The Effects of Message Board Information Dissemination on Stock Activity

By

Praveen Ghanta

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degrees of

Bachelor of Science in Electrical [Computer] Science and Engineering

and Master of Engineering in Electrical Engineering and Computer Science

at the Massachusetts Institute of Technology

May 16, 1999

June 1999]

Author_____
Department of Electrical Engineering and Computer Science
May 17, 1998

Certified by_____
Michael Siegel
Thesis Supervisor

Accepted by_____
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

The Effects of Message Board Information
Dissemination on Stock Activity
by
Praveen Ghanta

Submitted to the
Department of Electrical Engineering and Computer Science

May 16, 1999

In Partial Fulfillment of the Requirements for the Degree of
Bachelor of Science in Computer Science and Engineering
and Master of Engineering in Electrical Engineering and Computer Science

ABSTRACT

Internet stock message boards provide individual investors with a new information medium that could have several effects on stock activity. Message boards may provide investors with a new information source on small cap stocks. Message boards may affect the trading and volatility of certain stocks by influencing investors. The impact of message board information may vary according to characteristics of the message board. This thesis determines whether message board activity has statistical relationships with stock activity. The research examines how these relationships vary based on message board and stock characteristics. Regression analysis of message board and stock data was performed. The analysis shows that message boards track stock price and trading volume. The quality of the tracking relationship varies according to certain message board characteristics. Message boards have no systematic effects on trading activity, but have increased short term volatility for some small cap stocks.

Thesis Supervisor: Michael Siegel
Title: Senior Research Scientist, MIT Sloan School of Management

