Identifying Information in Stock Message Boards and Its Implications for Stock Market Efficiency

Bin Gu, Prabhudev Konana, Alex Liu, Balaji Rajagopalan, and Joydeep Ghosh

Extended Abstract for Consideration for WISE 2006

Abstract

The information value of stock message boards has often been debated. A main difficulty in assessing the value is the presence of a large number of posts with varying quality. This paper presents an intuitive approach to identify and aggregate information in stock message boards. We weigh each post's recommendation by its author's credibility based on accuracy of his past posts. We find that the weighted average recommendation of a stock message board has prediction power over future excessive returns of the stock. The effect is both statistically and economically significant. Interestingly, a simple average recommendation of a stock message board has no prediction power for future stock movements. These results indicate that there exist informed investors in stock message boards, but their information is neither fully incorporated into the market price, nor fully acknowledged by peers in stock message boards. An implementable trading strategy is developed to explore the information value of stock message boards. We find that significant economic gain can be achieved even after taking into consideration of trading costs. We also discuss various approaches to improve the information aggregation process and thus the performance of the trading strategy.

1. Introduction

Stock message boards provide an excellent forum for investors to interact, debate, and exchange stock information. The level of activities on stock message boards is unprecedented. Despite the attention investors paid to stock message boards, their information value is often debated. But it is well known that information in these boards may have significant noise, contradictory recommendations, rumors, and manipulations. Many "pump and dump" schemes investigated by the SEC involve investors disseminating false information through message boards and selling stocks at artificially inflated prices (SEC 2006). As SEC ex-chairman Arthur Levitt puts it: " … investors [should] take what they see over chat rooms - not with a grain of salt - but with a rock of salt." (Carson and Felton 2003) In addition to the significant level of noise and manipulation, the sheer amount of information contained in these stock message boards can easily inundate any investors. Thus, stock message boards raise two interesting questions:

1) Can useful information from large scale stock message boards be easily identified?

2) Do stock markets incorporate such information efficiently?

This paper attempts to uncover underlying information in stock message boards. The difficulty of this task lies in identifying useful information among millions of posts, most of which could be noise or manipulations. As a result, the overall sentiment of stock message boards often carry little information content for future stock movements (Das and Chen 2001, Antweiler and Frank 2004a, Antweiler and Frank 2004b). We propose an alternative approach to identify and aggregate information from large-scale stock message boards. We weigh each post by its author's credibility, based on forecasting accuracy of his past posts. We use the weighted average recommendation of daily posts in a stock message board to represent its information content and consider whether the weighted average recommendation has prediction power over future stock returns in a CAPM model.

2. Dataset

Our dataset consists of message posts collected from Yahoo! Finance from April 2005 to April 2006. The collection was conducted using a software crawler written by one of the authors. Given the large number of stock message boards and the amount of time it takes to collect posts, we need to choose a representative set of stock message boards for our analysis. Antweiler and Frank (2004a) consider 45 stocks consisting of Dow Jones Industrial Average (DJIA) and Dow Jones Internet Commerce Index (XLK), representing a collection of well-known large-cap stocks and internet stocks. To facilitate comparison, we use the same set of 45 stocks in our analysis.

Each Yahoo!Finance message post consists of a unique user ID/screen name, a time and date of post, a subject, and the text of the actual message. Yahoo!Finance also allows posters to specifically label their sentiment about the stock. That is, posters are allowed to indicate in the sentiment field whether they feel the stock is a "strong buy", "buy", "hold", "sell", or "strong sell". This sentiment tag is optional and reflects the poster's recommendation of the stock. We explore this self-reported sentiment for our analysis and ignore posts without sentiment tags. One major advantage of this approach is that a large percentage of message board posts are not related to the stock itself (e.g., discussion of unrelated current events). Using self-reported sentiment allows us an easy way to screen out these irrelevant posts.

3. Aggregating Information from Stock Message Boards

A large number of posters create posts in stock message boards every day. Each poster has the option of including a recommendation via the sentiment tag of whether one should buy or sell the stock. Our objective is to identify and aggregate information from posts added within a given day and test whether the information predicts future stock movements. To facilitate formation of an implementable trading strategy, we define each day t for a stock message board to start from 3:45pm of the previous day and end at 3:45pm that day. The purpose of this definition is to allow investors time to analyze posts and make trade decisions before closing time. To aggregate information from stock message boards, we note the universe of posters as E where the ith poster is indicated by the notation e_i. The information content of a stock message board on day t can be calculated based on the recommendation of each poster on day t (denoted as $e_i(t)$) and the weight of each poster (denoted as $w_i(t)$). The weight $w_i(t)$ of a poster reflects how accurate the poster has been in the past. Higher values of $w_i(t)$ indicate that the ith poster has been accurate and credible than a poster with a lower value of $w_i(t)$.

