return forecasts when several predictors are available.

In our empirical tests, trade sentiment performs better than text sentiment across all prediction methodologies and forecasting horizons for all four assets. We also demonstrate that all asset returns consistently underreact to trade sentiment.

As for out-of-sample tests, we find that trade sentiment produces superior return projections based on the out-of-sample R^2 measure and forecast-encompassing test for Bitcoin and government bonds. Trade and text sentiment fail to beat the historical mean benchmark in out-of-sample predictions for stocks and gold.

Finally, we investigate whether sentiment-based return projections may provide real-time mean-variance investors with economic value. According to our findings, Bitcoin is the only asset whose trade sentiment can create exceptional economic gains.

Acknowledgement: We thank seminar participants at Washington University in St. Louis and the University of Missouri-Columbia for helpful suggestions. The views expressed in this material are solely the views of the authors and are subject to change based on market and other conditions and factors.

1. See the Data section for details on the ETF list.
2. bitcoinity.org aggregates Bitcoin trading information from all major exchanges. The data can be downloaded here: https://data.bitcoinity.org/markets/price_volume/all/USD?t=lb&vu=curr.
3. See Appendix B for details on the construction of text sentiment.
4. For example, although Bitcoin has been around since January 2009, its text sentiment measures from social media and news headlines are unavailable until mid-2013. After we use the first 100 days to construct our sentiment variables, the final sample for Bitcoin starts from January 2014.
5. Because the contemporaneous return is included as an independent variable in regression (8), the contemporaneous correlations between the orthogonalized sentiment measures and asset returns are zero.
6. In unreported results, we find that using the raw sentiment measures yields robust predictability.
7. We thank Diego Garcia for making this data available.

REFERENCES

Athey, S., Parashkevov, I., Sarukkai, V., & Xia, J. (2016). Bitcoin pricing, adoption, and usage: Theory and evidence.

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

Baker, M., & Wurgler, J. (2007). Investor sentiment in the stock market. *Journal of Economic Perspectives*, *21*(2), 129–152.

Brock, W., Lakonishok, J., & LeBaron, B. (1992). Simple technical trading rules and the stochastic properties of stock returns. *Journal of finance*, *47*(5), 1731–1764.

Calomiris, C. W., & Mamaysky, H. (2019). How news and its context drive risk and returns around the world. *Journal of Financial Economics*, *133*(2), 299–336.

Campbell, J. Y., & Thompson, S. B. (2008). Predicting excess stock returns out of sample: Can anything beat the historical average? *Review of Financial Studies*, *21*(4), 1509–1531.

Chong, Y. Y., & Hendry, D. F. (1986). Econometric evaluation of linear macro-economic models. *Review of Economic Studies*, *53*(4), 671–690.

Clark, T. E., & West, K. D. (2007). Approximately normal tests for equal predictive accuracy in nested models. *Journal of Econometrics*, *138*(1), 291–311.

Cochrane, J. (2009). *Asset pricing: Revised edition*. Princeton university press.

Cochrane, J. H. (2008). The dog that did not bark: A defense of return predictability. *Review of Financial Studies*, *21*(4), 1533–1575.

DeMiguel, V., Garlappi, L., & Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Review of Financial studies*, *22*(5), 1915–1953.

Detzel, A., Liu, H., Strauss, J., Zhou, G., & Zhu, Y. (2021). Learning and predictability via technical analysis: Evidence from bitcoin and stocks with hard-to-value fundamentals. *Financial Management*, *50*(1), 107–137.

Dong, X., Li, Y., Rapach, D. E., & Zhou, G. (2022). Anomalies and the expected market return. *Journal of Finance*, *77*(1), 639–681.

Dwyer, G. P. (2015). The economics of Bitcoin and similar private digital currencies. *Journal of Financial Stability*, *17*, 81–91.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy*, *96*(2), 246–273.

Farrell, J. (1995). Talk is cheap. *American Economic Review*, *85*(2), 186–190.

Farrell, J., & Rabin, M. (1996). Cheap talk. *Journal of Economic perspectives*, *10*(3), 103–118.

Garcia, D. (2013). Sentiment during recessions. *Journal of Finance*, *68*(3), 1267–1300.

Goyal, A., & Welch, I. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies*, *21*(4), 1455–1508.

Goyal, A., Zafirov, A., & Welch, I. (2022). A comprehensive look at the empirical performance of equity premium prediction ii. *Working Paper*.

Greenwood, R., & Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies*, *27*(3), 714–746.

Harvey, D. I., Leybourne, S. J., & Newbold, P. (1998). Tests for forecast encompassing. *Journal of Business & Economic Statistics*, *16*(2), 254–259.

Huang, D., Li, J., & Wang, L. (2020). Are Disagreements Agreeable? Evidence from Information Aggregation. *Journal of Financial Economics*.

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

Huang, H., & Lee, T.-H. (2010). To combine forecasts or to combine information? *Econometric Reviews*, *29*(5-6), 534–570.

Jiang, F., Lee, J., Martin, X., & Zhou, G. (2019). Manager sentiment and stock returns. *Journal of Financial Economics*, *132*(1), 126–149.

Kelly, B., & Pruitt, S. (2013). Market expectations in the cross-section of present values. *Journal of Finance*, *68*(5), 1721–1756.

Kelly, B., & Pruitt, S. (2015). The three-pass regression filter: A new approach to forecasting using many predictors. *Journal of Econometrics*, *186*(2), 294–316.

Li, J., & Yu, J. (2012). Investor attention, psychological anchors, and stock return predictability. *Journal of Financial Economics*, *104*(2), 401–419.

Michaelides, A., Milidonis, A., & Nishiotis, G. P. (2019). Private information in currency markets. *Journal of Financial Economics*, *131*(3), 643–665.

Michaelides, A., Milidonis, A., Nishiotis, G. P., & Papakyriakou, P. (2015). The adverse effects of systematic leakage ahead of official sovereign debt rating announcements. *Journal of Financial Economics*, *116*(3), 526–547.

Neely, C. J., Rapach, D. E., Tu, J., & Zhou, G. (2014). Forecasting the equity risk premium: The role of technical indicators. *Management Science*, *60*(7), 1772–1791.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-Definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, *55*(3), 703–708.

Obaid, K., & Pukthuanthong, K. (2022). A picture is worth a thousand words: Measuring investor sentiment by combining machine learning and photos from news. *Journal of Financial Economics*.

Pagnotta, E., & Buraschi, A. (2018). An equilibrium valuation of bitcoin and decentralized network assets. *Working Paper*.

Rapach, D. E., Strauss, J. K., & Zhou, G. (2010). Out-of-sample equity premium prediction: Combination forecasts and links to the real economy. *Review of Financial Studies*, *23*(2), 821–862.

Shiller, R. J. (2005). *Irrational exuberance*. Princeton University Press, Princeton, NJ.

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

Stein, J. C. (1989). Cheap talk and the Fed: A theory of imprecise policy announcements. *American Economic Review*, 32–42.

