Investigating Predictive Power of Stock Micro Blog Sentiment in Forecasting Future Stock Price Directional Movement

Nominated for Best Paper in ICIS 2011, Shanghai, China.

Chong Oh
Ph.D. Candidate
Department of Operations and Information Systems
David Eccles School of Business
University of Utah
1655 Campus Center Drive
Salt Lake City
UT 84112
chong.oh@business.utah.edu

Olivia R Liu Sheng
Presidential Professor, Emma Eccles Jones Presidential Chair in Business
Department of Operations and Information Systems
David Eccles School of Business
University of Utah
1655 Campus Center Drive
Salt Lake City
UT 84112
olivia.sheng@business.utah.edu

Abstract

This study attempts to discover and evaluate the predictive power of stock micro blog sentiment on future stock price directional movements. We construct a set of robust models based on sentiment analysis and data mining algorithms. Using 72,221 micro blog postings for 1909 stock tickers and 3874 distinct authors, our study reveals not only that stock micro blog sentiments do have predictive power for simple and market-adjusted returns respectively, but also that this predictive accuracy is consistent with the underreaction hypothesis observed in behavioral finance. We establish that stock micro blog with its succinctness, high volume and real-time features do have predictive power over future stock price movements. Furthermore, this study provides support for the model of irrational investor sentiment, recommends a complimentary investing approach using user-generated content and validates an instrument that may contribute to the monetization schemes for Virtual Investing Communities.

Keywords: stock microblogging, sentiment identification, stock movement, predictive models. user-generated content, financial prediction, data mining, web mining

Acknowledgement

Introduction

The Internet, as a whole, has become an enabler that aggregates vital information for stock investor decision making. It is changing how information is delivered to investors and the ways in which investors can act upon that information (Barber & Odean, 2001). In essence, it alters the way that investors invest, trade, acquire and share information (Zhang & Swanson, 2010). Initially, it was more of aggregating public information such as financial data, market updates and public news. More recently, with the advent of social media and WEB 2.0 (Ullrich et al., 2008), user generated content (UGC) are incorporating private information in addition to public information (Tumarkin & Whitelaw, 2001). Thus we observe how such virtual investing communities (VIC) as Yahoo Finance and Raging Bull are publishing relevant and valuable UGC data such as investment recommendations and proprietary analysis. UGC in these channels enriches investors' ability in making better investing decisions by allowing investors to monitor the thought process and decision makings of others. Thus it is imperative for researchers and practitioners to understand how individuals in virtual communities interact with one another and how these behaviors relate to future predictive outcomes.

This is indeed an active area of research with many interesting streams of literature. One involves interactions among members of the communities focusing on topics such as homophily, network externalities and reputation (Chen et al., 2009; Gu et al., 2007; Zhang, 2009). Another correlates activities in virtual communities to stock market outcome investigating predictors such as volume, disagreement and bullishness of postings (Antweiler & Frank, 2004; Das & Chen, 2007; Das et al., 2005; Sabherwal et al., 2008). However, despite prevalent belief that sentiments from VICs have predictive value, scholars have provided little evidence that these sentiments play any significant role in predicting stock price movements (Tumarkin & Whitelaw, 2001; Das & Chen, 2007; Antweiler & Frank, 2004). Our study attempts to address this research gap by examining a popular stock microblogging channel, Stocktwits.com (http://www.stocktwits.com), with its succinct, real-time and high volume postings. We posit that these distinct features of stock micro blogs may present a nascent set of dimensions in uncovering the predictive value of online investor sentiment.

What is microblogging? Microblogging is a new social networking and blogging phenomenon on the Internet. Wikipedia defines microblogging as "A form of multimedia blogging that allows users to send brief text updates or micromedia…. to be viewed by anyone or by a restricted group." (Wikipedia, 2010a). The current most popular microblogging platform is Twitter (www.twitter.com) with over 100 million registered users (Romero et al., 2010). Most microbloggers converse about daily activities and seek and share information (Java et al., 2007) through short updates or postings up to 140 characters long. This constraint induces postings to be brief and succinct in response to the question, "What are you doing right now?" This reduction of time and thought investment increases the frequency of updating a typical micro blog, from one a day, as in regular blogging, to multiples per day, in microblogging (Java et al., 2007). In addition, microblogging's real-time feature is a key factor in its popularity (Claburn, 2009).

Why stock microblogging? StockTwits is a social, stock microblogging service that was established in October 2008. It is a variant platform of Twitter that aggregates only stock-related stock micro blog postings. This service allows users to monitor the activities of traders and investors, contribute to the conversation and build reputation as savvy market wizards (TechCrunch, 2010). It is a marketplace for investing ideas that allow amateurs to interact freely with professionals. It currently has more than 150,000 subscribers who spend an average of 32 minutes and contribute 20,000 messages per day (Zeledon, 2009). Stocktwits was recently awarded Time magazine's top 50 websites for 2010 (Time, 2010).

To study the predictive value of stock micro blog sentiment over future stock price movements, we put forward the following research questions:
1. How well can stock micro blog sentiment predict future directional stock price movements? Specifically, how well can a bullish (or bearish) sentiment extracted from a stock micro blog predict a future upward (or downward) stock price movement?
2. Is there any difference in the predictive power between bullish and bearish postings?

3. Can stock micro blog's predictive power and its difference between bullish and bearish postings be explained by any existing theoretical framework?

We propose a sentiment analysis (Das & Chen 2007) and predictive analytics (Shmueli & Koppius 2011) approach to extract relevant features in building a robust set of models in understanding the predictive relationship between stock micro blog sentiments and future ten days of stock price movements. We first manually label sentiment polarity of a subset of postings. Then using sentiment analysis and the labeled postings, we identify features from written text of micro blog to automatically predict sentiment polarity of other postings. Then we aggregate postings for each ticker on a daily basis. Thereafter we use data mining classifiers to build models from aggregated features of sentiment, posting, author, ticker and market to predict future stock price movements. Overall, our study shows that micro blog sentiments do contain valuable information for investing decision making and supports the investor sentiment hypothesis that irrational investors do influence market prices.

Related Work

We review three literature streams to identify the research gap for this study. These areas are: UGC in marketing, UGC in VICs and UGC in microblogging.

UGC in Marketing

One active area of research from the marketing discipline relating consumer behavior to economic outcome is closely associated with our study. It is popular for scholars in this area to study consumer behavior in the forms of customer reviews or electronic Word-of-mouth (eWOM), user ratings and blogs. The Internet's ability to reach out to vast audience at low cost has presented a new significance for Word-of-mouth (WOM) as a tool to influence and build trust (Dellarocas, 2003). We perused through a few notable studies in this area.

Chevalier & Mayzlin (2006) examined effect of consumer reviews on sales of books for two online rivals, Amazon.com and BarnesandNoble.com. The authors confirmed that customer WOM affects consumer purchasing behavior by observing existence of relationship between number of reviews and average ranking with sales outcome. Duan et al., (2008) studied the relationship between online user reviews for movies and box office sales. They found that box office sales are influenced by volume of postings but the rating of reviews has no significant impact on sales. They concluded that businesses should focus more effort on facilitating consumer WOM exchanges instead of user ratings. Godes & Mayzlin (2004) tested Usenet WOM community interactions with ratings of new television shows. They found that WOM that is more dispersed across communities may be better than WOM that concentrates within each community. Dhar & Chang (2009) investigated blog posts with sales of music. They found that blog posts volume correlates positively with future sales of music CDs.