Contents

1. Introduction **6**

1.1 Thesis Organization..... 7

1.2 Financial Information on the Internet..... 7

1.3 Introduction to Stock Message Boards..... 8

1.4 Message Boards and Stock Activity..... 10

1.4.1 Types of Users..... 10

1.4.2 Potential Effects..... 10

1.4.3 Aims of Quantitative Analysis..... 11

2. Research Design **12**

2.1 Design Variables..... 13

2.1.1 Stocks..... 13

2.1.2 Message Boards..... 14

2.1.3 Time Frame and Data Availability..... 14

2.2 Quantifying Messages..... 15

2.3 Quantitative Analysis..... 16

2.3.1 Price and Price Change Regressions..... 17

2.3.2 Trading Volume Regressions..... 19

2.3.3 Stock Volatility Regressions..... 19

2.3.4 Statistical Testing..... 20

3. Stock – Message Board Relationship **21**

3.1 Data..... 21

3.2 Stock Price – Message Relationship..... 23

3.2.1 Predictive or Reactive Relation? 25

3.2.2 Hourly Regressions..... 27

3.3 Regression Results by Industry..... 28

3.3.1 Software/Internet Companies..... 28

3.3.2 Biotech/Pharmaceutical Companies..... 29

3.3.3 Hi-Tech/Hardware Companies..... 30

3.3.4 Large Capitalization Stocks..... 32

3.3.5 Comparative Analysis..... 33

3.4 Trading Volume – Message Volume Relationship..... 35

4. Trading Effects of Message Boards **37**

4.1 An Example Scenario..... 37

4.2 Observed Effects..... 39

4.2.1 Effects on Information Dissemination..... 39

4.2.2 Hyping and Touting on Message Boards..... 40

4.2.3 Corporate Reaction..... 42

4.3 Quantitative Analysis..... 43

4.3.1 Volatility Regression Analysis..... 43

4.3.2 Specific Incident Hourly Regressions..... 47

5. Characteristics of Effective Message Boards **49**

5.1 Defining "Effective" Message Boards..... 50

5.1.1 Message Boards Following Large Cap Stocks..... 50

5.1.2 Message Boards Following Small Stocks..... 50

5.2 Identifying Effective Message Boards..... 51

5.3 Characterization of Observed Message Boards..... 53

5.3.1 Statistical Data..... 53

5.3.2 Observational Data..... 54

5.3.3 Analysis..... 58

6. Conclusions and Future Research **60**

6.1 Regression Analysis..... 60

6.2 Message Board Characteristics..... 62

6.3 Message Board Design and Business Model..... 62

6.4 Conclusion..... 63

7. Appendices **64**

7.1 Appendix 1..... 64

7.2 Appendix 2..... 66

7.3 Appendix 3..... 69

7.4 Appendix 4..... 71

7.5 Appendix 5..... 72

8. References **74**

1 Introduction

The recent rise of internet trading and financial sites on the web has given the individual investor considerably more sources of information. Much of the information on the web simply mirrors that available through traditional media sources. More recently, message boards that allow individuals to discuss stocks have gained in popularity. Sites like Yahoo, Silicon Investor, and AOL feature stock message boards. Message board traffic represents a new information source for individual investors, and may influence their investment decisions. With certain stocks, held primarily by individuals, this influence may be reflected in the stocks' performance and volatility. Articles in ZDNet, The Industry Standard, and CNet provide some evidence that message board traffic can have real effects.[1] In several instances companies charged individuals with libel and defamation after negative comments purportedly caused a drop in stock price. These and other cases stimulate the need for a systematic analysis of the effects of message board traffic on equities trading.

This thesis examines the relationship between message board activity and stock activity. The relationship between message boards and stock activity could potentially have several effects. Message board messages may have a predictive relationship with stock activity in cases where the boards provide new information. Message boards may affect the trading and volatility of certain stocks by influencing investors. This thesis examines stock and message board data in order to determine what these effects are, and how they vary across different message boards.

1.1 Thesis Organization

Chapter 2 provides a detailed description of the research design used to conduct analyses of the relationships between stocks and message boards. Chapter 3 describes the results of the stock price – message board relationship, and analyzes these results for differences across industries, company size, and other factors. Chapter 4 then discusses the effects of message boards on stock volatility, and examines whether certain cases of stock price spikes were caused by message boards. Chapter 5 analyzes which characteristics determine the effectiveness of message boards as sources of information dissemination. Conclusions and suggestions for further research are made in Chapter 6.

1.2 Financial Information on the Internet

With the advent of the Internet, society has been provided a network of information unparalleled in breadth and availability. At the end of 1997, there were 70 million regular users of the Internet, and the number continues to grow at the incredible rate of 71,000 per day [9]. One of the most important roles the internet plays is as a means of information dissemination for electronic commerce. Yahoo Finance, Quicken.com, and other sites have risen to provide a wealth of financial information. These sites contain information on equities investing, including stock quotes, charts, SEC filings, and other data.

Electronic message boards have recently experienced rapid growth on the Internet as well. Silicon Investor, a major stock message board web site, has 80 million web page

[1] See [10], [12], and [16].

hits per month.[2] Stock message boards differentiate themselves from other sources of financial information in a number of ways. Message boards contain interactive and opinion related content, while sites like Yahoo Finance provide factual and historical data. Unlike chat rooms, message boards offer users permanent access to information exchanged in the past. Message boards also differ from premium investment sites like theStreet.com, which charge a fee for detailed analyst opinions and recommendations. Message boards offer up much of the same analysis, albeit performed by anonymous individuals ranging from new investors to corporate insiders. While some message board posts are factually based, users also find posts filled with rumors, personal opinions, predictions, and at times slander and libel.

1.3 Introduction to Stock Message Boards

According to Wendy Lee, author of "Competing in the Net economy environment," information today is expected to be available in a Web-based format and easily accessible via the Internet. Message boards allow internet users to access information that would otherwise have been inaccessible or difficult to locate. For example, penny stocks that do not receive substantial print coverage now have several message boards dedicated solely to them. According to Herb Greenberg, senior columnist at theStreet.com, this is exactly what message boards are intended to do:

> "Message boards play off the interactive nature of the interactive nature of the Internet, and, in theory, at least, they're a good idea: They're supposed to level the playing field by giving the little guys a place to congregate. Rather than rely on brokers, analysts, and the news media, investors can

