The Relationship Between Stock Returns and Investor Sentiment: Evidence from Social Media

December 8, 2014

Zachary McGurk
College of Business and Economics, West Virginia University
Morgantown, WV 26506, United States
zcmcgurk@mix.wvu.edu

Adam Nowak
College of Business and Economics, West Virginia University
Morgantown, WV 26506, United States
adam.d.nowak@gmail.com

Abstract

The recent behavioral finance literature has found investor sentiment having some predictive ability in equity returns. This differs from the standard finance theory provides no role for investor sentiment. We examine the relationship between investor sentiment and stock returns by employing textual analysis on social media posts. Overall we find that our investor sentiment measure has a positive and significant effect on stock returns. These finds are consistent across a number of different models and specifications, thus finding further evidence against the standard finance theory.

JEL-Classification: G12, G13, G14
Keywords: Investor Sentiment, Stock Returns, Social Media, Predictive Regression
1 Introduction

The recent finance literature has tried to find the relationship between investor sentiment and stock returns. This literature has been limited to using macro estimates of sentiment using consumer confidence [5] or data mining methods like principal component analysis [1, 2]. Our paper utilizes textual analysis from Twitter to estimate a new daily measure of investor sentiment and then test its relationship with stock returns. Twitter is unique in that it is first activity used by over 500 million users in June 2012 ¹, and second it is used by individuals to expressed opinions and thoughts on number of different subject matters, including equity prices. Other measures of investor sentiment, like the AAII Investor Sentiment Survey, are limited in that it only asks individual investors weekly if they are bullish, bearish, and neutral on the stock market for the next six months ². Due to these measures of investor sentiment being survey in nature, investors are only allowed to respond to pre-written questions. Our measure is estimated on a daily interval and can show the sentiment for individual stocks. Also because this not a survey, individuals can tweet can contain any text as long as it is 140 characters and under.

Twitter itself has been shown to have an effect on specific stock returns. On January 10, 2011, the Business Insider reported that Hip Hop artist, 50-Cent(Curtis Jackson), tweeted "HNHI is the stock symbol for TVG there launching 15 different products. they are no joke get in now." The article went on to state "In the three months to the end of September the company was operating at a loss with cash of just $198,000 and a deficit of $3.3m. Then, on November 23, it said it would offer 180m shares to the public at a price of just 17 cents... trading under the stock name HNHI, were worth just 4 cents each. Spurred by the tweet, the stock took off. It hit nearly 50 cents on Monday, before closing at 39 cents."³

Overall we find a positive impact of our investor sentiment estimate and stock returns. With an increase in investor sentiment leading to an increase in daily stock returns.

The rest of this paper is structured as follows. First a discussion on the relevant literature. Next, a section detailing our estimation of investor sentiment. After that a description of the data and our empirical methodology.

¹This is reported by Semiocast
²This survey is conducted by the American Association of Individual Investors http://www.aaii.com/sentimentsurvey
Followed by explanation of the results and finally a conclusion.

2 Literature and Theory

Previous works ([1] [5]) detail classical finance theory having no role for investor sentiment determining stock returns.

3 Methodology

3.1 Sentiment

Each tweet contains 140 characters or less. Each tweet can be indexed by date, $t$, and company, $j$. On each date, there are tweets that discuss company. These tweets can be indexed by

$$i = 1, 2, \ldots N_{jt}$$

We therefore use the convention $x_{ijt}$ to refer to the $j$th tweet on date $t$ discussing company $i$. In order to indicate that an individual is discussing a company, he or she will append a “$” before the ticker symbol. For instance, when discussing Haliburton (Ticker symbol: HAL) an individual will use “$HAL” in the tweet. For example:

$$\text{Bought 100 shares of $HAL on a dip!!!}$$

In addition to indicating what company is being discussed, each $x_{ijt}$ contains words that describe sentiment regarding company $j$. That is, we can describe $x_{ijt}$ as a set of $W_{ijt}$ words. Dropping the subscript in $W_{ijt}$ we have:

$$x_{ijt} = \{w_{ijt}^1, w_{ijt}^2, \ldots, w_{ijt}^W\}$$

For instance, the tweet in indicates that the individual writing the tweet is taking a long position in Haliburton. This long position is indicated by the fact that “bought” appears in the tweet. We assume that the sentiment of each tweet can be represented as the words that appear in $x_{ijt}$.