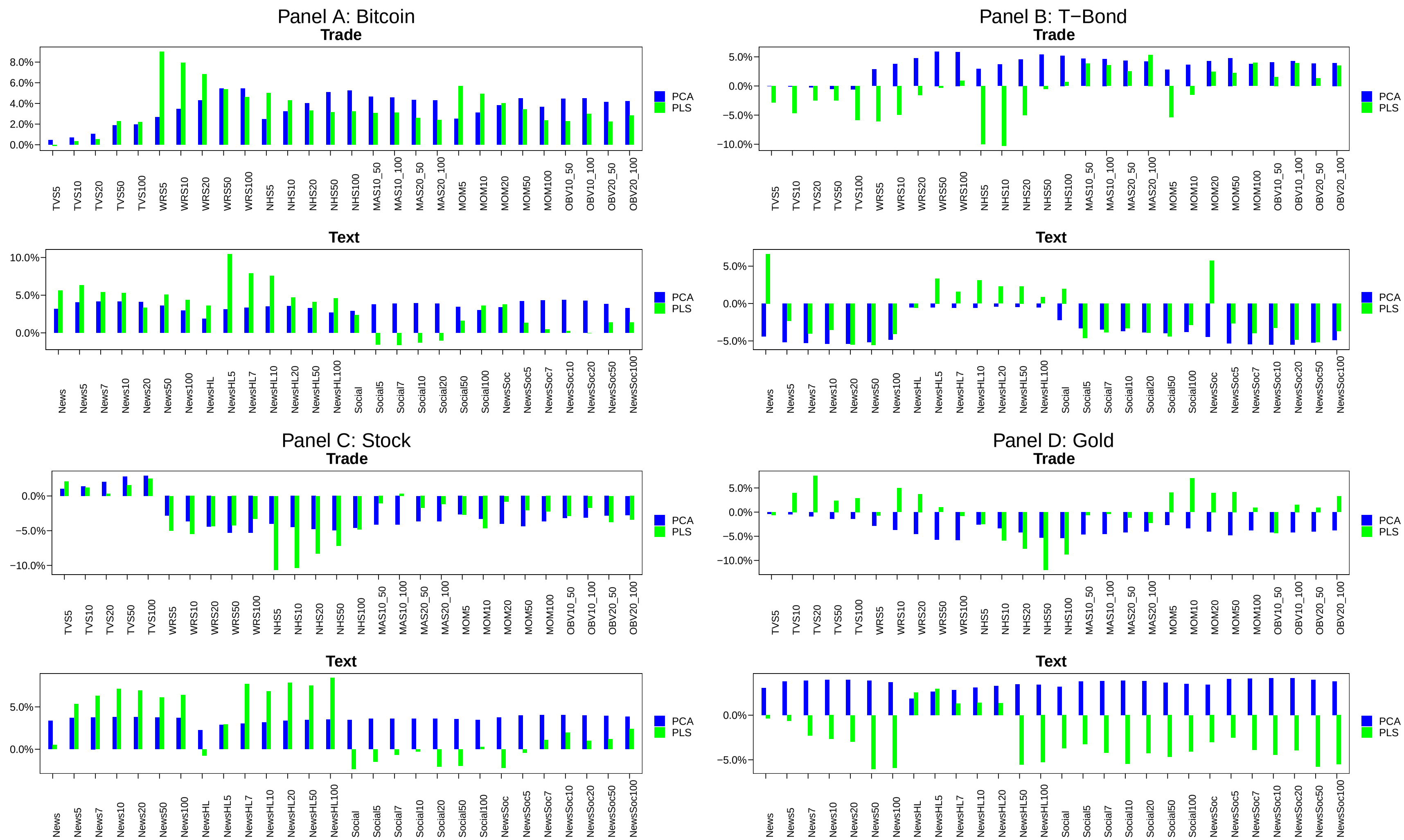
Tetlock, P. C. (2007). Giving content to investor sentiment: The role of media in the stock market. *Journal of Finance*, *62*(3), 1139–1168.

Williams, L. R. (1976). *How I made One million dollars in the commodity market last year*. Publikations-u. Handelsanst.

Zhou, G. (2018). Measuring investor sentiment. *Annual Review of Financial Economics*, *10*, 239–259.

Zhu, Y., & Zhou, G. (2009). Technical analysis: An asset allocation perspective on the use of moving averages. *Journal of Financial Economics*, *92*(3), 519–544.

Exhibit 1: PCA and PLS Weights



This figure plots the PCA (blue) and PLS (screen) weights of each orthogonalized individual sentiment measure. See section 6 for the construction of trade sentiment measures and PCA and PLS indexes. The number(s) trailing the trade sentiment variables indicate(s) the rolling window used to compute that variable. Text sentiment measures are constructed from news articles (*News*), news headlines (*NewsHL*), social media (*Social*), and both news and social media (*NewsSoc*). The number L trailing the text sentiment measures means moving average over the past L days. Returns are in percentages and independent variables are standardized to zero mean and unit standard deviation. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Exhibit 2: Contemporaneous Correlations Between Sentiment Measures and Returns

Panel A: Trade Sentiment Measures					Panel B: Text Sentiment Measures				
	Bitcoin	T-Bond	Stock	Gold		Bitcoin	T-Bond	Stock	Gold
TVS5	0.00	-0.02	-0.15 ***	-0.03 **	News	0.11 ***	0.02	0.08 ***	0.15 ***
TVS10	0.02	-0.02	-0.13 ***	-0.03 *	News5	0.05 ***	0.01	0.03 **	0.05 ***
TVS20	0.02	-0.02	-0.13 ***	-0.02	News7	0.04 **	0.00	0.03 **	0.04 **
TVS50	0.05 ***	-0.02	-0.11 ***	-0.02	News10	0.03 *	-0.01	0.02 *	0.02
TVS100	0.05 **	-0.02	-0.09 ***	-0.03 *	News20	0.02	-0.01	0.02	0.00
WRS5	0.57 ***	0.62 ***	0.56 ***	0.60 ***	News50	0.02	-0.01	0.01	-0.02
WRS10	0.46 ***	0.49 ***	0.44 ***	0.47 ***	News100	0.02	-0.01	0.01	-0.02
WRS20	0.36 ***	0.37 ***	0.34 ***	0.36 ***	NewsHL	0.11 ***	0.17 ***	0.08 ***	0.18 ***
WRS50	0.26 ***	0.26 ***	0.24 ***	0.25 ***	NewsHL5	0.06 ***	0.08 ***	0.02 *	0.08 ***
WRS100	0.20 ***	0.20 ***	0.18 ***	0.19 ***	NewsHL7	0.06 ***	0.06 ***	0.02 *	0.06 ***
NHS5	0.54 ***	0.57 ***	0.55 ***	0.55 ***	NewsHL10	0.05 ***	0.05 ***	0.02 *	0.05 ***
NHS10	0.41 ***	0.42 ***	0.39 ***	0.41 ***	NewsHL20	0.03 *	0.04 ***	0.02 *	0.03 **
NHS20	0.31 ***	0.33 ***	0.27 ***	0.31 ***	NewsHL50	0.02	0.02 *	0.02	-0.01
NHS50	0.23 ***	0.23 ***	0.18 ***	0.22 ***	NewsHL100	0.02	0.01	0.02	-0.01
NHS100	0.19 ***	0.17 ***	0.14 ***	0.18 ***	Social	0.11 ***	0.01	0.09 ***	0.13 ***
MAS10_50	0.07 ***	0.03 **	0.01	0.02	Social5	0.02	-0.00	0.02 *	0.03 **
MAS10_100	0.08 ***	0.03 **	0.01	0.01	Social7	0.02	-0.01	0.02	0.02
MAS20_50	0.06 ***	0.03 **	-0.00	0.01	Social10	0.01	-0.01	0.01	0.00
MAS20_100	0.05 ***	0.03 **	0.00	-0.00	Social20	0.00	-0.01	0.00	-0.00
MOM5	0.37 ***	0.37 ***	0.34 ***	0.35 ***	Social50	0.01	-0.01	-0.00	-0.01
MOM10	0.25 ***	0.23 ***	0.21 ***	0.24 ***	Social100	0.01	-0.01	0.00	-0.01
MOM20	0.18 ***	0.17 ***	0.14 ***	0.16 ***	NewsSoc	0.13 ***	0.02	0.09 ***	0.15 ***
MOM50	0.12 ***	0.10 ***	0.10 ***	0.09 ***	NewsSoc5	0.04 **	0.01	0.03 **	0.04 ***
MOM100	0.08 ***	0.06 ***	0.07 ***	0.07 ***	NewsSoc7	0.03	-0.00	0.02 *	0.03 **
OBV10_50	0.06 ***	0.02	-0.01	0.00	NewsSoc10	0.02	-0.00	0.02	0.01
OBV10_100	0.07 ***	0.03 **	-0.00	0.01	NewsSoc20	0.01	-0.01	0.01	0.00
OBV20_50	0.05 ***	0.02	-0.01	0.01	NewsSoc50	0.01	-0.01	0.00	-0.02
OBV20_100	0.07 ***	0.03 *	-0.02	0.01	NewsSoc100	0.01	-0.01	0.01	-0.02