[2] Figures courtesy Silicon Investor, http://www.siliconinvestor.com

> compare notes and, if all goes well, beat Wall Street at its own game. Some posts can be well written and informative. On small stocks, it's not uncommon for them to break news before Dow Jones, Bloomberg, or Reuters. And if you're lucky, corporate insiders... will post confidential inside information." [8]

There are several major web sites featuring well used message boards, each offering a different set of features. The Motley Fool (http://www.fool.com) is among the better known financial resource sites. Motley Fool boards tend to have very little jargon and contain information ranking high on a scale of "pleasantness and practicality." [15] Its message boards, which cater to the general user, tend to provide utilitarian information pertinent to daily financial issues. The most popular message board site is Yahoo Messages (http://finance.yahoo.com). Yahoo provides additional services such as chat rooms and background data on corporations. Yahoo's message boards carry a lot of traffic and hence a lot of excess information.

Silicon Investor (http://www.siliconinvestor.com) charges an annual fee of $120 to its users. Its boards focus on computer and biotech companies and discussions tend to be more technically advanced than other message boards. In addition, members also have access to a company profile database. America Online (Keyword: Personal Finance) provides a service exclusive to its members. With a variety of live formal and informal chat rooms, message boards, investing forums, and financial advice, AOL provides an incredible spectrum of information.

1.4 Message Boards and Stock Activity

1.4.1 Types of Users

The broad appeal of stock message boards has caused many types of individuals to become regular users. Many are casual investors seeking advice or the personal opinion of fellow investors. Day traders and experienced shareholders also frequent message boards. Message boards occasionally draw corporate employees who provide insights into company performance and work to dispel rumors. Some individuals frequent message boards for the express purpose of hyping their stock positions [18]. These individuals and the associated storms of hype on boards have allegedly caused a number of volatile swings in small company stocks. An interesting example of these individuals is Joe "Smokin'" Park. According to Salowe [5], Park made a reputation through day trading, and now has a following thousands of investors on his Silicon Investor message board. Park has more influence than simply handing out free advice. On December 11, 1998, an analyst at Standard and Poor's observed that shares of FileNET Corporation had jumped $8 a share. FileNET volume also rose to about 3.5 million, double the average volume. The analyst discovered that Park had posted messages on Silicon Investor his belief that the stock price was going to rise.

1.4.2 Potential Effects

Traditional market theory holds that the market prices already take into account all publicly available information. As information has become more available to individual investors, this might seemingly cause market volatility to drop, as all investors

receive better valuation information. If much of the new information generated is of questionable value, as on message boards, this noise might simply have no effect on the market. With stocks that do not have a significant institutional following, however, this conventional logic may prove less true. News articles at CNet [10], Wired [13], and ZDNet [12] seem to indicate that both news and defamy released via message boards can have a significant impact. With thinly traded stocks that have little news, analyst, or institutional following, message boards represent a new information source for investors. Whether this causes the markets to value these equities more efficiently remains to be tested.

1.4.3 Aims of Quantitative Analysis

The stock – message board relationship could be very complex. Do message boards consistently provide new information for any class of stocks? Are there differences between message boards which affect the quality of information provided? Can a statistical relationship between stock price activity and messages be determined, and how does this relationship differ across types of stocks? Finally, do message boards have any systematic effects on stock volatility or trading patterns? Quantitative analysis of stock price data and message content over time will help to answer a number of these questions. The analysis will also generate further topics of exploration, such as creating effective designs for message boards based on the results.

2 Research Design

There are many potential avenues of research concerning the relation between stock message board information and equities trading. The research must be further defined through the specification of several end goals:

1. Determine whether a statistical relationship exists between message "predictions" and stock price changes
2. Determine whether a statistical relationship exists between message board traffic levels and trading volume or stock volatility
3. Determine for which stocks the above relationships are most likely to hold
4. Characterize message boards and identify the differences in and relationships between types of message boards
5. Examine differences in message board - stock relationships based on the characterization of the message board

The above list leaves the actual methods of data collection and analysis to be determined. Numerous issues arose related to variable choice, data collection, message characterization, and actual regression analysis. Stocks of smaller, less followed companies were selected, in accordance with the hypothesis that message boards would have the strongest impact here. The regressions and statistical analysis performed tested the simple message prediction – price change relation as well as other more complex relations. These design issues are explored in detail below. Section 2.1 discusses the different design variables and their selection, while section 2.2 discusses the system used to quantify messages. Section 2.3 then details the methods used in the regression analysis of the data.