This sample of literature confirms that volume of UGC does have significant positive correlation with economic outcomes. However findings regarding user rankings or sentiment is inconsistent. With this, we next review UGC in virtual investing communities (VIC) to understand its relationship with regard to investing outcomes.

UGC in Virtual Investing Communities

Virtual investing communities (VIC) are a popular social media for online investors. It has blossomed with the growth of the Internet and its popularity stems from providing an environment where investors can collaborate and discuss, monitor what others are doing, or simply to seek fellowship (Wasko & Faraj, 2005). We peruse a few samples of studies undertaken to understand the relationship between behavior of community participants and stock market outcomes.

One of the earlier studies in this area is from Wysocki (1998), which used a sample of 3,000 stocks on Yahoo! message board, and found that previous day returns, changes in trading volume, and changes in

previous day postings have no predictive ability on stock returns. He did, however, found that an increase in volume of overnight postings correlated to a 0.18% average abnormal return. In addition, he concluded that total posting volume is higher for firms with high short-seller activity, extreme past stock returns and accounting performance, higher price earnings and book-to-market ratios, higher past volatility and trading volume, higher analyst following, and lower institutional holding (Wysocki, 1998). In another study, Tumarkin and Whitelaw (2001), using 181,000 postings from RagingBull.com found that, in general, message board activity does not predict industry-adjusted returns or abnormal trading volume. However they found that it is possible to predict the number of postings using previous day's trading volume, number of postings and weighted opinion (Tumarkin & Whitelaw, 2001).

A well-referenced paper, Antweiler and Frank (2004), using 1.5 million postings from Yahoo! Finance and RagingBull.com message boards, found significant but negative contemporaneous correlation between number of postings and stock returns on the next day. The return, however, is economically very small in comparison to transaction costs. Nevertheless, message posting activities do help to predict volatility and trading volume. In addition, the authors concluded that volume of postings is positively correlated with volatility and bullishness. Similarly, Koski et al., (2004), apart from confirming that day traders are noise traders, also found that day-trading volume increases volatility but concluded no predictive relationship with stock returns. Das and Chen (2007) developed a methodology using five classifier algorithms to extract sentiment from stock message boards but found no significant predictive relationship between sentiment and stock prices. However, consistent with findings of Antweiler and Frank (2004), Das and Chen (2007) reaffirmed the existence of a significant correlation between posting volume and volatility but asserted that sentiment does not predict stock movements. Interestingly, Das et al. (2005) found that sentiment does not predict returns but instead returns drive sentiments. They implied that members of virtual community are more likely to extrapolate past returns rather than to be contrarian, which ultimately leads to a behavior consistent with the representativeness heuristic (Das et al., 2005; Lakonishok et al., 1994; Kahneman & Tversky, 1973).

Sabherwal et al. (2008) downloaded 160,000 postings from TheLion.com stock message board and conducted an event study to estimate daily abnormal returns. The authors found that posting volume positively correlates with stock's abnormal return on the same day and also predict next day's abnormal returns. They concluded that online investors focused on thinly traded micro-cap stocks with low institutional holdings and low analyst coverage.

Overall, although significant correlations exists between VIC activities such as posting volume and stock market movements, prior literature has yet to establish existence of any predictive power between sentiment and stock market outcome. This is the research gap we seek to answer through investigating the relationship between sentiments of stock micro blog postings with future stock price movements.

UGC in Microblogging

Although microblogging is a nascent UGC channel, a few scholars have attempted to examine the relationship between predictors extracted from micro blogs with future outcomes such as movie revenue, events and stock prices. For example, Bollen et al. (2010a) used extended version of Profile of Mood States (POMS) to extract six mood dimensions from over 9 million Twitter postings. They aggregated mood components on a daily scale and compared them to the timeline of cultural, social, economic and political events in the same time period. They found significant correlations between extracted mood dimensions and those occurring events. Bollen et al. (2010b) further expanded their prior study in Bollen et al. (2010a) specifically towards predicting DJIA index over the same time period and conclude an accuracy of 87.6%. Another example is Asur and Huberman (2010) which extracted sentiment from 6 million Twitter postings to predict box office revenue for movies. They benchmarked against Hollywood Stock Exchange (HSX) and obtained an accuracy of 0.94.

Research correlating predictors in microblogging with future outcomes is still in its infancy. In this study we seek to understand the relationship between sentiment of stock micro blogs with future stock price movements.

Theoretical Foundation

We present three theoretical discussions to support our assertion on the existence of predictive power of stock micro blog, the difference between predictive power of bearish and bullish sentiment and the relationship to the underreaction hypothesis.

Predictive Power of Stock Micro Blog Sentiment

We assert that stock micro blog has predictive power over future stock price movements. We first discuss microblogging distinct features as a support for this assertion. These three distinct characteristics, succinctness, high volume and real-time, greatly facilitate the diffusion of investing information in the community (Java et al., 2007; Bollen et al., 2010a). Since users may only send short updates or postings up to 140 characters long, postings are brief and succinct by design. This may reduce noise and transmits a more relevant message to the recipients. Additionally, this reduction of time and thought investment lead to high volume of postings by increasing the frequency of updating a typical micro blog, from one a day, as in regular blogging, to multiple per day, in microblogging (Java et al., 2007). Furthermore, postings are usually real-time as they are posted very close to the occurrence of the events. Being real-time is a key factor in microblogging's popularity (Claburn, 2009) because over time information becomes less relevant and less useful for decision and planning purposes (Ballou & Pazer, 1995). Furthermore, since public investing information such as company press releases, earning announcements and analyst recommendations is sporadic and infrequent, the continuous streaming of micro blog represents new information at a constant and greater temporal frequency to the investors that is otherwise unavailable. This would blend naturally with stock trading activities. Hence, we postulate that these three microblogging features: succinctness, high volume and real-time, positively contribute to the predictive power of micro blog sentiment.

The second reason for our assertion stems from the behavioral finance literature. There are two types of traders in the financial markets, the irrational noise traders or day traders (Koski et al., 2001) who hold random beliefs about future dividends--and the rational arbitrageurs who hold Bayesian beliefs (DeLong et al., 1990). Noise traders frequently engage in discussions or conversations on investing information. Conversation, in the context of online investing, involves discussing alternatives, making predictions, asking questions, reporting observations, contributing opinions, sharing analysis and announcing decisions. Conversation is critical in the contagion of popular ideas about financial markets as people tend to pay more attention to ideas or facts that are reinforced by conversations, rituals and symbols (Hirshleifer, 2001). DeLong et al. (1990) stated that when sentiments of noise traders are correlated with each other, they create risk. Furthermore, the noise trader literature asserts that presence of noise traders increase volatility (Koski et al. 2004). In short, noise traders have the ability to affect stock prices whenever information can be shared among investors and spread quickly through web channels (Zhang & Swanson 2010). Microblogging is one such channel where sentiments of irrational investors spread quickly with the continuous streaming of information. Specifically sentiment, represented by opinions, plays a vital role in this diffusion.

Based on these discussions on features of stock micro blogs and impact of irrational noise traders, we postulate that stock micro blog sentiments extracted from online investor conversations have predictive power to influence future stock price movements. Hence we postulate that:

> H1: The predictive performance of a model with sentiment is higher than a model without sentiment.