This table reports contemporaneous correlations between trade (text) sentiment measures and returns in Panel A (B). See section Method for the construction of trade sentiment measures. The number(s) trailing the trade sentiment variables indicate(s) the rolling window used to compute that variable. Text sentiment measures are constructed from news articles (*News*), news headlines (*NewsHL*), social media (*Social*), and both news and social media (*NewsSoc*). The number L trailing the text sentiment measures means moving average over the past L days. ***, **, and * indicate significance at 1%, 5%, and 10% respectively. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Exhibit 3: Predicting Returns with Individual Sentiment Measures

Panel A: Trade Sentiment Measures

	Bitcoin		T-Bond		Stock		Gold	
	h=1	h=20	h=1	h=20	h=1	h=20	h=1	h=20
TVS5	-0.00	0.27	0.00	-0.08 *	0.01	-0.11 **	0.01	0.06
TVS10	-0.05	0.12	-0.01	-0.07	-0.01	-0.10 *	0.02	0.07
TVS20	-0.05	0.12	-0.00	-0.09 *	-0.01	-0.06	0.02	0.12
TVS50	0.02	0.69 **	0.00	-0.10 *	-0.01	-0.04	-0.00	0.03
TVS100	0.01	0.42	-0.01	-0.13 **	0.01	-0.02	0.01	0.02
WRS5	0.32 ***	1.25 ***	0.02	-0.00	0.04 **	0.04	0.01	0.14 *
WRS10	0.29 ***	1.14 ***	0.02	-0.02	0.06 ***	0.12 *	0.02	0.16 **
WRS20	0.18 ***	1.09 ***	0.02	0.06	0.06 **	0.13 **	0.02	0.10
WRS50	0.10	0.86 **	0.01	0.05	0.04	0.14 **	0.03 *	0.15 **
WRS100	0.04	0.68 **	0.00	0.10 **	0.03	0.15 **	0.02	0.05
NHS5	-0.04	-0.31	0.00	0.02	0.00	-0.06	-0.00	-0.03
NHS10	-0.06	-0.33	-0.01	-0.02	0.00	-0.14 **	-0.01	-0.07
NHS20	-0.04	-0.19	-0.02	0.07	0.05 *	-0.01	-0.02	-0.05
NHS50	-0.05	-0.20	-0.01	0.08	0.01	-0.01	-0.00	-0.08
NHS100	-0.12 *	-0.08	-0.02	0.08	-0.01	0.05	-0.01	-0.09
MAS10_50	0.09 *	0.74 **	0.01	0.06	0.00	0.03	0.01	-0.02
MAS10_100	0.05	0.45 *	-0.00	0.07	0.01	0.01	0.00	-0.04
MAS20_50	0.06	0.56 *	0.01	0.02	-0.01	-0.04	-0.01	-0.02
MAS20_100	0.06	0.31	0.03 **	0.15 ***	0.04 *	-0.05	0.02	-0.05
MOM5	0.19 ***	0.78 **	0.01	0.01	0.05 ***	0.08	0.02	0.15 **
MOM10	0.13 ***	0.85 ***	0.02	-0.00	0.01	0.08	0.03 *	0.05
MOM20	0.11 **	0.70 **	0.00	0.03	0.03 *	0.09	0.02	0.05
MOM50	0.03	0.39	0.00	0.08 *	0.02	0.01	0.02	0.02
MOM100	0.04	0.23	0.02 **	0.09 *	-0.03 *	-0.01	0.03	0.01
OBV10_50	0.04	0.35	0.01	0.02	0.00	-0.05	0.01	-0.08
OBV10_100	-0.02	0.64 **	0.02	0.07	0.04 *	-0.04	-0.00	-0.03
OBV20_50	0.02	0.69 **	0.03 **	0.02	-0.00	-0.14 **	-0.01	-0.02
OBV20_100	0.10 *	0.75 **	0.01	0.05	-0.00	-0.12 *	-0.00	-0.07