2.1 Design Variables

2.1.1 Stocks

The choice of which stocks to analyze has a major impact on the direction of the research. Since blue chip and S&P 500 companies have a considerable volume of traditional news coverage, they are unlikely to be moved by message board traffic. Considerable media attention suggests that small or micro cap stocks may show the strongest results [2]. These stocks make better candidates since they have few or no analysts and little substantial daily news. Stocks with low institutional ownership are more likely to move as a result of individual investor activity as well. Thinly traded stocks make better choices, since a small number of investors could potentially move the stock price [5]. While these stocks were the focus of the research, a sampling of companies including those in the S&P 500 were included for comparative purposes. Companies were also chosen across industries to help determine whether effects are concentrated in one industry. Specific criteria were determined for the selection of small cap stocks:

1. Market capitalization less than $500,000,000
2. Average trading volume less than 100,000
3. Less than 5% institutional ownership
4. Two or fewer analysts

For all stocks, message board activity was an important selection factor. Stocks with inactive message boards obviously could not be used. Since messages required manual grading, stocks with overly active message boards (e.g. EBAY or AMZN, with > 100 messages per day) could not be used either.

2.1.2 Message Boards

Message board selection could potentially have an impact on the research, since different message boards may attract different numbers and kinds of users. Message board readership was an important factor in determining a board's value. Yahoo, Silicon Investor, and The Motley Fool have emerged as the top three message board destinations, with Silicon Investor recording 80 million web page hits and 2 million user logins monthly.[3] Of these, Yahoo and Silicon Investor have significantly larger discussions for many small and micro cap companies. These two message boards were used for all message data. While this limited comparison across different boards, it allowed for better comparison of results across stocks. For smaller companies, generally only one message board was active enough to merit consideration, so that message board selection was not really an issue.

2.1.3 Time Frame and Data Availability

The time frame for message and stock data was largely determined by the message boards. Since message boards were created recently (particularly for some stocks), the time frame used must be recent. Since messages must be manually graded the volume of messages on some boards necessitated a small time span. Three months proved to be an effective compromise, since it included at least one earnings report for a company. Stock price data was available at the daily and trade-by-trade level. Daily data

[3] Courtesy Silicon Investor

was available from many sources, and was obtained from AOL. Trade-by-trade data was obtained from the NYSE TAQ data set, and was used to construct hourly data sets.[4]

2.2 Quantifying Messages

Message board messages vary immensely in their detail, quality, and even the nature of information they provide, making it difficult to develop a uniform grading system. Even when messages display strong opinions it can be difficult to compare them directly. The following two message quotes from Nov. 6, 1998 express positive opinions on the Silicon Investor Xybernaut board:

> "some times you have to wait for the good ones-- good luck to us all regards john"
> - JHP, Message 1572
>
> "Nice rise on XYBR today. It stayed up all day, so we should expect more on Monday."
> - CanynGirl, Message 1576

One message is positive for the long run, while the other predicts gains on the next trading day. The simplest system for quantifying messages grades each message as positive, negative, or neutral towards the stock. Beyond this, each message contains a certain level of factual information, and possesses a degree of opinion based judgement. The grading system used assigns each message a -1, 0, or 1 based on its overall prediction for the stock. Each message is then assigned fact and opinion values on scales from –3 to 3. Negative values correspond to facts or opinions which reflect negatively on the stock. This system allows each of the grading variables to vary independently.