Difference between predictive power of bullish and bearish sentiment

Bearish sentiment is pessimistic or negative news as it implies declining stock prices. Bullish sentiment, which is related to overconfidence and over-optimism, on the other hand, often lead to wishful thinking, a phenomenon that is related to speculative bubbles and information mirages (Seybert & Bloomfield, 2009). "Wishful thinking is the formation of beliefs and decision making according to what might be pleasing to imagine instead of appealing to evidence, rationality and reality" (Wikipedia, 2010b). Wishful thinking is highly influential in markets where many traders who each hold a small bit of information have to rely on inferences from observed behavior in order to estimate asset values (Seybert & Bloomfield, 2009). Since VICs have such characteristic, they are susceptible to wishful thinking. Thus we postulate that bullish sentiment to have a lower predictive accuracy than that of bearish sentiment.

In addition, people pay more attention to negative than positive news (Luo, 2007). According to prospect theory (Kahneman & Tversky, 1979), losses weigh more than gains. Chevalier & Mayzlin (2006) also echoed that an incremental negative review is more powerful in decreasing sales than an incremental positive review is in increasing sales. Essentially negative words have more impact and are more thoroughly processed than positive information (Baumeister et al., 2001; Rozin & Royzman, 2001).

Both the phenomena of wishful thinking and impact of negative news support our assertion that bearish postings should have a higher predictive accuracy than bullish postings. Thus, we hypothesize that:

> H2: The predictive performance for postings of bearish sentiment is higher than that of bullish sentiment.

Underreaction Hypothesis

Is there a theoretical framework to explain the presence of predictive power in stock microblogging? Prior findings confirm that stock messages such as from Yahoo Finance message board reflect current news extremely rapidly (Antweiler & Frank, 2004). With close similarity between stock micro blogs and message board postings, we stipulate that stock micro blogs also exhibit the same behavior in reflecting news. However, scholars emphasized that market prices do not fully incorporate stock posting information (Gu et al., 2007), hence this information has predictive power over future returns. Stock prices tend to drift or underreact after news for a period of time (Chan, 2003) because investors tend to hold too strongly to their own prior information and discount public signals (Daniel et al., 1998) influenced by conservatism heuristic (Barberis et al., 1998). Further, Hong and Stein (1999) stated that this underreaction hypothesis result in a hump-shaped equilibrium impulse response (hump-shaped curve) of the stock price movement. This hump-shaped curve is measurable by the underreaction coefficient (URC) as defined by Cohen & Frazzini (2008).

The Efficient Market Hypothesis (EMH) stipulates that security prices always fully reflect all available information (Fama 1970). Hence we would expect predictive accuracy to be at its peak the day of posting ($t=0$) and starts to decline thereafter. However, with the effects of underreaction, we may observe that since investor exhibit conservatism after receiving a particular sentiment, the prediction accuracy may not be at its peak until a few periods later. Only after the new information is largely absorbed into the stock price, will the predictive accuracy drop from peak. Based on this discussion, we postulate that prediction accuracy for both simple and market-adjusted return to exhibit underreaction hypothesis in the short run. Thus, we hypothesize that:

> H3: The predictive performance of microblog features for both simple and market-adjusted returns form a hump-shaped curve measured by the underreaction coefficient (URC) where URC>0 but URC<1.

Data & Methodology

Data

The primary data for this study was downloaded from Stocktwits.com (http://www.stocktwits.com) and Yahoo Finance (http://finance.yahoo.com) for the period May 11, 2010 to August 7, 2010 (89 days). We obtained over 208,278 stock micro blog postings, 5,981 author profiles, and 720,840 inter-day stock price posts.

Stock micro blog postings were pre-processed -- those without any ticker, more than one ticker, or not in NASDAQ and NYSE exchanges were removed leaving 72,221 valid postings, consisting 1,909 stock tickers; 1,474 in NASDAQ and 435 in NYSE, posted by 3,874 distinct authors. A list of top 10 stock tickers with corresponding number of postings is shown in Table 1 while a description of all attributes is in Table 2. Interestingly, top 10% of all the stock tickers are responsible for over 70% of all postings. These are popular stocks, consistent with the finding that people invest in the familiar while often ignore principles of portfolio theory (Huberman, 2001).
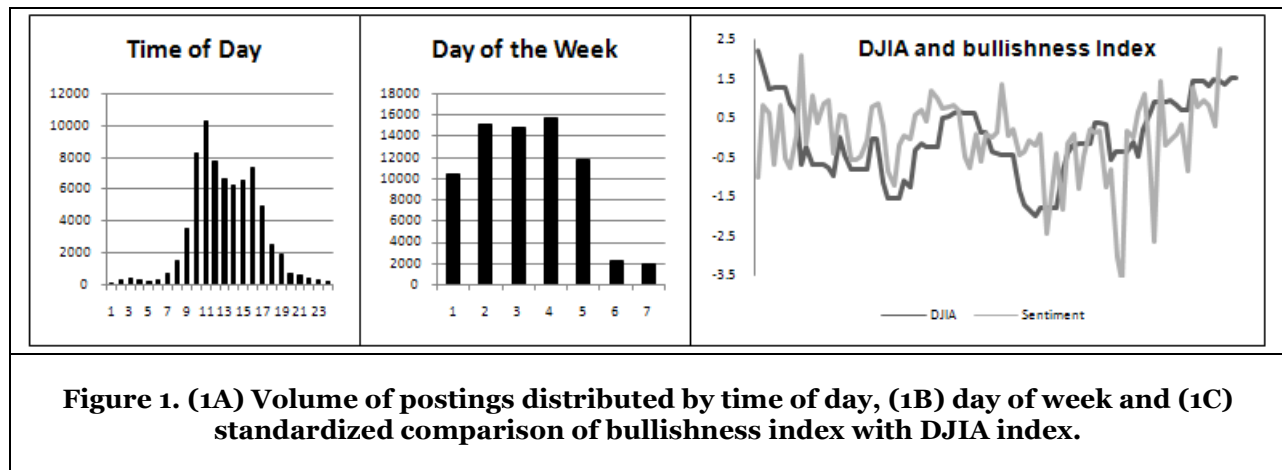
	Ticker	Total	Exchange
	Table 1. Distribution of postings by top 10 tickers.		
1	AAPL	7212	NASDAQ
2	BP	6220	NYSE
3	GS	2502	NYSE
4	GOOG	1803	NASDAQ
5	NFLX	1576	NASDAQ
6	RIMM	1558	NASDAQ
7	GLD	1532	NYSE
8	BIDU	1373	NASDAQ
9	AMZN	1071	NASDAQ
10	FAS	827	NYSE

| Variable	Description
Table 2. Description of posting attributes	
Sentiment	0-neutral , 1-bullish, -1-bearish (manually labeled)
Posting	Posting id, post date, day of the week, time of the day, market hours, text of posting.
Author	Expert, bio, url, location, follower, following, total postings, posting per day, retweet, direct, mention, etc.
Ticker	Exchange, volume, past 5 days closing prices and volumes.
Market	Past 5 days DJIA index.

In line with Antweiler & Frank (2004), Figure 1A shows that most postings occur during working hours, typically between 10:00 am and 4:00 pm, indicating high day trading activity. This is analogous with Antweiler & Frank (2004) and Cao et al. (2002)'s assertion that most people post messages within minutes of conducting a trade. Consistent with earlier studies, we find high volume of postings during trading days, peaking on Thursdays, and low activity during weekends and holidays (Figure 1B). This illustrations support the assertion that the distribution of the postings is highly representative of investor behavior.

Figure 1. (1A) Volume of postings distributed by time of day, (1B) day of week and (1C) standardized comparison of bullishness index with DJIA index.