In order to calculate sentiment we use a list of positive and negative words. The Positive Word List (PWL) is a list of words or phrases that signify a positive outlook for the stock. This can include: bought, bullish, breakout, etc. In addition, we also use a Negative Word List (NWL); words in NWL include: sold, bearish, dismal, etc. We write PWL and NWL as:
\[
\text{PWL}_i = \{\text{PWL}_1, \text{PWL}_2, \ldots\} = \{\text{bought, bullish, breakout, } \ldots\} \tag{2a}
\]
\[
\text{NWL}_i = \{\text{NWL}_1, \text{NWL}_2, \ldots\} = \{\text{old, bearish, dismal, } \ldots\} \tag{2b}
\]
\[
\text{s}_{ijt}^c = \#(x_{ijt} \cap \text{PWL}) - \#(x_{ijt} \cap \text{NWL}) \tag{3}
\]

We define sentiment as the difference in the number of words from the PWL and NWL. This is written as:

Here, \(\#(X)\) is the number of words in the set \(X\). Equation (3) is a base equation which simply counts the difference in the number of positive and negative words lists. If there are more positive words than negative words in \(x_{ijt}\), then \(s_{ijt}^c > 0\) and vice-versa.

An alternative scheme weights the words in the word list. Define the variable, \(\text{PWL}_{ijt}^k\), as an indicator variable for the \(k^{th}\) word of \(\text{PWL}\) and likewise for \(\text{NWL}_{ijt}^k\).

\[
\text{PWL}_{ijt}^k = \begin{cases} 
0 & : \text{PWL}_k \notin x_{ijt} \\
1 & : \text{PWL}_k \in x_{ijt}
\end{cases} \tag{4a}
\]
\[
\text{NWL}_{ijt}^k = \begin{cases} 
0 & : \text{NWL}_k \notin x_{ijt} \\
1 & : \text{NWL}_k \in x_{ijt}
\end{cases} \tag{4b}
\]

Using the indicator variables, an alternative measure for weighting the sentiment is found by:

\[
\text{s}_{ijt}^w = \sum_k \phi^k \text{NWL}_{ijt}^k + \sum_k \theta^k \text{PWL}_{ijt}^k \tag{5}
\]

### 3.2 Determining if the Tweet is Relevant

We can assign relevance weight to each tweet. Define the Relevant Word List (RWL) as the list of words that most frequently appear in tweets discussing stocks. This might look like:

\[
\text{RWL} = \{\text{stock, shares, debt, sales} \ldots\}
\]

We will need relevance in order to weight tweets based on whether or not we believe them to be talking about stocks. This is important due to the way in which we captured the data. First, we collect a random set of 2000 tweets, in \(TS\). We then manually look through \(TS\) and code each \(x_{ijt}\) in \(S\) as \(Y_{ijt} = 1\) if the tweet clearly discusses stocks and 0 otherwise. From this we obtain a list of most commonly used words and weight each by its occurrence in the sample. Next we then take the top 200 words, provided in Table 1, and then weight each tweet by relevance. Finally we drop the bottom 10% and thus only keep relevant tweets.
4 Data and Descriptive Statistics

5 Empirical Approach

We are interested in whether or not sentiment is correlated with current or future stock returns. For security $j$ at time $t$ we estimate:

$$ R_{jt} = \gamma_0 + \gamma_1 R_{mt} + \gamma_2 s_{ijt} + u_{ijt} \quad (6) $$

$$ R_{jt} = \gamma_0 + \gamma_1 R_{mt} + \gamma_2 s_{ijt} + \gamma_3 MC_{ijt} + u_{ijt} \quad (7) $$

$$ R_{jt} = \gamma_0 + \gamma_1 R_{mt} + \gamma_2 s_{ijt} + \gamma_3 MC_{ijt} + \gamma_4 F_{mt} + u_{ijt} \quad (8) $$

6 Results

6.1 Results for Full Sample OLS

Table 4 detail the results from $EQ : 6 - 8$ using a standard OLS estimation method. Overall we find that the mean of estimated twitter investor sentiment has a positive and significant effect on stock returns with an increase in the average of one positive word increasing stock returns by 200 basis points. This is interpreted as the more positive investor sentiment is the larger the return. When including Market Capitalization, and the Frama-French factors, as discussed in $EQ : 8 - 9$, still produces a similar estimated coefficient and does not decrease statistical significance. This consistent with the literature with investor sentiment having a positive effect on stock returns ( 1).