[4] Information on the TAQ data can be found at http://risk.mit.edu:8080/web/

The grading system used does not take into account much of the more specific information in each message. Each of the two messages quoted above were graded 1 for the prediction, 0 for the fact level, and 1 for the opinion level (Appendix 5 has more grading examples). The first message from JHP makes a long term prediction about stock value, while the second message from CanynGirl makes a short term prediction. Since the time frame of message predictions is difficult to measure (is JHP talking about next month or next year?), no method for taking this into account was developed. More complex grading systems might be devised, but these systems introduce noise since message data is imprecise. Human graders introduce noise through their interpretations, and a more complex system would exacerbate this.

2.3 Quantitative Analysis

The simplest level of quantitative analysis involved simply plotting stock prices and trading volumes against message predictions and message volumes respectively. This allowed for a rough inspection to determine whether interesting results might follow from regression analysis. Several sets of regression analysis were performed. Stock price and price changes were regressed against a combination of message volume, message prediction, and individual message predictors like fact and opinion level. To help control for the movements of the market, the movement of market indices were included as regressors as well. Time delay regressions were used to determine whether messages had any predictive value over stock prices, or whether price movements caused messages to be written. Trading volumes were regressed against message volume to determine the relation between the two variables. Monthly stock volatility was regressed against

monthly message traffic rates and index volatility to determine whether message traffic moved with volatility.

2.3.1 Price and Price Change Regressions

To investigate the key stock price – message board relationship, stock price and price change were regressed against the message predictive variables. A basic linear regression form was used. The following equation shows the general linear form for the regressions performed. The actual regressions involved different independent variables in different regressions.

$$\Delta price_i = B_1 + B_2 * message\_volume_i + B_3 * opinion_i + B_4 * fact_i + B_5 * \Delta index\_price_i$$

A linear relation is not unreasonable, as some linear proportion of message board readers might act on its advice on average. For both daily and hourly regressions, one observation had to be created from all messages within the time unit. This was done by simple summation of each message's prediction, fact, and opinion values respectively. Since each message's values were weighted equally, it was possible to have a group of messages with varying opinions cancel each other out in terms of the total values for the time period. Also, since trading occurs only during trading hours, many time units had message traffic without any corresponding trading. Since price change and trading volume were 0 for these time periods, this added some noise to the daily regressions, and considerable noise to the hourly regressions (where less than 1/3 of all time units were during trading hours).

In addition to the message values, the Nasdaq or S&P 500 index was used in each regression to take into account the effect of market conditions. The Nasdaq was used as a regressor for technology stocks, while the S&P 500 was used for large cap industrial stocks. Time lag regressions were performed in order to determine whether messages were predictive of price changes, or whether they simply reacted to stock movements. The three previous time units' message variables were used for predictive regressions, and the three subsequent time units' variables were used for reactive regressions. A table of the important regressions follows. Appendix 1 contains the complete details of the regressions performed.

Reg. #	Regressand	Regressor 1	Regressor 2	Regressor 3	Regressor 4	Regressor 5	Regressor 6
1	price	mess. volume	prediction	index			
2	price	Y mess. volume	Y prediction	index			
3	price change	mess. volume	prediction	index change			
4	price change	mess. volume	fact	opinion	index change		
5	price change	Y mess. volume	Y2 mess. volume	Y fact	Y2 fact	Y opinion	Y2 opinion
6	price change	Y mess. volume	Y2 mess. volume	Y prediction	Y2 prediction	index change	

Table 2-1 Price and Price Change Regressions

Table 2-1 shows the six most important price and price change regressions. The Regressand column contains the dependent variable for each regression model, and columns Regressor 1 – 6 contain the independent variables for each regression model. The "Y" and "Y2" in some of the Regressor columns indicate time lagged variables used in predictive regressions. "Y" indicates the previous day's value, while "Y2" indicates the value from two days prior. Regression 5, for example, regresses price change against the past two days' message volume, fact, and opinion in order to determine whether these variables have predictive value.

2.3.2 Trading Volume Regressions

Trading volume was regressed against message volume in order to determine whether a relationship existed between the variables. Other message variables were used as regressors here as well. Index volume was to be used in these regressions, but was not available in the data, unfortunately. Time lag regressions were performed to determine whether message volume predicted or reacted to trading volume movements. Table 2-2 lists the primary trading volume regressions performed. The table is patterned after Table 2-1 (explained above). "T" indicates the next day's value, while "T2" indicates the value two days ahead.