Another interesting observation is illustrated in Figure 1C which highlights the similarity of the trends between standardized Dow Jones Industrial Average (DJIA) and Stocktwit's community bullishness index, an aggregated stock sentiment measure to be explained in the later section. This correlation alludes to the predictive power of stock micro blog sentiment. In addition, note that even during times of declining trends, there was still more bullish than bearish sentiment overall represented by overwhelmingly positive bullishness index. This may be explained by the over-confidence and over-optimism biases discussed earlier indicating that online investors are overly optimistic even during market downturns.

System Pipeline

The system pipeline diagram, illustrated in Figure 2, outlines the overall methodology for this study. There are essentially five phases: 1. Downloading data, 2. Pre-processing, 3. Sentiment Analysis, 4. Prediction classification, and 5. Evaluation and analysis.

Figure 2. System Pipeline
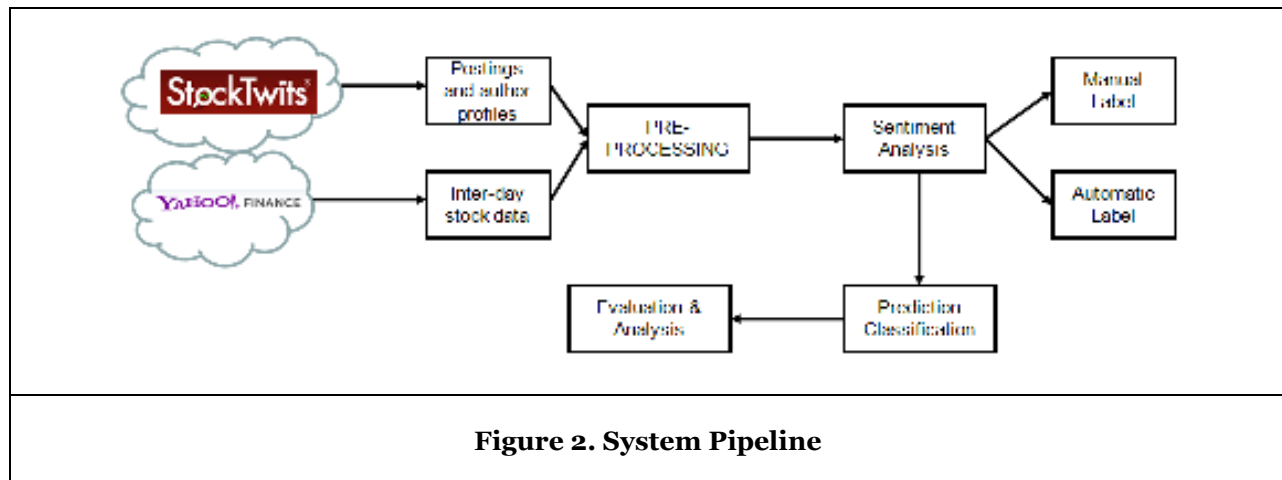
Automatic Extraction of Micro Blog Sentiment

The key predictor in our study is the sentiment extracted from each micro blog posting. Since these sentiments are derived from postings and are not explicitly provided by the authors, we have to manually and automatically extract them. We first manually label 7,109 postings (about 10% of all valid postings) to three distinct sentiments, 1 for bullish, -1 for bearish, and 0 for neutral sentiment--referred to as "manual labels". This initiative was with the consensus of three ressearchers so as to maintain consistency and reliability. This gold standard is used to evaluate the result of the automatic (system) labeling process on the remainder of the postings. We use the lexical scorer (Kim & Hovy, 2006), bag of words (BOW) (Schumaker & Chen, 2009) and a combination of both approaches to automatically label the 65,112 remaining postings--referred to as "system labels". The lexical knowledge involves scoring every word in each posting with a manually crafted lexicon of bullish and bearish keywords where bullish word is 1 while bearish word is -1. An aggregate score is then derived for each posting. The bag of words (BOW) involves generating a feature vector of all words in the posting set and then a machine learning algorithm J48 classifier (10-fold cross validation) is used to produce a learning model. The combination approach uses words produced by the Lexical classifier to be generated into a feature vector. Table 2A displays the results of the 3 approaches. The BOW approach has the highest score of .663 F-measure. Due to the challenging nature of sentiment classification, this result is encouraging as accuracy range of between .6 to .7 is already considered good (Das & Chen 2007; Pang, Lee & Vaithyanathan, 2002). A few examples of the top features generated by BOW are trading words like short, sold, long, out and bought.

Table 2A. Results of Sentiment Analysis					
	True Positive	False Positive	Precision	Recall	F
Lexical	.603	.229	.639	.603	.596
BOW	**.664**	**.191**	**.663**	**.664**	**.663**
Combo	.634	.18	.663	.634	.639

We opt for BOW model to automatically assign labels for the remainder unlabeled postings producing results of bullish (30018), bearish (16732), and neutral (18362) postings. Since only postings with sentiment are useful as per Antweiler and Frank (2004) we removed neutral postings obtaining 46750 system labeled postings (30018 + 16732) for prediction classification. At this juncture, we divide both manual and system labeled dataset into two parts: in-sample and hold out set. This data are then aggregated for daily ticker level analysis.

Aggregation of Sentiment

To facilitate analysis at daily ticker level, we generate aggregated sentiment bullishness for each ticker for each day. To do this we adopted the bullishness index that was introduced by Antweiler & Frank (2004).

$$\text{Bullishness Index} = \ln\left[\frac{1 + M^{BULL}}{1 + M^{BEAR}}\right]$$

M^{BULL} is the total bullish postings while M^{BEAR} is the total bearish postings. An index that is more than 0 is bullish, while 0 is neutral and less than 0 is bearish. This index accounts for large number of postings expressing a particular sentiment. Antweiler and Frank (2004) found this measure to be most robust out of three sentiment indices proposed in their study and is well used in other studies.

Aggregation of daily ticker

To study the impact of UGC sentiment on future stock price movements, we aggregate postings into daily ticker level analysis. A brief description of attributes for ticker level aggregated data is outlined in Table 3 below. These are the independent variables for our models. Due to the nature of the analysis neutral labels are removed as they do not have any discriminatory contribution (Antweiler & Frank 2004). We ended up with manual labels consist of 1148 records for the in-sample set and 230 records for hold-out set while the system labels consist of 11000 for in-sample and 1141 for the hold-out set (Table 4). The approach of evaluating with hold-out set is an important feature of modal fitting where the period of fit (in-sample) is separated from the period of evaluation. Specifically the hold-out set is from a period in the future, used to compare the forecasting accuracy of models fit to past data (SAS, 2011). Note that the manual labeled postings are not randomly selected due to the design constraint of our dependent variable.

Table 3. Description of attributes for daily ticker	
Variable	**Description**
Bullishness Index	Aggregated sentiment index where 0-neutral , >0-bullish, <0-bearish
Posting	Average posting volume, day of the week, time of the day, market hours, message length.
Author	Average expert, bio, url, location, follower, following, total postings, posting per day, retweet, direct, mention, etc.
Ticker	Exchange, volume, past 5 days closing prices and volumes.
Market	Past 5 days DJIA index.