Our weighted average approach works by repeating three basic steps: getting individual recommendation from each poster, aggregating the recommendations of all posters, and updating the weights of the posters based on their latest recommendations. We will now discuss each of these steps in more details.

3.1 Individual Recommendations

The first step in our weighted average approach is to obtain the individual recommendations of each poster for day t. That is, we need to determine $e_i(t)$ for all i. The recommendations of each poster are based on the sentiment tags created by the poster. If a poster makes a single recommendation, then $e_i(t)$ is 2, 1, 0, -1, -2 if the poster recommends "strong buy", "buy", "hold", "sell" and "strong sell" respectively. If a poster may create more than one post in a given day, we calculate the average sentiment of his posts in a given day as his recommendation.

If a poster does not have any posts on day t, the above discussion provides no indication of his recommendation. There are two alternatives in handling such a situation. We could either carry over his past recommendations or we can exclude the poster from information aggregation for day t. We choose not to carry over past recommendations because our focus is on identifying information contained in posts created in day t. In addition, as prices change constantly, we can not reasonably assume that a poster's recommendations would remain unchanged.

3.2 Aggregating Recommendations

We now aggregate the information across all posters to form an overall recommendation $y(t)$. The most intuitive definition of $y(t)$ is the weighted average of the recommendations $e_i(t)$ for all posters who create posts on day t as shown below in equation 1. The weights $w_i(t)$ for each $e_i(t)$ are given by the accuracy of the poster's historical recommendations. We will discuss the weight updating rule in section 3.3. Generally, the higher values of $w_i(t)$ correspond to posters that have historically been more correct. In particular, posters with a weight of zero have never been correct in the past.

$$y(t) = \sum_{i:M_i^{total}(t)>0} w_i(t) \times e_i(t) \Big/ \sum_{i:M_i^{total}(t)>0} w_i(t) \qquad (1)$$

For comparison, we also create an "unweighted" recommendation which reflects the overall sentiment of the discussion boards. This recommendation is similar to bullishness and other sentiment measures used in earlier

studies on message boards that give every post an equal weight[1]. We define the unweighted recommendation as an average of all poster recommendations on day t.

$$\bar{e}(t) = \sum_{i:M_i^{total}(t)>0} e_i(t) \Big/ \sum_{i:M_i^{total}(t)>0} 1 \qquad (2)$$

3.3 Updating weights

The aggregate recommendation $y(t)$ defined in (2) does not indicate a time horizon. We do not know whether the posters recommend holding the stocks for one day, one week, one month or longer. We start by using a one-day holding period to illustrate the weight update process. We then show that the weighted average approach can accommodate any holding periods. In a one-day holding period, one observes the realized log stock returns $r_i(t+1)$ the next day. As stock prices fluctuate with market prices, we use beta adjusted returns for our analysis, where beta is calculated using data during one year prior to the study period. S&P500 are used as the market index for beta calculation. Given the realized beta-adjusted log return $r_i^{adjusted}(t+1)$, the weights of all posters are updated. Weights on posters who accurately predicted the direction of beta-adjusted return are increased and weights on those with wrong recommendation are decreased.

We used a simple additive update rule to update weights of individual posters. Additive weight updates have been studied extensively in a variety of research context and can be traced at least as far back to Rosenblatt (1958). The additive update rule we use is as follows:

$$w_i(t+1) = \alpha \times w_i(t) + (1-\alpha) \times I\left(sign(e_i(t-1)) = sign\left(r_i^{adjusted}(t)\right)\right) \qquad (3)$$

where α is a parameter often known as the "learning rate" and $0 \leq \alpha \leq 1$. α controls and limits how quickly w_i changes from day t to $t+1$. To assure the robustness of our results, we use a wide range of α (0.1, 0.3, 0.5, 0.7 and 0.9) and find our result remain qualitatively unchanged. Equation (3) describes the weight updating rule when the holding period is one day. We need to revise the equation if the holding period is longer. For a holding period of h days, $e_i(t)$ is the prediction of the direction of beta-adjusted return for buying a stock at the closing of day t and selling it at closing of day $t+h$. Thus, the realized beta-adjusted return will not be known until day $t+h+1$. In order for our weight updates to use only known information, we use the following updating rule: $w_i(t+1) = \alpha \times w_i(t) + (1-\alpha) \times I\left(sign(e_i(t-h)) = sign\left(\sum_{k=1}^{h} r_i^{adjusted}(t-h+k)\right)\right) \qquad (4)$