Reg. #	Regressand	Regressor 1	Regressor 2	Regressor 3
1	volume	mess. volume	index change	prediction
2	volume	mess. volume		
3	volume	Y mess. volume	Y2 mess. volume	
4	volume	T mess. volume	T2 mess. volume	

Table 2-2 Trading Volume Regressions

2.3.3 Stock Volatility Regressions

The stock volatility analysis attempted to determine a relationship between monthly stock volatility and the monthly message rate. Monthly message rate (instead of message volume) was used in order to normalize for the difference in length between months. In general, message rates have been rising over the past two years as message boards have risen in popularity. If stock volatility has trended along similar patterns in certain stocks (once market volatility has been accounted for), there is evidence to suggest that message boards may be implicit in the change. Monthly stock volatility was

regressed against monthly message rate and monthly index volatility. Message rate was also regressed against a time trend to determine whether it trended upward over time.

2.3.4 Statistical Testing

The above statistical analyses were prone to a number of problems that might adversely affect the accuracy of results. Multicollinearity among independent variables seemed to be the biggest potential problem for a number of the regressions. Multicollinearity might arise in these regressions since a number of the message predictive variables were in some cases highly correlated. Message prediction and opinion tend in general to be highly correlated, as message opinion usually differs only in magnitude from the prediction. No single test for multicollinearity has gained widespread acceptance [14], so observation of independent variable cross-correlations and t values in multivariate regressions was used. Message prediction was found to cause multicollinearity problems when used with either fact or opinion, and message fact and opinion occasionally caused multicollinearity when used together.

Other potential problems included autocorrelation in time lag regressions, and heteroscedasticity of error in those regressions where stocks experienced large price spikes. The Durbin-Watson test for autocorrelation was used to test time lagged regressions [4]. The null hypothesis of no autocorrelation could not be rejected. Heteroscedasticity did not appear to be an issue, as in time series regressions the changes in independent variables and the dependent variable generally occur at the same order of magnitude [14]. Applications of the Goldfeld-Quandt test confirmed this hypothesis [6].

3 Stock – Message Board Relationship

Twelve stocks from a range of industries were examined to determine the relation between stock prices and message board activity. The time period observed for all stocks was October 1, 1998 to December 31, 1998. In all but one case, message boards were found to accurately track stock activity. No evidence of a predictive relation was found for any company, however. Contrary to initial hypotheses, regression relations found were very similar across all types of companies. No substantial difference was found between large companies and small companies, or across industries. Even stronger than the price – message relationship was the relationship between trading volume and message volume. Message board traffic volumes were strongly linked to stock trading volume in almost every case. Regression results also provided some insights into the accuracy of message grading schemes used. The simple message prediction grading proved as accurate as the more complex system involving fact and opinion values.

Section 3.1 details the specific stock and message board data used in the regression analysis. Section 3.2 discusses the general stock price – message relationship, while Section 3.3 describes more detailed results by industry. Finally, Section 3.4 examines the trading volume – message volume relationship.

3.1 Data

The twelve stocks examined included companies from a broad range of industries. A number of companies were selected to investigate the hypothesis that message boards would have stronger relationships with small company stocks (see 2.11). These

companies were selected from the biotech and pharmaceutical, software and internet, and electronics and computer hardware industries. Several large capitalization stocks were also selected. These companies were drawn from the consumer products, retail, oil, and semiconductor manufacturing industries. Table 3-1 provides a description of the companies examined. The companies were selected from four industry categories: internet/software, biotech/pharmaceutical, computer hardware, and large capitalization. While the large cap stocks vary in industry, they are distinguished by the fact that they are all large, well established corporations.