Priors for manual labels and system labels are outlined below:

Table 4. Priors for Manual and System Daily Ticker Labels				
	Manual		**System**	
	In-sample	**Hold-out**	**In-sample**	**Hold-out**
Total labels	1148	230	11000	1141
Bullish labels	796 (.69)	173 (.75)	7690 (.7)	873 (.77)
Bearish labels	352 (.31)	57 (.25)	3310 (.3)	268 (.23)
Period	May 11 – May 17	May 18 – May 19	May 19 – July 31	Aug 1- Aug 8

We generate two types of dependent variables – simple return and market-adjusted return. These dependent variables are derived by matching bullishness index per ticker against the stock price movements for future 10 days. Tetlock (2008) stated that five days is a good cutoff time for considering informational events to be potentially related and seven is natural for news to be diffused. In this study, we extend to 10 future days for better coverage. The explanation is described below:

Table 5. Dependent Variables	
Simple return (1,0)	Prediction outcome matching bullishness index (sentiment) from ticker level aggregated postings, bullish (>0) or bearish (<0), against actual stock price movement, upward (1) or downward (-1) trend. 1 for correct prediction, 0 for incorrect prediction. Each record has 10 DVs, each for period *(t+1)* to *(t+10)*.
Market-adjusted return (1,0)	Prediction outcome as described in Simple return above but accounts for market return (Dow Jones Industrial Average). Specifically, the system matches ticker level sentiment against the net trend between simple return and market return. 1 for correct prediction, 0 for incorrect prediction. Each record has 10 DVs, each for period *(t+1)* to *(t+10)*.

Stock Price Prediction Classification

In this study we analyze predictive value of stock micro blog postings by daily aggregated ticker level analysis. The evaluation output is based on directional accuracy approach (Schumaker & Chen, 2009) by matching the predicted movement against actual stock movement trends. This approach of correlating to

future stock prices is based on Gu et al. (2007)'s remark that market prices do not fully incorporate stock posting information; hence, this information has predictive power over future returns.

Models

We attempt to answer research questions H1 and H2 with three predictive models: M1, M2 and M3. The main Model M1 is the de facto model encompassing all the predictor variables (Table 3) in addition to micro blog sentiment, measured by aggregated bullishness index. The second model, Model M2 is M1 without bullishness index while Model M3 is M1 with additional past five days' bullishness index metric. M3 is included to test the current findings that lagged sentiment is correlated with stock indicators (Das & Chen 2007; Antweiler & Frank 2004). They are tested over 4 datasets: manual label simple return (manual-simple), manual label market-adjusted return (manual-market), system label simple return (system-simple) and system label market-adjusted return (system-adjusted).

Performance measures

All classifications are performed using Weka data mining software (Whitten & Frank, 2005) with 10-fold cross-validation to strengthen validity of the results. In line with standard metrics from information retrieval (Ma et al., 2009), we report precision, recall and F-measure measures (Whitten & Frank, 2005) to evaluate the performance of the predictive models.

There are three types of F-measures in our results: weighted, class-1 and class-0. Weighted is the average while class-1 measures the accuracy of the correct predictions and class-0 measures accuracy of incorrect predictions. For most of our analysis we used the weighted F-measure as the default accuracy measure. However when comparing accuracy of bearish and bullish postings we used class-1 F-measure instead. This is because class-1 F-measure is a measure of accuracy of correct predictions which is a measure that is more appropriate in that context.

Results

Baseline Classifications

We first process manually labeled dataset using default Model M1 (in-sample) with 8 types of classifiers to obtain a baseline for the best classifier and predictive accuracy with reference to the majority rule benchmark. Majority rule is when the majority class (1 or 0) decides the accuracy of the entire dataset and is the benchmark chosen for our study. The 8 classifiers are Adaboost, Bagging-J48, Cost sensitive, J48, Logistic, Random Forest, SMO and ZeroR where ZeroR is the benchmark majority rule classifier for this study. Table 6 shows the result for simple return where SMO is the best classifier with weighted F-measure of .853. Incidentally, all classifiers performed better than the benchmark classifier: ZeroR (F=.495). The overall accuracy results provide support that features from micro blog postings do have predictive power. Period (t+6) is the best with weighted F-measure of .863. A pairwise t test was conducted but yielded no significant difference (p>.05) between SMO and Bagging-J48, thus we proceed to use Bagging-J48 for future classifications as it is efficient and reliable, especially for processing of large datasets.
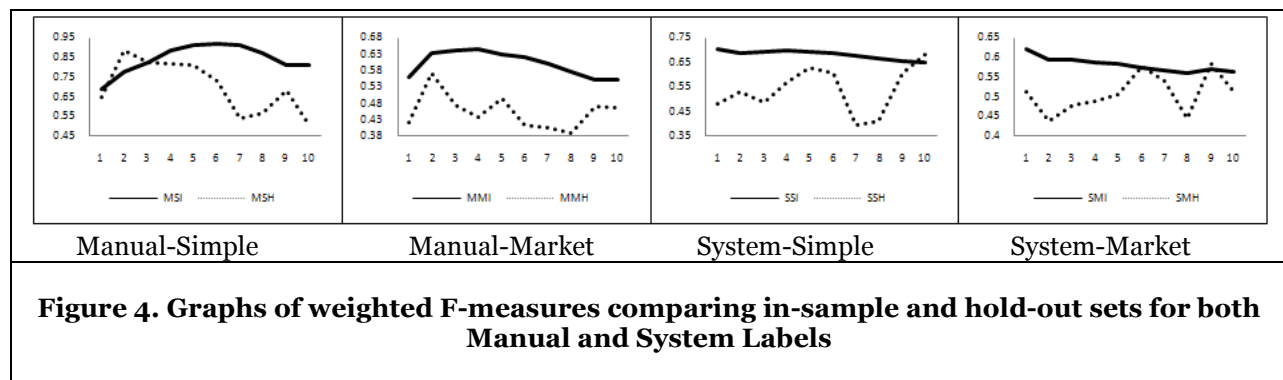
Table 6. Weighted F-measures for 8 classifiers on M1 Model									
Period	Adaboost	Bag	Cost	J48	Logistic	RF	SMO	ZeroR	Mean
1	0.666	0.691	0.601	0.676	0.694	0.677	0.724	0.429	0.645
2	0.730	0.782	0.736	0.746	0.792	0.753	0.796	0.442	0.722
3	0.792	0.827	0.789	0.818	0.838	0.797	0.841	0.466	0.771
4	0.867	0.889	0.862	0.867	0.881	0.869	0.890	0.507	0.829
5	0.891	0.915	0.888	0.911	0.908	0.884	0.909	0.529	0.854
6	0.903	0.921	0.901	0.926	0.912	0.891	0.915	0.535	**0.863**
7	0.893	0.917	0.889	0.915	0.908	0.889	0.910	0.527	0.856
8	0.843	0.874	0.829	0.870	0.868	0.847	0.878	0.521	0.816
9	0.791	0.813	0.765	0.814	0.830	0.782	0.837	0.507	0.767
10	0.773	0.814	0.743	0.808	0.825	0.773	0.828	0.483	0.756
Mean	0.815	0.844	0.800	0.835	0.846	0.816	**0.853**	0.495	
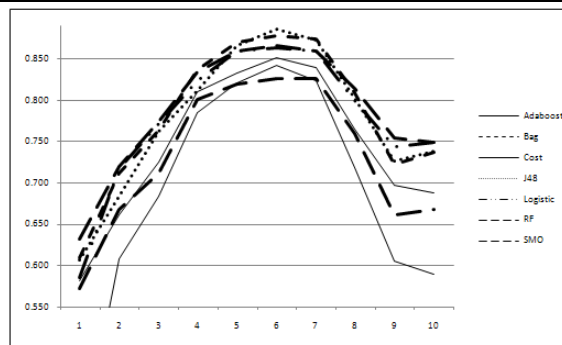
Figure 3. Weighted F-measures for 8 classifiers on M1 Model

We then compare predictive accuracy for the in-sample and hold-out set for four groups of datasets (Table 7): manual-simple, manual-market, system-simple and system-market using Bagging-J48 with M1 model. Overall the results show existence of a higher predictive value for simple returns (Manual-simple = .8443 and system-simple = .6827) but lower for the market adjusted return (Manual-market = .6031 and system-manual = .5815). This confirms the fact that it is a significant challenge to consistently beat the market as previously outlined by other scholars (Brock et al., 1992; Borodin et al., 2004). Nevertheless, the encouraging accuracy obtained from simple return implies an optimistic note that micro blog sentiment does positively impact the return of investors. We also observe the presence of hump shaped curves in manual labels but not in the system labels (Figure 4).