4. Empirical Validation

To evaluate our framework empirically, we estimate predictability of $y(t)$ on future stock returns. We use the CAPM model to identify factors that predict stock returns on day $t+1$. The CAPM model suggests that excess stock returns (i.e. net of risk-free returns) equal to beta times excess market returns (i.e market returns minus risk-free returns). We therefore include excess return of market index S&P500 on day $t+1$ into our regression model. We also allow the coefficients on the excess market return to vary across stocks as different stocks have different beta values. The variable of interest is the aggregate recommendations we estimated from each stock message board the day before $y(t)$. If the market is efficient, the aggregation recommendations shall have no prediction power over future excess stock returns. On the other hand, the presence of significant prediction power would indicate that aggregation recommendations contain useful information on stock future movement. Besides including the most recent recommendations, we also include earlier recommendations from stock message boards to assess the long term effect of the recommendations. We include 5-day lags of earlier recommendations for that purpose. To simplify notation, we use L5[2] operator to represent 5-day lags of a

---

[1] One difference exists between the "unweighted" prediction and the sentiment measures used in earlier studies. In our definition, each poster is weighted equally, while in early studies, each post is weighted equally. The difference makes our approach less influenced by individuals who "pump" a stock by posting multiple times in a given day,.

[2] This operator has been used in earlier studies, e.g. Tetlock (2006)

variable, i.e. L5(x(t)) = [x(t-1) x(t-2) x(t-3) x(t-4) x(t-5)]. The equation for assessing predictability of information from stock message boards can be expressed as follows:

$$\left(r_{it} - r_{ft}\right) = \lambda \cdot L5\left(y_{it}\right) + \beta_i\left(r_{mt} - r_{ft}\right) + \varepsilon_{it} \tag{5}$$

In the above equation, r_{it} represents the return of stock i on day t. r_{it} is the risk-free rate on day t. L5(y_i) stands for the most recent five predictions on the direction of stock movement. r_{mt} is the market return on day t and its coefficients β_i capture the betas for individual stocks.

Equation (5) considers a holding period of one day. Accordingly, the weighting update gives more weight to posters who can accurately predict one-day returns. Many online posters however are not necessarily short-term traders. A casual browsing of online posts reveals that many of the posts focus on fundamentals of a stock and recommend long-term holding. This raises an interesting question – can information aggregated from the stock message boards predict excessive stock returns in the longer term?

Here our approach provides an advantage over earlier studies that often use the same aggregate information measure to test its relationship with stock returns over various forecasting horizons. Instead of using one aggregate measure, the weighted average approach tailors the information aggregation process for specific holding period. For a one-day holding period, the weight update process automatically identifies posters who are historically accurate in recommending stocks for one-day returns and aggregate only their information to form the overall recommendation. Likewise, for a one-month holding period, the weight update process only aggregates those who are credible in forecasting one-month returns. This approach allows us to dissect information for different stock holding periods from a group of posts without knowing ex ante their recommended holding periods.

We empirically test two longer periods for this study. We first test prediction power of stock message boards for stock returns for the next week, i.e. 5 trading days. We then test stock message boards for a one-month holding period, i.e. 20 trading days. For both tests, we first use equations (1), (2) and (4) to calculate individual poster weights and aggregate recommendations embedded in stock message boards. We define $y_{it}^{(5)}$ and $y_{it}^{(20)}$ as the aggregated recommendation for the next 5 trading days and the next 20 trading days respectively. We also define $r_{it}^{(5)}$ and $r_{it}^{(20)}$ as the excess stock returns for the same period. To assess whether information from stock message boards can be used to predict next week or next month returns, we run the following regressions:

$$\left(r_{it}^{(5)} - r_{ft}^{(5)}\right) = \lambda \cdot L5\left(y_{it}^{(5)}\right) + \beta_i\left(r_{mt}^{(5)} - r_{ft}^{(5)}\right) + \varepsilon_{it} \tag{6}$$

$$\left(r_{it}^{(20)} - r_{ft}^{(20)}\right) = \lambda \cdot L5\left(y_{it}^{(20)}\right) + \beta_i\left(r_{mt}^{(20)} - r_{ft}^{(20)}\right) + \varepsilon_{it} \tag{7}$$

Table 1 presents the results of this regression.