Panel B: Text Sentiment Measures

	Bitcoin		T-Bond		Stock		Gold	
	h=1	h=20	h=1	h=20	h=1	h=20	h=1	h=20
News	0.01	-0.11	0.03 **	0.03	-0.01	0.07	0.00	-0.02
News5	0.01	-0.16	0.02	0.01	-0.03	0.09	0.01	0.01
News7	-0.00	-0.20	0.01	0.01	-0.01	0.08	0.00	-0.02
News10	-0.02	-0.09	0.01	0.01	-0.01	0.09	0.02	-0.00
News20	0.02	-0.29	0.02	0.01	-0.03 *	0.10	0.00	0.06
News50	-0.00	-0.26	0.03 **	0.00	-0.03 *	0.05	0.00	-0.00
News100	0.02	-0.25	0.03 **	-0.03	-0.04 ***	0.11	0.02	-0.01
NewsHL	-0.01	-0.10	0.01	0.03	-0.00	0.06	0.02	0.02
NewsHL5	0.00	-0.08	0.02	0.04	-0.01	0.10	0.01	0.04
NewsHL7	0.02	-0.12	0.01	0.03	0.00	0.09	0.01	0.00
NewsHL10	-0.04	-0.04	0.02	0.03	-0.00	0.09	0.02	0.03
NewsHL20	0.04	-0.27	0.02 *	0.01	-0.01	0.07	0.01	0.09
NewsHL50	0.01	-0.16	0.01	0.05	-0.02	0.04	0.01	0.01
NewsHL100	-0.04	-0.19	0.01	0.04	-0.01	0.12 *	0.03 *	-0.01
Social	0.01	-0.40	0.01	-0.05	0.01	0.03	0.00	-0.04
Social5	0.01	-0.51	0.01	-0.03	-0.01	0.05	-0.01	-0.01
Social7	-0.06	-0.52	-0.00	-0.04	-0.01	0.05	0.01	-0.01
Social10	0.00	-0.40	0.02	-0.03	-0.01	0.05	-0.02	-0.03
Social20	0.03	-0.37	0.01	-0.02	-0.02	0.04	-0.01	-0.03
Social50	0.04	-0.32	0.01	-0.05	-0.04 **	-0.06	-0.00	-0.06
Social100	-0.01	-0.91 **	0.01	-0.07	-0.05 ***	-0.06	-0.00	-0.06
NewsSoc	0.01	-0.35	0.03 **	0.03	-0.00	0.04	-0.00	-0.02
NewsSoc5	0.01	-0.42	0.01	0.01	-0.03 **	0.05	-0.00	0.02
NewsSoc7	-0.02	-0.49	0.01	0.00	-0.02	0.05	0.00	0.00
NewsSoc10	-0.02	-0.34	0.01	0.02	-0.02	0.05	-0.01	-0.00
NewsSoc20	0.03	-0.35	0.02	0.01	-0.03 *	0.05	-0.02	0.03
NewsSoc50	0.02	-0.30	0.03 **	0.01	-0.04 ***	-0.02	-0.01	-0.03
NewsSoc100	-0.01	-0.54	0.03 **	-0.03	-0.06 ***	0.03	0.01	-0.02

This table reports β from the following regression

$$r_{t+1\to t+h} = \alpha + \beta \times x_t + \epsilon_{t+1\to t+h},$$

where $r_{t+1\to t+h}$ is the cumulative asset returns over the next h days and x_t is an orthogonalized trade (text) sentiment measure in Panel A (B). See section Method for the construction of trade sentiment measures. The number(s) trailing the trade sentiment variables indicate(s) the rolling window used to compute that variable. Text sentiment measures are constructed from news articles (*News*), news headlines (*NewsHL*), social media (*Social*), and both news and social media (*NewsSoc*). The number L trailing the text sentiment measures means moving average over the past L days. Returns are in percentages and independent variables are standardized to zero mean and unit standard deviation. ***, **, and * indicate significance at 1%, 5%, and 10% respectively based on Newey and West (1987)'s standard errors with h lags (if $h < 5$, lag is set to 5). The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Exhibit 4: Predicting Returns with Aggregate Sentiment Measures

	Panel A: Bitcoin				Panel B: T-Bond				Panel C: Stock				Panel D: Gold			
	h=1	h=5	h=10	h=20	h=1	h=5	h=10	h=20	h=1	h=5	h=10	h=20	h=1	h=5	h=10	h=20
Average Variable																
Trade AV	0.16 **	0.66 ***	1.03 ***	1.44 ***	0.02	0.09 ***	0.06	0.07	0.05 **	0.05	0.02	-0.01	0.03	0.08 *	0.08	0.06
	(2.53)	(4.04)	(3.72)	(3.37)	(0.93)	(2.92)	(1.34)	(1.03)	(2.15)	(1.19)	(0.34)	(-0.11)	(1.63)	(1.94)	(1.28)	(0.66)
IS R2 (%)	0.23	0.61	0.66	0.52	0.02	0.24	0.03	0.01	0.14	0.02	-0.01	-0.02	0.04	0.09	0.03	-0.01
Text AV	0.00	-0.10	-0.24	-0.52	0.03 **	0.01	0.00	0.00	-0.03 *	0.05	0.05	0.09	0.01	0.01	0.01	-0.00
	(0.08)	(-0.63)	(-0.81)	(-1.18)	(2.27)	(0.34)	(0.10)	(0.06)	(-1.79)	(1.43)	(1.18)	(1.15)	(0.36)	(0.33)	(0.17)	(-0.03)
IS R2 (%)	-0.03	-0.02	0.01	0.04	0.08	-0.02	-0.02	-0.02	0.04	0.02	0.01	0.02	-0.02	-0.02	-0.02	-0.02
Trade AV	0.16 **	0.67 ***	1.05 ***	1.47 ***	0.02	0.09 ***	0.06	0.07	0.05 **	0.05	0.02	-0.01	0.03	0.08 *	0.08	0.06
	(2.54)	(4.13)	(3.83)	(3.51)	(0.87)	(2.93)	(1.34)	(1.03)	(2.16)	(1.18)	(0.33)	(-0.12)	(1.62)	(1.94)	(1.28)	(0.66)
Text AV	-0.00	-0.14	-0.30	-0.61	0.03 **	0.01	0.00	0.00	-0.03 *	0.05	0.05	0.09	0.01	0.01	0.01	-0.00
	(-0.09)	(-0.88)	(-1.03)	(-1.39)	(2.21)	(0.24)	(0.06)	(0.03)	(-1.83)	(1.41)	(1.18)	(1.15)	(0.32)	(0.29)	(0.15)	(-0.04)
IS R2 (%)	0.20	0.61	0.69	0.59	0.09	0.22	0.01	-0.01	0.18	0.04	-0.01	0.00	0.02	0.07	0.01	-0.03
PCA																
Trade PCA	0.14 **	0.47 ***	0.71 ***	0.81 **	0.01	0.11 ***	0.07 *	0.09	0.05 **	0.09 **	0.07	0.13 **	0.01	0.09 **	0.07	0.05
	(2.00)	(2.77)	(2.72)	(2.07)	(0.44)	(3.08)	(1.66)	(1.49)	(2.14)	(2.22)	(1.40)	(2.09)	(0.83)	(2.40)	(1.22)	(0.77)
IS R2 (%)	0.18	0.30	0.30	0.14	-0.01	0.30	0.06	0.04	0.14	0.11	0.02	0.06	-0.01	0.12	0.01	-0.01
Text PCA	0.00	-0.10	-0.24	-0.49	0.03 **	0.01	0.00	0.01	-0.03 *	0.05	0.05	0.09	0.01	0.01	0.01	-0.00
	(0.09)	(-0.64)	(-0.81)	(-1.14)	(2.15)	(0.26)	(0.06)	(0.11)	(-1.82)	(1.35)	(1.15)	(1.14)	(0.32)	(0.33)	(0.16)	(-0.01)
IS R2 (%)	-0.03	-0.02	0.01	0.03	0.07	-0.02	-0.02	-0.02	0.04	0.02	0.01	0.02	-0.02	-0.02	-0.02	-0.02
Trade PCA	0.14 **	0.47 ***	0.72 ***	0.81 **	0.01	0.11 ***	0.07 *	0.09	0.05 **	0.09 **	0.07	0.13 **	0.01	0.09 **	0.07	0.06
	(2.00)	(2.78)	(2.75)	(2.09)	(0.38)	(3.09)	(1.67)	(1.49)	(2.15)	(2.22)	(1.40)	(2.09)	(0.81)	(2.40)	(1.23)	(0.78)
Text PCA	0.00	-0.11	-0.25	-0.50	0.03 **	0.00	-0.00	0.00	-0.03 *	0.05	0.05	0.08	0.01	0.01	0.01	-0.00
	(0.06)	(-0.68)	(-0.85)	(-1.16)	(2.09)	(0.10)	(-0.02)	(0.04)	(-1.85)	(1.34)	(1.14)	(1.14)	(0.27)	(0.20)	(0.10)	(-0.05)
IS R2 (%)	0.15	0.29	0.31	0.18	0.06	0.28	0.04	0.02	0.19	0.12	0.03	0.08	-0.03	0.10	-0.01	-0.03
PLS																
Trade PLS	0.28 ***	0.59 **	1.10 ***	2.44 ***	0.04 ***	0.08 *	0.12 *	0.03	0.05 ***	0.07	0.04	0.21 *	0.03 *	0.03	0.17	0.35 **
	(4.82)	(2.57)	(2.74)	(3.38)	(3.10)	(1.77)	(1.74)	(0.25)	(2.97)	(1.31)	(0.51)	(1.79)	(1.74)	(0.53)	(1.55)	(2.21)
IS R2 (%)	0.79	0.47	0.76	1.56	0.16	0.16	0.20	-0.01	0.17	0.07	-0.00	0.18	0.05	-0.01	0.20	0.50
Text PLS	0.08	0.19	-0.15	-0.10	0.01	-0.03	-0.02	-0.00	0.03 **	0.02	0.03	0.05	-0.01	-0.04	0.03	0.04
	(1.34)	(0.81)	(-0.38)	(-0.15)	(0.67)	(-0.85)	(-0.34)	(-0.02)	(2.09)	(0.35)	(0.30)	(0.32)	(-0.74)	(-0.59)	(0.28)	(0.23)
IS R2 (%)	0.04	0.02	-0.02	-0.03	-0.01	0.01	-0.01	-0.02	0.05	-0.01	-0.01	-0.00	-0.01	0.00	-0.02	-0.02
Trade PLS	0.28 ***	0.59 ***	1.10 ***	2.44 ***	0.04 ***	0.08 *	0.12 *	0.03	0.05 ***	0.07	0.04	0.21 *	0.03 *	0.03	0.17	0.35 **
	(4.84)	(2.58)	(2.74)	(3.38)	(3.10)	(1.77)	(1.74)	(0.25)	(2.96)	(1.31)	(0.51)	(1.79)	(1.74)	(0.53)	(1.56)	(2.21)
Text PLS	0.09	0.20	-0.14	-0.09	0.01	-0.03	-0.02	-0.00	0.03 **	0.02	0.03	0.05	-0.01	-0.04	0.03	0.04
	(1.39)	(0.83)	(-0.36)	(-0.14)	(0.67)	(-0.86)	(-0.35)	(-0.02)	(2.09)	(0.34)	(0.30)	(0.32)	(-0.73)	(-0.59)	(0.29)	(0.24)
IS R2 (%)	0.83	0.50	0.75	1.53	0.15	0.17	0.19	-0.03	0.22	0.06	-0.01	0.18	0.04	-0.00	0.19	0.48