	Manual				System			
	Simple		Market		Simple		Market	
Period	In sample	Hold out	In sample	Hold out	In sample	Hold out	In sample	Hold out
1	0.691	0.65	0.563	0.421	0.705	0.482	0.621	0.513
2	0.782	0.889	0.636	0.571	0.689	0.531	0.595	0.439
3	0.827	0.825	0.641	0.473	0.694	0.484	0.595	0.478
4	0.889	0.817	0.647	0.437	0.702	0.567	0.587	0.489
5	0.915	0.81	0.63	0.494	0.695	0.631	0.583	0.505
6	0.921	0.731	0.624	0.413	0.689	0.604	0.574	0.576
7	0.917	0.535	0.602	0.406	0.678	0.394	0.567	0.538
8	0.874	0.567	0.58	0.39	0.668	0.413	0.56	0.446
9	0.813	0.683	0.554	0.471	0.655	0.609	0.571	0.583
10	0.814	0.504	0.554	0.466	0.652	0.683	0.562	0.511
Mean	0.8443	0.7011	0.6031	0.4542	0.6827	0.5398	0.5815	0.5078

Table 7. Weighted F-measures comparing in-sample and hold-out sets for both Manual and System Labels

Manual-Simple   Manual-Market   System-Simple   System-Market

Figure 4. Graphs of weighted F-measures comparing in-sample and hold-out sets for both Manual and System Labels

(H1) Predictive Value in stock micro blog sentiment

With the baseline ascertained, we examine the effect of sentiment on predictive accuracy by testing both manual and system labeled datasets on the 3 models (M1, M2 and M3). We find that sentiment has a strong influence on the predictive power, as results drop significantly when sentiment is withdrawn from the predictive model in comparing M1 with M2 and increased in comparing M1 and M3 (Table 8). Weighted F-measure dropped from .8443 (MM1) to .5819 (MM2) and .6827 (SM1) to .5501 (SM2). Adding past 5 days sentiment as predictors increased the prediction from Weighted F-measure of .8443 (MM1) to .851 (MM3) and .6827 (SM1) to .7041 (SM3). It is clear that sentiment have strong predictive influence. This finding supports hypothesis H1 indicating that predictive model with sentiment has higher accuracy than model without sentiment.

We support the tenet that irrational investor conversations and such distinct features of microblogging as succinctness, high volume and real-time contribute to the predictive value of micro blog sentiments. In short, valuable stock investing information, both private and public information, does exist in the sentiment of stock micro blog postings.

Table 8. Weighted F-measures of the 3 models						
	Manual			System		
Period	MM1	MM2	MM3	SM1	SM2	SM3
1	0.691	0.539	0.704	0.705	0.557	0.718
2	0.782	0.546	0.78	0.689	0.558	0.713
3	0.827	0.571	0.838	0.694	0.557	0.724
4	0.889	0.574	0.892	0.702	0.55	0.721
5	0.915	0.601	0.916	0.695	0.546	0.718
6	0.921	0.623	0.924	0.689	0.549	0.706
7	0.917	0.609	0.921	0.678	0.541	0.694
8	0.874	0.622	0.877	0.668	0.545	0.697
9	0.813	0.569	0.83	0.655	0.545	0.678
10	0.814	0.565	0.828	0.652	0.553	0.672
Mean	0.8443	0.5819	0.851	0.6827	0.5501	0.7041

(H2) Bearish versus Bullish

In comparing the accuracy of aggregated bullish and bearish daily ticker data, we analyzed two datasets: daily ticker bearish and bullish with J48 classifier on both manual and system labels. We first test using weighted F-measure and found no significant difference between bullish and bearish ($p>.05$). However when we test using class-1 F measure we found a significant difference. Class-1 F measure is appropriate as it focuses on the accuracy of correct predictions (class-1) – as explained previously. We realized that bull labels have a high class-0 F measure that artificially pushes up its weighted F measure. In other word, bull labels are predicting wrong predictions correctly (class-0). The result is shown in Table 9. Note that bearish has a higher class-1 F-measure as compared to bullish for all future 10 days for both manual and system labels ($p<.05$ as per pairwise t-test). This supports H2 stating that bearish labels have higher predictive accuracy than bullish labels supporting the fact that bearish or pessimistic information has higher predictive value and gets more attention as compared to bullish information. Furthermore, bullish sentiment is heavily influenced by phenomenon of wishful thinking that reduces its predictive accuracy.

Table 9. Class-1 F-measures for comparing bullish and bearish sentiments				
	Manual Simple		System Simple	
	bull	bear	bull	bear
1	0.28	0.787	0.463	0.695
2	0.212	0.907	0.441	0.779
3	0.11	0.943	0.374	0.808
4	0.124	0.955	0.377	0.782
5	0.133	0.957	0.332	0.722
6	0.147	0.954	0.344	0.697
7	0.123	0.957	0.406	0.673
8	0.137	0.921	0.447	0.651
9	0.084	0.89	0.516	0.603
10	0.095	0.905	0.566	0.582
mean	0.1445	0.9176	0.4266	0.6992

(H3) Underreaction Hypothesis

We test the underreaction hypothesis by measuring the underreaction coefficient (URC) (Cohen & Frazzini, 2008) and observing the presence of a hump-shaped curve (Hong & Stein, 1999) in the predictive value of the models for the manual labels. URC is defined by Cohen and Frazzini (2008) as "the fraction of total returns from month t to month t+last month that occurs in month t". A URC that is more than 0 and less than 1 is a sign of underreaction but URC outside this range is a sign of no underreaction. A low URC measure is a sign of severe underreaction. The intuition of URC is that total return of future periods (t>0) should be more than the first period (t=1). In our study we calculate simple and market-adjusted normalized returns for manual labels for t=1 to t=10 and calculate *URC = Return for t1/Total Return.

Table 10. Underreaction coefficients											
	URC*	1	2	3	4	5	6	7	8	9	10	Total
Simple	0.02	4.00	9.44	15.92	22.59	24.66	27.28	27.29	24.36	19.94	20.95	196.43
Market Adjusted	0.10	3.36	3.85	5.74	5.64	4.24	2.17	4.09	1.81	1.84	0.52	33.26

URC estimations (Table 10) show a clear presence of high underreaction for both simple (URC=.02) and market-adjusted (URC=.1) normalized returns. Furthermore, the hump-shaped curves are also observed for both simple and market-adjusted returns as shown in Figures 4 (Manual). This supports our assertion that the predictive power from stock microblog is explained by the underreaction hypothesis. As investors display conservatism, stock prices reflect the same. Thus our finding supports H3.

Discussion

"In models of investor sentiment on stock market prices, uninformed traders rely on various information sources in forming their beliefs. In equilibrium, their beliefs, albeit noisy, influence prices" (Xu & Zhang, 2009; De Long et al, 1990). Our study echoes the assertions of scholars that stock discussions are not noise (Antweiler & Frank, 2004; Sabherwal et al., 2008), and that they do have predictive value, and that their effect on economic outcome is real and substantial. In this section, we touch on two topics for further discussions: implications of this study and future directions.