Ticker	Company Name	Industry	Description
GLCCF (now GLXW)	Galaxiworld Casinos Ltd.	Internet/Software	Operates a digital casino on the Internet (http://www.galaxiworld.com)
INNI	I/Net Inc	Internet/Software	Develops Internet computer systems and software; principally AS/400 software
VINF	VISTA Info Solutions	Internet/Software	Provides environmental risk information about real estate in the U.S. to bankers, corporations
BASEA	Base Ten Systems	Biotech/ Pharmaceutical*	Creates software solutions for pharmaceutical and medical device manufacturing industries
HEB	Hemispherx Biopharma	Biotech/ Pharmaceutical	Uses nucleic acid technologies to develop treatments of viral infections and some cancers
TCLN	Techniclone Corp	Biotech/ Pharmaceutical	Makes monoclonal antibodies designed to attack cancer cells
SRSL	SRS Labs	Computer Hardware	Develops audio technologies for use in home, car, computer, and professional sound systems
XYBR	Xybernaut Corp	Computer Hardware	Creates wearable computer systems to enhance productivity in commercial applications
CPU	CompUSA Inc	Large Cap Retail	Sells personal computer hardware and software and related products through 207 superstores
EK	Eastman Kodak	Large Cap Consumer	Develops and sells consumer and commercial photographic imaging products
NVLS	Novellus Systems	Large Cap Semiconductor	Creates chemical vapor deposition equipment used in the fabrication of integrated circuits
TX	Texaco	Large Cap Oil	Produces crude oil, natural gas and petroleum products
* BASEA software is tied closely to pharmaceutical manufacturing processes, hence its designation			

Table 3-1 Company Descriptions[5]

[5] Company profiles from http://quicken.excite.com/investments/quotes/

The time period for observation for all message boards was October 1, 1998 to December 31, 1998. Seven Silicon Investor and five Yahoo! Message boards were used. The stock data and message data did not align perfectly since trading only occurs during market hours. The data were augmented with data points for weekends and nontrading hours (for hourly regressions) so that all messages could be included.

3.2 Stock Price – Message Relationship

The daily regressions on stock price and message data revealed some strong relationships. For all but one message board, at least one of the message predictive variables (prediction, fact, opinion) had a statistically significant positive relationship with daily stock price change. Table 3-2 provides some details on the quality of the price change – message relationship for all twelve stocks.

Stock Name	R^2 for Price Change Regression	Prediction Significant?	Fact Sig?.	Opinion Sig.?	R^2 for Volume Regression	Message Vol. Sig.?
INNI	0.5193	Yes	Yes	Yes	0.9125	Yes
GLCCF	0.5108	Yes	Yes	Yes	0.4592	Yes
EK	0.3358	Yes	No	Yes	0.5458	Yes
NVLS	0.3273	Yes	Yes	Yes	0.2584	Yes
XYBR	0.2764	Yes	Yes	Yes	0.5308	Yes
HEB	0.2474	Yes	No	Yes	0.4335	Yes
VINF	0.2358	Yes	No	Yes	0.1476	Yes
CPU	0.2011	No	Yes	No	0.2974	Yes
SRSL	0.18	Yes	No	Yes	0.2574	Yes
BASEA	0.1729	Yes	Yes	Yes	0.1599	Yes
TX	0.1438	Yes	No	No	0.1996	Yes
TCLN	0.0285	No	No	No	0.0238	No

Table 3-2 Price Change Regression Results

The R^2 for price change regression column shows the R^2 for either regression 3 or 4 listed in Table 2-1, depending on which was more significant. The different predictive variables