This table reports the following regression

$$r_{t+1\to t+h} = \alpha + \beta \times x_t + \epsilon_{t+1\to t+h},$$

where $r_{t+1\to t+h}$ is the cumulative asset returns over the next h days and x_t is an orthogonalized aggregate sentiment measure. Average Variable (AV) means average across all trade sentiment variables (Trade AV) and separately average across all text sentiment variables (Text AV). Principal Component Analysis (PCA) uses the PCA index constructed with all trade sentiment variables (Trade PCA) and separately with all text sentiment variables (Text PCA). Partial Least Squares (PLS) is constructed using the next one day returns to supervise the constructions. PCA and PLS indexes for trade (text) sentiment measures are adjusted to have the same sign as Trade (Text) AV to remain the sentiment interpretation. Returns are in percentages and independent variables are standardized to zero mean and unit standard deviation. Reported in parenthesis are t-statistics corrected for Newey and West (1987)'s standard errors with h lags (if $h < 5$, lag is set to 5). ***, **, and * indicate significance at 1%, 5%, and 10% respectively. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Exhibit 5: Out-Of-Sample R-Squared

Panel A: Bitcoin					**Panel B: T-Bond**				
	h=1	h=5	h=10	h=20		h=1	h=5	h=10	h=20
Trade AV	1.57 ***	2.09 ***	2.71 ***	3.65 ***	Trade AV	-0.05	-0.19	-0.11	-0.09
Trade CF	0.22 ***	0.19 ***	0.15 **	0.13 ***	Trade CF	-0.02	0.06 **	0.03	0.04 *
Trade PCA	0.18 **	0.09 **	-0.11 *	-0.25	Trade PCA	-0.13	0.16 **	-0.03	-0.12
Trade PLS	0.48 ***	-0.35 ***	-0.54 *	-0.53 **	Trade PLS	-0.73	-0.06	-0.12	-0.08 **
Text AV	-0.03	-0.21	-0.33	-0.74	Text AV	-0.06	-0.16	-0.39	-0.69
Text CF	-0.03	-0.04	-0.02	-0.00	Text CF	0.02	-0.03	-0.03	-0.03
Text PCA	-0.04	-0.16	-0.08	-0.02	Text PCA	0.06 **	-0.06	-0.06	-0.05
Text PLS	-0.58	-0.08	0.04	-0.04	Text PLS	-0.27	-0.23	-0.15	-0.29
Panel C: Stock					**Panel D: Gold**				
	h=1	h=5	h=10	h=20		h=1	h=5	h=10	h=20
Trade AV	-0.01 *	-0.28	-0.68	-1.67	Trade AV	-0.11	-0.58	-0.95	-1.27
Trade CF	0.05	0.01	-0.02	0.00	Trade CF	-0.03	0.03	0.05 *	0.03
Trade PCA	0.12 *	-0.08	-0.09	-0.14	Trade PCA	-0.06	0.07 *	-0.04	-0.06
Trade PLS	-0.55	-0.61	-0.80	-0.57 *	Trade PLS	-1.07	-0.79	-0.66 *	-0.53 *
Text AV	-0.09	-0.38	-0.85	-1.59	Text AV	-0.17	-0.77	-1.17	-0.77
Text CF	0.00	0.01	0.01	0.04 *	Text CF	-0.04	-0.07	-0.07	-0.13
Text PCA	-0.04	-0.01	-0.02	-0.07	Text PCA	-0.09	-0.15	-0.23	-0.42
Text PLS	-0.03 **	-0.28	-0.31	-0.24	Text PLS	-0.27	-0.38	-0.29	-0.35

This table reports the out-of-sample R^2_{OS} in percentages from recursively predicting the out-of-sample returns over the next h days using the following methods: average of forecasting variables (AV), average of individual forecasts (CF), principal component analysis (PCA), and partial least squares (PLS). The estimation employs an expanding window with an initial training window of 750 days. ***, **, and * indicate significance at 1%, 5%, and 10% respectively, based on the Clark and West (2007)'s statistic. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Exhibit 6: Forecast Encompassing Tests