Implications

This study has implications to the following three groups of stakeholders, namely, researchers, investors and managers.

Contribution to Research Community

This study contributes a decision support artifact using a nascent social media to the research community. The models and approaches constructed herein may become the groundwork for future research in understanding the impact of sentiment, specifically micro blog sentiment, on other economic outcomes such as firm financial performance, product and movie sales. Moreover Agarwal and Lucas, Jr. (2005) exhort IS scholars to initiate research with a greater macro focus in establishing a stronger IS credibility and identity. We heed this call in studying and explaining the transformational impact of a nascent IT artifact, stock microblogging in strengthening ties with reference disciplines, in particular in addressing the on-going debate between efficient market hypothesis (EMH) (Fama, 1970) and behavioral finance in supporting the latter in its tenet that investor sentiment does correlate with future stock price movements (De Long et al., 1990; Tetlock, 2007).

Contribution to Practitioners

There are two groups of practitioners addressed by this study: investors and managers. Investors, both individuals and institutions, have always been searching for effective approaches to predict stock

outcome. We present a nascent approach in predicting future stock movements by utilizing stock micro blog sentiment. This assists investors with an additional robust methodology in making informed and pre-calculated decisions in stock investing. In addition, this study helps platform providers such as Stocktwits.com or Twitter to know how to extract value from UGC. Sentiment may also be a viable criterion for hedge-fund managers to evaluate candidate tickers to be included in their portfolios.

Furthermore, managers may opt to design investing tools to enable investors in analyzing stock postings. For monetization purposes they may also incorporate advertisements with postings deemed with higher predictive value or even establish partnerships with selected authors of predictive postings. These findings may be valuable for UGC channels such as Twitter which has yet to establish any significant revenue stream (Miller, 2010).

Future Improvements

One extension to our study is to explore the magnitude of stock price change in addition to directional accuracy of stock movements. This understanding is important as even with a model that gets 80% of predictions correct, 20% wrong predictions could lead to huge losses which vastly undermine the value of the underlying model. In addition, it would be interesting to explore narrower time ranges for predictions such as minutes and hours after posting. Long-term time ranges such as weeks may also be important to examine as well. In addition, we would like to extend our study to predict market sentiment as well. This is expected to help improve the performance of ticker sentiment prediction.

Another extension of this study is to specify the sources of public and private information: news, analyst forecasts, stock message boards and financial blogs. This may or may not strengthen the predictive model but will reveal the relationship between these sources and predictive accuracy of the community. We could also include stock characteristics (small and large capitalization, etc.) in the model as well as other investing instruments such as foreign exchange and futures to further enrich the dataset and discover deeper relationships between predictive accuracy stock microblogging and characteristics of investment types.

Another pertinent extension is to explore the influence of social network features on future stock price movements. Intuitively we acknowledge that attributes of different authors may have different impact on their predictive power. This is due to their diverse experiences, investing knowledge and strategies. These features may be examined from two dimensions of social network: structural and relational. While structural refers to the position of the actor in the network, relational refers to the quality of the relationship among actors in the network (Grannovetter, 1992; Lin et al., 2010). Examples of these measures are outlinks and inlinks of retweets, direct postings, mentions, followers and tickers.

An on-going effort is to increase the accuracy of automated sentiment extraction process. We foresee much room to improve beyond the current F-measure of .67 and this will ultimately improve the overall predictive capability of the system. In addition, exploring different degrees of bullishness could also be a beneficial direction. This differentiates the level of author confidence thus may lead to increased predictive performance.

As a caveat, we acknowledge that our findings are based on data collected from a relatively short period of time--3 months. The predictive power may be biased by economic sway, cyclical stocks and other seasonal economic and political influences. Future work should cover a wider time range to establish validity.

Conclusion

In conclusion, investor sentiment as expressed in stock micro blogs postings does appear to have strong predictive value for future market directions. We concluded this by studying sentiment from 72,221 micro blog postings from Stocktwits.com, a stock microblogging service, over a period of three months. The

principal contribution of this study is to explain the transformational impact of a nascent IT artifact, stock microblogging. We also provide support for the model of irrational investor sentiment, to recommend a supplementary investing approach using user-generated content (UGC) and to propose an approach that may contribute to the monetization schemes for Virtual Investing Communities (VIC).

References

Agarwal, R. & Lucas, Jr., H. (2005) "The Information Systems Identity Crisis: Focusing on High-Visibility and High-Impact Research" *MIS Quarterly* 29(3):381-398.

Antweiler, W. & Frank, M. (2004) "Is all that talk just noise? The information content of internet stock message boards." *Journal of Finance.* 59(3):1259-1295.

Asur, S., Huberman, B. (2010) "Predicting the Future with Social Media" *HP Labs* Palo Alto, CA.

Ballou, D. & Pazer, H. (1995) "Designing Information Systems to Optimize the Accuracy-timeliness tradeoff" *Information Systems Research* 6(1):51.

Barber, B. & Odean, T. (2001) "The Internet and the Investor." *Journal of Economic Perspective* 15:41-54.

Barberis, N., Shleifer, A. & Vishny, R. (1998) "A model of investor sentiment" *Journal of Financial Economics* 49:307-43.

Baumeister, R., Bratslavsky, E., Finkenauer, C., & Vohs, K.D. (2001). "Bad is stronger than good." *Review of General Psychology*, 5, 323-370.

Bollen, J., Mao, H. & Zeng, X. (2010b) "Twitter mood predicts the stock market" *arXiv Server.* Retrieved 25 Oct, 2010 from http://arxiv.org/abs/1010.3003

Bollen, J., Pepe, A. & Mao, H. (2010a) "Modeling public mood and emotion: Twitter sentiment and socio-economic phenomena" *Proceedings from 19th International World Wide Web Conference* Raleigh, North Carolina.

Borodin, A., El-Yaniv, R. & Gogan, V. (2004) "Can we learn to beat the best stock" *Journal of Artificial Intelligence Research* 21:579-594.

Brock, L., Lakonishok, J., & LeBaron, B. (1992) "Simple technical trading rules and the stochastic properties of stock returns" *Journal of Finance,* 47, 1731{1764.

Cao, H., Coval, J. & D. Hirshleifer "Sidelined Investors, Trading-Generated News, and Security Returns," *Review of Financial Studies*, 2002, 15, 615-648.

Chan, W. (2003) "Stock price reaction to news and no-news: drift and reversal after headlines" *Journal of Financial Economics* 70, 223-260.

Chen, M., Gu, B. & Konana, P. (2009) "Social Capital, Social Identity and Homophily Behavior in Virtual Communities: An Analysis of User Interactions in Stock Message Boards" Revise and Resubmit *Information Systems Research.*

Chevalier, J. & Mayzlin, D. (2006) "The effect of word of mouth on sales: online book reviews" *Journal of Marketing Research* 43(3):345-354.

Claburn, T. (2009) "Twitter growth surges 131% in March" *Information Week*. Retrieved 25 Oct, 2010 from
http://www.informationweek.com/news/internet/social_network/showArticle.jhtml?articleID=216500968

Cohen, L. & Frazzini, A. (2008) "Economic Links and Predictable Returns" *Journal of Finance* 63:1977-2011.

Daniel, K., Hirshleifer, D. & Subrahmanyam, A. (1998) "Investor psychology and security market under- and overreactions" *Journal of Finance* 53:1839-85.