6.2 Results for Panel

Table 5 details the results from $EQ : 6 - 9$ using fixed effects panel regression. Again this result provide similar estimation results as the full sample OLS estimates. Again with an increase in the average of one positive word increasing stock returns by 200 basis points. These results provide robustness check in our investor sentiment measure.

6.3 Results for Individual Stocks

6.4 Out of Sample Tests

To further test the validity of our investor sentiment measure we include standard nested out of sample tests. Clark and McCracken (2001, 2005) [3,4]
and McCracken(2007) provide good basis for studying nested out of sample tests. In these papers they outline two basic out of sample tests for nested models each with two different probability distribution. The first being the Mean square error test (MSE), in this test the null is states the two models being estimated have equal prediction accuracy. This can be viewed as an out of sample Granger causality test. They include an F and T test for this MSE test. They show that MSE-F has the most power out of the two. As the value of out of sample increases until T, the MSE-F test converges to normal and becomes the in sample Granger causality test. They also note that in sample Granger causality has the greatest power.

We test three models:

\[ R_{jt} = \gamma_0 + \gamma_1 R_{mt} + \gamma_2 s_{ijt} + u_{ijt} \quad \text{where} \quad \gamma_2 = 0 \quad (9) \]

\[ R_{jt} = \gamma_0 + \gamma_1 R_{mt} + \gamma_2 s_{ijt} + \gamma_3 M_{ijt} + u_{ijt} \quad \text{where} \quad \gamma_2 = 0 \quad (10) \]

\[ R_{jt} = \gamma_0 + \gamma_1 R_{mt} + \gamma_2 s_{ijt} + \gamma_3 M_{ijt} + \gamma_4 F_{mt} + u_{ijt} \quad \text{where} \quad \gamma_2 = 0 \quad (11) \]

where We select a one for one ratio of in sample to out of sample. The results are shown in Table (number).

7 Conclusion

References


Table 4: Full Sample OLS

<table>
<thead>
<tr>
<th></th>
<th>EQ(6)</th>
<th>EQ(7)</th>
<th>EQ(8)</th>
<th>EQ(9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable: Stock Returns</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>returnsp500</td>
<td>1.023***</td>
<td>1.023***</td>
<td>0.935***</td>
<td>-3.459***</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.032)</td>
<td>(0.257)</td>
</tr>
<tr>
<td>s_it_mean</td>
<td>0.002***</td>
<td>0.002***</td>
<td>0.002***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>markcap</td>
<td>0.000</td>
<td>-0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MktRF</td>
<td>0.003***</td>
<td>0.046***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.003)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L.returns</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.000</td>
<td>0.000</td>
<td>-0.000</td>
<td>-0.000*</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>181,103</td>
<td>181,103</td>
<td>181,103</td>
<td>88,601</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.050</td>
<td>0.050</td>
<td>0.051</td>
<td>0.036</td>
</tr>
<tr>
<td>r2_a</td>
<td>0.0496</td>
<td>0.0498</td>
<td>0.0506</td>
<td>0.0356</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1
<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>returns</td>
<td>returns</td>
<td>returns</td>
<td>returns</td>
<td>returns</td>
</tr>
<tr>
<td>returns_ssp500</td>
<td>1.062***</td>
<td>1.061***</td>
<td>0.978***</td>
<td>-3.404***</td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.023)</td>
<td>(0.032)</td>
<td>(0.270)</td>
</tr>
<tr>
<td>s_jt_mean</td>
<td>0.002***</td>
<td>0.002***</td>
<td>0.002***</td>
<td>0.002***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>markcap</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>MktRF</td>
<td>0.002***</td>
<td>0.045***</td>
<td>0.002***</td>
<td>0.045***</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.003)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>L.returns</td>
<td>-0.000</td>
<td>-0.000</td>
<td>-0.000</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.000</td>
<td>-0.000</td>
<td>-0.000</td>
<td>-0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.000)</td>
</tr>
<tr>
<td>Observations</td>
<td>181,103</td>
<td>181,103</td>
<td>181,103</td>
<td>88,601</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.050</td>
<td>0.050</td>
<td>0.051</td>
<td>0.037</td>
</tr>
<tr>
<td>Number of ticker1</td>
<td>4.345</td>
<td>4.345</td>
<td>4.345</td>
<td>3.526</td>
</tr>
<tr>
<td>r2_a</td>
<td>0.0500</td>
<td>0.0501</td>
<td>0.0508</td>
<td>0.0366</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

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