Panel A: Bitcoin

	Trade AV	Trade CF	Trade PCA	Trade PLS	Text AV	Text CF	Text PCA	Text PLS
h=1								
Trade AV		0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trade CF	**0.34**		**0.14**	0.00	0.00	0.00	0.00	0.00
Trade PCA	**0.28**	**0.23**		0.00	0.03	0.03	0.03	0.00
Trade PLS	0.00	0.00	0.00		0.00	0.00	0.00	0.00
Text AV	**0.53**	**1.00**	**0.38**	0.00		**0.46**	**0.31**	0.00
Text CF	**0.46**	**1.00**	**0.37**	0.00	**0.33**		**0.22**	0.00
Text PCA	**0.46**	**1.00**	**0.37**	0.00	**0.36**	**0.60**		0.00
Text PLS	**0.64**	**1.00**	**0.80**	0.01	**0.95**	**0.98**	**0.96**	
h=20								
Trade AV		0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trade CF	**0.45**		0.01	0.00	0.00	0.01	0.02	0.02
Trade PCA	**0.45**	**0.56**		0.00	0.00	**0.24**	**0.22**	**0.17**
Trade PLS	**0.20**	0.10	0.03		0.00	0.03	0.02	0.02
Text AV	**0.56**	**1.00**	**0.41**	0.00		**1.00**	**1.00**	**1.00**
Text CF	**0.47**	**0.96**	0.02	0.00	0.00		**0.30**	**0.18**
Text PCA	**0.43**	**0.88**	0.02	0.00	0.00	**0.51**		**0.18**
Text PLS	**0.53**	**0.82**	0.02	0.00	0.00	**0.46**	**0.32**	

Panel B: T-Bond

	Trade AV	Trade CF	Trade PCA	Trade PLS	Text AV	Text CF	Text PCA	Text PLS
h=1								
Trade AV		**0.61**	0.07	0.00	**0.21**	**0.90**	**0.89**	0.00
Trade CF	**0.14**		0.03	0.00	**0.14**	**0.89**	**0.84**	0.00
Trade PCA	**0.62**	**0.88**		0.00	**0.45**	**0.88**	**0.89**	0.01
Trade PLS	**0.82**	**0.91**	**0.87**		**0.76**	**0.88**	**0.89**	**0.30**
Text AV	**0.33**	**0.50**	**0.10**	0.00		**0.68**	**0.71**	0.00
Text CF	0.02	0.06	0.02	0.00	0.05		**0.82**	0.00
Text PCA	0.01	0.03	0.01	0.00	0.02	0.06		0.00
Text PLS	**0.53**	**0.65**	**0.28**	0.00	**0.45**	**0.82**	**0.86**	
h=20								
Trade AV		**0.60**	0.10	0.01	0.00	**0.39**	**0.29**	0.00
Trade CF	0.03		0.06	0.01	0.00	0.02	0.02	0.00
Trade PCA	**0.14**	**0.52**		0.05	0.00	**0.30**	**0.24**	0.02
Trade PLS	0.01	0.08	0.03		0.00	0.03	0.02	0.00
Text AV	**0.36**	**0.67**	**0.23**	0.06		**0.59**	**0.58**	**0.11**
Text CF	0.09	**0.93**	0.10	0.01	0.00		**0.12**	0.00
Text PCA	**0.12**	**0.91**	**0.11**	0.01	0.00	**0.75**		0.00
Text PLS	**0.41**	**0.94**	**0.25**	0.04	0.00	**0.89**	**0.79**	

Panel C: Stock

	Trade AV	Trade CF	Trade PCA	Trade PLS	Text AV	Text CF	Text PCA	Text PLS
h=1								
Trade AV		**0.13**	0.06	0.00	0.04	**0.12**	0.08	0.01
Trade CF	0.04		**0.33**	0.00	0.01	**0.11**	0.05	0.01
Trade PCA	0.01	**0.11**		0.00	0.03	0.06	0.04	0.01
Trade PLS	0.09	**0.44**	**0.90**		**0.18**	**0.26**	**0.20**	0.08
Text AV	**0.15**	**0.94**	**0.45**	0.00		**0.88**	**0.42**	0.07
Text CF	0.10	**0.72**	**0.31**	0.00	0.01		**0.11**	0.03
Text PCA	**0.13**	**0.66**	**0.33**	0.00	0.06	**0.65**		0.04
Text PLS	0.01	0.08	0.08	0.00	0.01	0.07	0.03	
h=20								
Trade AV		**0.47**	**0.23**	0.00	0.00	**0.45**	**0.20**	**0.13**
Trade CF	0.00		0.00	0.00	0.00	**0.47**	0.01	0.00
Trade PCA	0.00	**0.88**		0.00	0.00	**0.73**	0.06	0.01
Trade PLS	0.00	**0.14**	**0.15**		0.00	0.09	0.02	0.01
Text AV	0.00	**1.00**	**0.98**	0.02		**1.00**	**0.98**	**0.88**
Text CF	0.00	0.08	0.00	0.00	0.00		0.01	0.00
Text PCA	0.00	0.10	0.01	0.00	0.00	**0.60**		0.00
Text PLS	0.00	**0.51**	**0.12**	0.00	0.00	**0.54**	0.04	

Panel D: Gold

	Trade AV	Trade CF	Trade PCA	Trade PLS	Text AV	Text CF	Text PCA	Text PLS
h=1								
Trade AV		**0.83**	**0.66**	0.00	**0.20**	**0.78**	**0.42**	0.02
Trade CF	0.04		**0.11**	0.00	**0.12**	**0.35**	**0.12**	0.01
Trade PCA	**0.13**	**0.80**		0.00	**0.16**	**0.68**	**0.25**	0.01
Trade PLS	**0.70**	**0.86**	**0.79**		**0.57**	**0.78**	**0.72**	**0.32**
Text AV	**0.39**	**0.61**	**0.54**	0.00		**0.60**	**0.45**	0.07
Text CF	0.06	**0.50**	**0.18**	0.00	**0.12**		0.07	0.01
Text PCA	**0.28**	**0.75**	**0.57**	0.00	**0.17**	**0.87**		0.02
Text PLS	**0.25**	**0.43**	**0.36**	0.00	**0.21**	**0.53**	**0.34**	
h=20								
Trade AV		**0.84**	**0.76**	0.01	0.00	**0.68**	**0.32**	**0.30**
Trade CF	0.00		0.01	0.00	0.00	0.02	0.00	0.00
Trade PCA	0.00	**0.98**		0.00	0.00	**0.11**	0.01	0.01
Trade PLS	0.00	0.08	0.03		0.00	0.02	0.00	0.00
Text AV	0.00	0.03	0.02	0.00		0.01	0.00	0.00
Text CF	0.00	**0.93**	**0.71**	0.00	0.00		0.01	0.02
Text PCA	0.00	**0.90**	**0.82**	0.00	0.00	**0.94**		**0.18**
Text PLS	0.00	**0.71**	**0.57**	0.00	0.00	**0.57**	0.07	

This table reports the p-value of testing the null hypothesis that the out-of-sample forecasts over the next h-day returns made by the models in the columns encompass the forecasts made by the models in the rows. p-value greater than 10% are in bold, i.e. those entries indicates the failure to reject the null that the forecast made by the model in the column encompasses the forecast made by the model in the row. For trade and text sentiment measures separately, the four prediction models include average of forecasting variables (AV), average of individual forecasts (CF), principal component analysis (PCA), and partial least squares (PLS). Grey areas indicate where the trade and text sentiment measures are tested against each other. The estimation employs an expanding window with an initial training window of 750 days. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Exhibit 7: Asset Allocation