Das, S. & Chen, M. (2007) "Yahoo! For Amazon: Sentiment Extraction from Small Talk on the Web" *Management Science* 53(9):1375-1388.

Das, S., Martinez-Jerez, A. & Tufano, P. (2005) "eInformation: A Clinical Study of Investor Discussion and Sentiment" *Financial Management* 34(3):103-137.

De Long, J., Shleifer, A., Summers, L. & Waldmann, R. (1990) "Noise trader risk in financial markets" *Journal of Political Economy* 98:703-738.

Dellarocas, C. (2003) "The Digitization of Word of Mouth: Promise and Challenges of Online Feedback Mechanisms" *Management Science* 49(10):1407-1424.

Dhar, V. & Chang, E. (2009) "Does Chatter Matter? The Impact of User-Generated Content on Music Sales" *Journal of Interactive Marketing* 23(4):300-307.

Duan, W., Gu, B. & Whinston, A. (2008) "Do Online Reviews Matter? An Empirical Investigation of Panel Data" *Decision Support Systems* 45(4):1007-1016.

Fama, E. (1970) "Efficient capital markets: a review of theory and empirical work" *Journal of Finance.* 25:383-417.

Godes, D. & Mayzlin, D. (2004) "Using Online Conversation to Study Word-of-Mouth Communication" *Marketing Science* 23(4):545-560.

Godes, D. & Mayzlin, D. (2009) "Firm Created Word-of-Mouth Communication: Evidence from a Field Test" *Marketing Science* 28(4):721-739.

Granovetter, M. (2005) "The Impact of Social Structure on Economic Outcomes" *Journal of Economic Perspectives* 19(1):33-50.

Gu, B., Konana, P., Rajagopalan, B. & Chen, M. (2007) "Competition Among Virtual Communities and User Valuation: The Case of Investing-Related Communities" *Information Systems Research* 18(1):68-85.

Hirshleifer, D. (2001) "Investor Psychology and Asset Pricing" *Journal of Finance* 56(4):1533-1597.

Hong, H. & Stein, J. (1999) "A Unified Theory of Underreaction, Momentum Trading and Overreaction in Asset Markets" *Journal of Finance* 54:2143-2184.

Huberman, G. (2001) "Familiarity breeds investment" *Review of Financial Studies* 14(3):659-680.

Java, A., Song, X., Finin, T., & Tseng, B. (2007). "Why We Twitter: Understanding Microblogging Usage and Communities" *Joint 9th WEBKDD and 1st SNA-KDD Workshop*. San Jose, CA.

Kahneman, D. & Tversky, A. (1973) "On the Psychology of Prediction" *Psychological Review* 80:237-25.

Kim, S. & Hovy, E. (2006) "Extracting Opinions Expressed in Online News Media Text with Opinion Holders and Topics" *Proceedings of the Workshop on Sentiment and Subjectivity in Text* at the joint COLING-ACL conference. Sydney, Australia.

Koski, J., Rice, E. & Tarhouni, A. (2004) "Noise Trading and Volatility: Evidence from Day Trading and Message Boards" *Working Paper, University of Washington.*

Lakonishok, J., Shleifer, A. & Vishny, R. (1994) "Contrarian Investment, Extrapolation and Risk" *Journal of Finance* 49:1541-1578.

Lin, M., Prabhala, N.R. & Viswanathan, S. (2010) "Judging Borrowers by the Company they keep: Social Networks and Adverse Selection in Online Peer-to-peer Lending" *Working Paper* University of Maryland.

Luo, X. (2007) "Consumer Negative Voice and Firm-Idiosyncratic Stock Returns" *Journal of Marketing* 71:75-88.

Ma, Z., Sheng, O. & Pant, G. (2009) "Discovering company revenue relations from news: A network approach" *Decision Support Systems* 47:408-414.

Miller, C. (2010) "Twitter unveils plan to draw money from ads" *The NY Times* retrieved 1 Nov, 2010 from http://www.nytimes.com/2010/04/13/technology/internet/13twitter.html

Pang, B., Lee, L. & Vaithyanathan, S. (2002) "Thumbs up? Sentiment classification using machine learning techniques" *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing* Pennsylvania, PA.

Romero, D., Galuba, W., Asur, S. & Huberman, B. (2010) "Influence and Passivity in Social Media" *HP Labs Research.*

Rozin, P. & Royzman, B. (2001). "Negativity bias, negativity dominance, and contagion." *Personality and Social Psychology Review,* **5**, 296-320.

SAS (2011) "Using Hold-out Samples" Retrieved May 3, 2011 from http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/viewer.htm#etsug_tfbstmod_sect008.htm

Sabherwal, S., Sarkar, S. & Zhang, Y. (2008) "Online talk: does it matter?" *Managerial Finance* 34(6):423-436.

Seybert, N. & Bloomfield, R. (2009) "Contagion of Wishful Thinking in Markets" *Management Science* 55(5):738-751.

Schumaker, R. & Chen, H. (2009) "Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZFinText System" *ACM Transactions on Information Systems* 27(2).

Shmueli, G. & Koppius, O. (2011) "Predictive Analytics in Information Systems" *MISQ Forthcoming.*

TechCrunch (2010) "StockTwits Profile" Retrieved Sept 23, 2010 from http://www.crunchbase.com/company/stocktwits

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

Tetlock, P. (2008) "All the News That's Fit to Reprint: Do Investors React to Stale Information?" *AFA 2009 San Francisco Meetings Paper.*

Time (2010) "50 best websites 2010" *Time Magazine.* retrieved 1 Nov 2010 from http://www.time.com/time/specials/packages/completelist/0,29569,2012721,00.html

Tumarkin, R. & Whitelaw, R. (2001) "News or noise? Internet postings and stock prices" *Journal of Financial* Analyst 57(3) 41-51.

Tversky, A. & Kahneman, D. (1974) "Judgment under uncertainty: Heuristics and biases" *Science* 185:1124-1131.

Ullrich, C., Borau, K., Luo, H., Tan, X., Shen, L., Shen, R. (2008) "Why web 2.0 is good for learning and for research: principles and prototypes" *In Proceedings of the 17th International Conference on World Wide Web,* 705-714.

Wasko, M. & Faraj, S. (2005) "Why Should I Share? Examining Social Capital and Knowledge Contribution in Electronic Networks of Practice" *MIS Quarterly* 29(1):35-57.

Whitten, I. & Frank, E. (2005) "Data Mining: Practical Machine Learning Tools and Techniques" 2nd ed. *Morgan Kaufmann,* San Francisco.

Wikipedia (2010a) "Microblogging" Retrieved Sept 23, 2010 from http://en.wikipedia.org/wiki/Microblogging

Wikipedia (2010b) "Wishful thinking" Retrieved Nov 1, 2010 from http://en.wikipedia.org/wiki/Wishful_thinking

Wysocki, P. (1998) "Cheap Talk on the Web: The Determinants of Postings on Stock Message Boards" University of Michigan Business School Working Paper.

Xu, S. & Zhang, X. (2009) "How do social media shape the information environment in the financial market?" *ICIS 2009.*

Zeledon, M. (2009) "StockTwits may change how you trade" *Bloomberg Businessweek.* Retrieved Oct 25, 2010 from http://www.businessweek.com/technology/content/feb2009/tc20090210_875439.htm

Zhang, Y. (2009) "Determinants of Poster Reputation on Internet Stock Message Boards" *American Journal of Economics and Business Administration* 1(2):114-121.

Zhang, Y. & Swanson, P. (2010) "Are day traders bias free? – evidence from internet stock message boards" *Journal of Economics Finance* 34:96-112.