Table 1: Prediction Stock Returns Using Weighted Average Prediction From Stock Message Boards

Variables	One-Day Holding	One-Week Holding	One-Month Holding
Recommendation (t-1)	33.86**	85.81**	187.26**
Recommendation (t-2)	-8.50**	56.09 **	135.03**
Recommendation (t-3)	-6.79*	20.31**	92.68**
Recommendation (t-4)	-5.15	-11.72	54.91**
Recommendation (t-5)	-4.35	-45.35**	37.27**

** p-value < 0.01. * p-value < 0.05
Coefficients on market returns not reported.
Coefficients are reported in basis points (0.01%)

The result shows that information aggregated from stock message boards can predict future excessive stock return. Column 1 in the table shows a recommendation that changes from "Neutral" to "Buy" is associated with an increase of 34 basis points in the next day's excessive return. This result reflects the presence of informed investors in online message boards and their information can be discovered from millions of posts by considering posters' historical recommendation performance. Interestingly, the positive recommendation is

partially reversed in the next four days. The coefficients on recommendations from day *t-2* to *t-5* are all negative. This result suggests that the informed investors who are accurate in recommendation stocks for the next day largely build their recommendations on market short-term overreaction rather than on private information. Columns 2 and 3 show prediction power of recommendations over longer horizons. Information aggregated from tock message boards has significant prediction power on one-week stock returns and one-month stock returns. For the one-week holding period, an increase in recommendation from "Neutral" to "Buy" is associated with 0.88% increase in weekly excess returns. For the one-month holding period, an increase in recommendation from "Neutral" to "Buy" is associated with 1.87% increase in weekly excess returns. These results indicate that our weighted average approach can identify informed investors in stock message boards.

We now compare the effectiveness of the weighted average approach against a simple average approach in aggregating information from stock message boards. Using equation (3), we calculate average recommendation across all posters for a given stock on day *t* without considering posters' historical recommendation performance. We apply the same approach as in equation (6) to assess its predictability on future stock returns:

$$\left(r_{it} - r_{ft}\right) = \lambda \cdot L5\left(\bar{e}_{it}\right) + \beta_i\left(r_{mt} - r_{ft}\right) + \varepsilon_{it} \tag{8}$$

Table 2: Prediction Stock Returns Using Simple Average Prediction From Stock Message Boards

Variables	Coefficients	t-statistics	p-value
Recommendation (t-1)	0.71	3.61	0.84
Recommendation (t-2)	4.74	3.67	0.20
Recommendation (t-3)	-3.72	3.67	0.31
Recommendation (t-4)	-2.58	3.66	0.48
Recommendation (t-5)	-1.73	3.60	0.63

** p-value < 0.01. * p-value < 0.05
Coefficients on lagged stock returns, market returns and fixed effect not reported.
Coefficients are reported in basis points (0.01%)

Table 2 presents the results of the simple average recommendations. It shows that recommendations formed by simply aggregating individual recommendation on stock message boards have no prediction power on future excessive stock returns, a result consistent with findings in early studies (e.g. Antweiler and Frank 2004) that bullishness in stock message boards has no impact on stock returns. Our results indicate that most of online investors who posts on stock message boards are not informed and their recommendations have little informational value.

5. Conclusion

In this paper, we develop a methodology to identify and aggregation useful information from financial message boards. Our approach uses each poster's historical performance as weight to aggregate information from stock message boards. We find the aggregate recommendations have significant prediction power over future stock returns. Our finding has implications for market efficiency. The efficient market hypothesis indicates that public information shall be fully incorporated into market prices and has no prediction power over future stock movements. We show evidence that information available in stock message boards has not been incorporated into stock prices even though identification of the information does not require sophisticated financial or statistical knowledge. Our results suggest that informed investors are presence and active in stock message boards, but their information are not fully reflected in market prices.

Reference

Antweiler W. and M. Z. Frank. 2004. Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards. Journal of Finance 59 (3), 1259-1294.

Bagnoli, Mark, Messod D. Beneish, and Susan G. Watts, Whisper forecasts of quarterly earnings per share, Journal of Accounting and Economics 28, 27–50.

Tumakin R. and R. F. Whitelaw 2001. New or Noise? Internet Postings and Stock Prices. Financial Analyst Journal, 57, 41-51