Panel A: Bitcoin					Panel B: T-Bond				
	$\gamma = 3$		$\gamma = 5$			$\gamma = 3$		$\gamma = 5$	
	UG	SR	UG	SR		UG	SR	UG	SR
Trade AV	61.72 ***	2.00 ***	48.42 ***	2.08 ***	Trade AV	-0.37	0.22	-0.39	0.22
Trade CF	16.17 ***	0.92 ***	13.33 ***	0.90 ***	Trade CF	-0.21	0.23	-0.02	0.25
Trade PCA	35.76 ***	1.49 ***	33.36 ***	1.72 ***	Trade PCA	-2.20	0.10	-2.44	0.05
Trade PLS	50.29 ***	1.73 ***	37.06 ***	1.77 ***	Trade PLS	-0.40	0.21	-0.67	0.24
Text AV	2.53	0.50	2.27 **	0.41 *	Text AV	-1.20	0.19	-1.23	0.20
Text CF	-1.49	0.38	-0.87	0.29	Text CF	0.03	0.25	0.11	0.26
Text PCA	1.06	0.47	2.01	0.44 *	Text PCA	1.22	0.32	1.16	0.35
Text PLS	-11.60	0.10	-9.05	-0.00	Text PLS	-1.07	0.16	-1.52	0.15
Buy-Hold		0.76		0.76	Buy-Hold		0.32		0.32

Panel C: Stock					Panel D: Gold				
	$\gamma = 3$		$\gamma = 5$			$\gamma = 3$		$\gamma = 5$	
	UG	SR	UG	SR		UG	SR	UG	SR
Trade AV	-0.04	0.49	0.21	0.52	Trade AV	-0.77	0.26	-1.23	0.25
Trade CF	-0.37	0.49	-0.89	0.40	Trade CF	-0.40	0.27	-0.49	0.28
Trade PCA	-1.66	0.38	-2.21	0.29	Trade PCA	-0.44	0.27	-0.42	0.28
Trade PLS	-1.49	0.39	-1.37	0.40	Trade PLS	-2.15	0.15	-3.22	0.12
Text AV	-1.23	0.41	-0.78	0.43	Text AV	-1.95	0.13	-0.46	0.19
Text CF	-0.82	0.44	-0.95	0.39	Text CF	-0.93	0.24	-0.37	0.28
Text PCA	-1.59	0.38	-1.70	0.32	Text PCA	-1.85	0.19	-1.11	0.23
Text PLS	-0.88	0.43	-0.35	0.48	Text PLS	-0.17	0.25	-1.30	0.21
Buy-Hold		0.36		0.36	Buy-Hold		0.45		0.45

This table reports utility gains for a mean-variance investor who allocates between a risky asset and a risk-free asset using the one-day return forecast by a predictive model against the forecast using the historical mean benchmark. **UG** means annualized utility gains in percentages while **SR** means annualized Sharpe ratio of the resulting portfolio. ***, **, and * indicate significance at 1%, 5%, and 10% respectively, based on the tests in DeMiguel, Garlappi, and Uppal (2009). Stars under the UG columns test the significance of the utility gains and stars under the SR columns test whether the Sharpe ratio of the portfolio using the predictive model is different from that using the historical mean return. The estimation employs an expanding window with an initial training window of 750 days. The sample period for stocks, T-Bond, gold, and Bitcoin is daily from May 1998, January 2003, May 2005, and January 2014, respectively to September 2022.

Appendix
Investor Sentiment and Asset Returns: Actions Speak Louder than Words

A Replication Of Garcia (2013)

Exhibit 8: Predicting Returns with News Sentiment from Garcia (2013)

	β_1	β_2	β_3	β_4	β_5	R^2	Obs
			Panel A: 1905-2005				
POS	0.04 ***	0.00	-0.01	-0.01 *	-0.00	1.83	27427
	(5.43)	(0.55)	(-0.90)	(-1.85)	(-0.54)		
NEG	-0.04 ***	0.00	0.01	0.01	0.01	1.82	27427
	(-5.24)	(0.36)	(0.70)	(1.29)	(1.56)		
SENT	0.05 ***	-0.00	-0.01	-0.02 *	-0.01	1.91	27427
	(6.53)	(-0.12)	(-0.98)	(-1.89)	(-1.51)		
			Panel B: 1984-1999				
POS	0.02	0.01	-0.00	0.00	-0.00	4.41	4044
	(1.52)	(0.45)	(-0.21)	(0.08)	(-0.22)		
NEG	-0.05 ***	-0.03 **	-0.01	0.02	0.01	4.72	4044
	(-3.04)	(-1.99)	(-0.39)	(1.51)	(0.48)		
SENT	0.05 ***	0.03 *	0.00	-0.02	-0.01	4.69	4044
	(3.03)	(1.78)	(0.22)	(-1.08)	(-0.57)		
			Panel C: 2000-2005				
POS	-0.03	-0.02	-0.00	-0.03	0.07 *	0.38	1501
	(-0.83)	(-0.61)	(-0.12)	(-0.81)	(1.89)		
NEG	0.01	0.03	0.04	-0.06 **	0.01	0.46	1501
	(0.31)	(0.94)	(1.44)	(-2.21)	(0.41)		
SENT	-0.02	-0.03	-0.04	0.04	0.02	0.32	1501
	(-0.54)	(-1.00)	(-1.21)	(1.38)	(0.62)		

This table reports the results of the following predictive regression:

$$r_t = \alpha + \beta' L_{1-5}(X_t) + \gamma' Z_t + \epsilon_t,$$

where r_t is the return of the DJIA index on the day t, $L_{1-5}(X_t)$ is five lags of one of news sentiment variables—positive (POS), negative (NEG), or POS-NEG (SENT)—from Garcia (2013), and Z_t is a set of control variables including five lags of daily returns, five lags of squared daily returns, and weekday indicators. To conserve space, only β's are reported. Returns and in-sample R^2's are in percentage points. Reported in parenthesis are t-statistics corrected for Newey and West (1987) standard errors with five lags. ***, **, and * indicate significance at 1%, 5%, and 10% respectively. Panel A reports the sample from 1905 to 2005 as in Garcia (2013), Panel B reports the sample from 1984 to 1999 as in Tetlock (2007), and Panel C reports the sample from 2000 to 2005.

B Refinitiv MarketPsych Indices (RMI) *Text Sentiment* Construction

Text sentiment is a news content index gathered by Refinitiv MarketPsych Indices (RMI). It is based on word recognition techniques designed to extract relevant economics, finance, business, and psychology data from financial news, social media, conference call transcripts, and executive interviews. Since 1998, statistics have been accessible for 45 currencies, over 12,000 firms from over 75 countries, and index-level data for the 15 largest markets. Also accessible are data for 187 nations and regions. Only English-language text is used for analysis.

B.1 Data Analysis

Most previous research uses a lexical analysis technique in which an article's text is compared to a context-relevant vocabulary that has been pre-specified, and their frequencies are then tallied. This method's major shortcoming is focusing solely on one component of *Text sentiment* (positive versus negative) while ignoring other partially linked variables. Also, certain open-source *Text sentiment* dictionaries may misclassify topics linked to business and money. RMI word recognition technique uses extensive customization and curation of lexicons. It aims to identify and score hundreds of *Text sentiment* dimensions using different grammatical frameworks to various text sources based on their unique characteristics and takes sentence and article structure into account. The convergence of the text processing techniques described below, when applied to text, yields approximately 400 psychological variables (*PsychVars*)), each of which has the potential to be applied to a different entity. It should be noted that the method seeks to find word interrelationships and calculates the score based on these instead of employing frequency analysis of individual words. Following is a discussion of the word recognition and content quantification techniques used in data generation.

B.2 Source Type Differences

The fact that RMI data are taken from traditional news and social media sources creates a problematic environment for word recognition techniques. Due to the differences in communication patterns between social and news media, each source's data analysis is subsequently tailored.

In contrast to mainstream media, which employ standard terminology and acceptable tone in their material, social media are rife with sarcasm, irony, colloquial meanings, incomplete thoughts, incorrect punctuation, misspellings, non-standard syntax, case insensitivity, and harsh language. The expressions of emotions differ between the news and social media. In social media, authors frequently express themselves through the use of emoticons (e.g., ":)") and acronyms (e.g., "LOL").

Moreover, unlike journalists of traditional news media, who are educated to provide many perspectives on the underlying issue and undergo an editorial process, authors in social media tend to convey their thoughts and emotional states more directly.

The duty of journalists in the news is to describe the emotional states of individuals being reported on. Consequently, social media content is often less inclusive of opposing perspectives and more emotionally expressive from the first-person perspective than news information.

Due to these variances, text analytical models are utilized to calibrate text analytics by source type, and separate models are employed for news, social media forums, tweets, SEC filings, and earnings conference call transcripts.

B.3 Entity Identification

A list including more than 60,000 entity names and their aliases, such as language-specific writing forms of locations and businesses, is used to identify entities.

RMI utilizes anti-correlate filters to eliminate irrelevant elements. For instance, gold and silver are typically referenced every two years at the Olympic Games, although they have no relation to gold and silver commodities. Similarly, allusions to the South Korean Won may allude to the successes of South Korean athletes or the country's currency. Additionally, the entities must have accurate co-references. For instance, "breakfast oats" is not a valid reference to the commodity Oats unless it is accompanied by crucial identification correlates such as "prices" or "futures."

B.4 Timing

RMI has various expectations-related variables. The text-analytics program is tuned to identify verb tenses in each phrase and recognize whether references are future-oriented and related to anticipations. For instance, "Optimism" distinguishes between references to future-oriented positive and negative statements, but "Uncertainty" excludes references to historical uncertainties from its analysis.

B.5 Modifiers and Negations

By changing the impact of a phrase or sentence, modifiers alter its meaning or tone. For instance, words or phrases that raise the importance of an adjective, such as "big" (e.g. "great loss"), are multiplicative on the weighting of the modified word, but minimizers such as "a little" and "a handful" multiply the meaning score by 0.5. Modifiers may enhance (maximizers) or decrease (minimizers) the score of a critical term, influencing the scoring of textual meaning. In the case of frequent interpretations based on fundamental word relationships, such as "new," a multiplier is employed to reduce the significance of the meaning. In the Innovation index, the multiplier for the word "new" is 0.1 when used with this meaning. The data also considers the distinctions between news headlines and article bodies. In the data creation, headlines have a multiplicative weight of 3, subtitles a weight of 2, and bodies a weight of 1. In addition to maximization and minimization, the data construction considers four negations. For instance, "I'm not worried about the earnings release" implies that the author is not fearful; hence, the extracted fear score is nullified (-1).

B.6 Source Texts

The RMI measures are derived from two categories of sources, news, and social media, and the data consists of three feeds: a news feed, a social media feed, and a combined news and social media feed. RMI analyses about two million articles every day, and the data are updated on a minute-by-minute basis. Each minute number represents the average of the previous 24 hours' observations regarding the target entity.

RMI uses data from the largest and most influential news organizations, global online news coverage, and a wide variety of trustworthy social media sites as its sources. The RMI news indexes are derived from content provided by Thomson Reuters News Feed Direct and two Thomson Reuters news archives: a Reuters-only archive from 1998 to 2002 and one containing Reuters and chosen third-party wires from 2003 onwards. In 2005, Moreover Technologies' aggregate news feed was added to the data, increasing the number of sources by 50,000 online news sites. In addition, the data includes hundreds of financial news websites and finance-specific sources that professional investors extensively read. Since 1998, MarketPsych has downloaded social media content from public social media sites. In 2009, RMI added data from Moreover Technologies' aggregate social media feed, collected from more than four million social media sites. v

B.7 Quantifying the Articles

Each RMI variable constitutes a combination of minor components called *PsychVars*. To form an RMI variable, the absolute values of *PsychVars* are first determined using 24 hours of observations. These absolute values are then summed for all constituents to get the Buzz variable. More specifically, Buzz is calculated as follows:

$$Buzz_t(a) = \sum_{c \in C(a), p \in P(s)} |PsychVar_{p,t}(c)|,$$

where t denotes time, a is the entity, $C(a)$ is the set of all constituents of a (for example, for indices, $C(a)$ consists of all individual assets comprising the index and for respective companies $C(a) = a$) and $P(s)$ is the set of all *PsychVars* relevant to a particular RMI variable s.

The value of an RMI variable s of an entity an at time t is calculated as follows:

$$RMI_{s,t}(a) = \frac{\sum_{c \in C(a), p \in P(s)} (I_t(s,p) * PsychVar_{p,t}(c))}{Buzz_t(a)}$$

where $I_t(s,p)$ defines whether a *PsychVar* $p \in P(s)$ is additive or subtractive to an RMI variable s at time t:

$$I_t(s,p) = \begin{cases} +1, \text{if additive at time t} \\ -0, \text{ if subtractive at time t} \end{cases}$$

A single *PsychVar* can contribute to multiple RMI variables. For example, EarningsUp_f *PsychVar* (see below for an example) is a constituent of all, EarningsForecast, Text Sentiment, Optimism, and FundamentalStrength RMI variables.

B.8 Sentence-level Example

The following shows an example of the RMI data-generating process using the abovementioned principles. Consider the following sentence: "Analysts expect Mattel to report much higher earnings next quarter." The language analyzer performs the following sequence:

[1] Associates ticker symbol MAT with entity reference "Mattel."

[2] Identifies "earnings" as an Earnings word in the lexicon.

[3] Identifies "expect" as a future-oriented word and assigns future tense to the phrase.

[4] Identifies "higher" as an Up-Word (positive reference).

[5] Multiplies "higher" by 2 due to presence of the modifier word "much."

[6] Associates "higher" (Up-Word) with "earnings" (Earnings) due to proximity.

The analysis algorithm will report the following:

Date	Time	Ticker	PsychVar	Score
20110804	15:00.123	MAT	EarningsUp_f	2

In the example above, 2 is the raw score produced for EarningsUp_f.