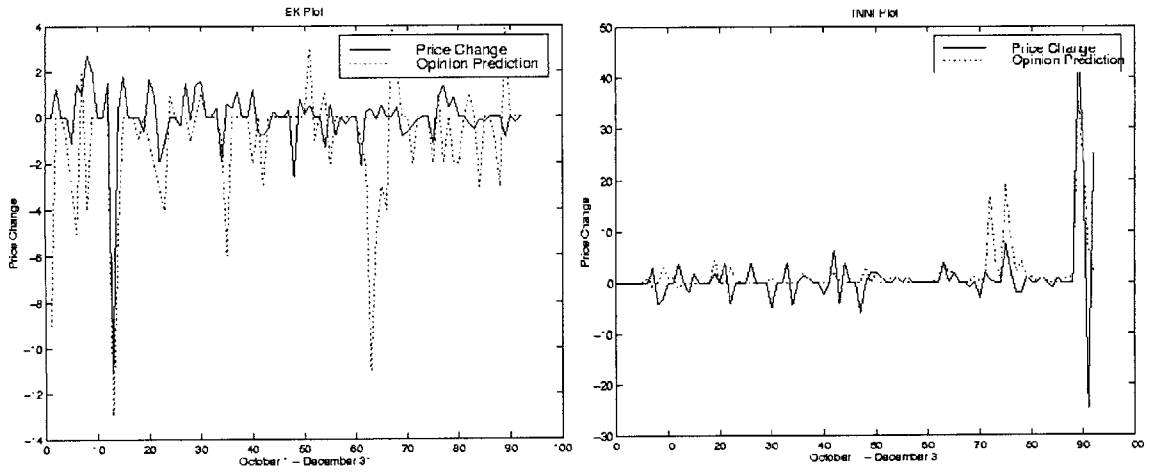
varied in their accuracy on different boards, so the better of the two regressions was taken. The R^2 value for a regression measures the quality of the regression fit. An R^2 of 1 indicates that the independent variables explain all variation in the dependent variable, while and R^2 of 0 indicates that the independent variables have no relationship to the dependent variable. The next three columns show whether the prediction, fact, and opinion variables were statistically significant in the price change regression. This means that the coefficients of these variables had a 95% chance of being statistically nonzero, as measured using the t statistic. [14] The R^2 for volume regression column shows the R^2 for regression 1 shown in Table 2-2. The last column of Table 3-2 shows whether the message volume was statistically significant in the volume regression.

The majority of message boards showed a statistically significant relationship between both prediction and opinion and the daily price change. For all but two message boards message prediction was significant. Message opinion was significant for all but one of the same message boards. This indicates that the message prediction alone provided as good a measure of message board activity as message opinion and fact. The simple grading system was as effective as the more complex system. This is not surprising, since the more complex system introduced more potential for human error.

In general, the message boards with a higher quality relationship (measured by the R^2 of price change regressions) tended to also show a significant positive relation between factual content and price change. Message boards that tracked price change effectively also tracked message volume more effectively. Certain message boards tracked stock activity considerably better than others overall. Section 3.35 and Chapter 5 attempt to explain why these differences in message board effectiveness exist.

Only one stock message board showed no significant results at all. The TCLN message board showed both an extremely low quality of fit for all regressions and no statistically significant coefficients. This message board seemed to be an exception, however. The great majority of the boards showed that there is a significant tracking relationship between message board content and stock price activity. Table 3-2 also shows that message prediction and opinion provided consistent results, and that message fact was generally significant for message boards with a stronger stock price - message relationship.

3.2.1 Predictive or Reactive Relation?



Graphs 3-3 and 3-4: EK and INNI Price Change – Opinion Prediction Plots

The above graphs show two of the more accurate message boards, EK and INNI. Opinion prediction is measured as an integer in the graphs. Graphs 3-3 and 3-4 indicate that there may be a reactive relationship between opinion and price change, since opinion swings occur in general concurrently with or after major price movements. These results are confirmed by graphs of the remaining stocks, and by the results of the predictive and

reactive regressions (see Graphs 3-3 – 3-12 and Appendix 2). The predictive regressions regressed the present day's stock data against previous days' message data. For a number of message boards significant results were found for predictive regressions of today's price change against past message predictive variables. This indicates that the message boards are predicting the price changes. In every case, however, reactive regressions show even stronger relationships between today's price change and future message predictive variables. Since price movements in stocks often last for several days, an early price movement could cause positive message predictions. These positive message predictions would then correlate with later parts of the same price movement, causing a "predictive" relation. The reactive relation would always be stronger, since this measures the true relationship.

A specific example using INNI will help clarify this effect. Graph 3-3 shows opinion prediction vs. price change for INNI. The main movement in INNI's price occurs in a prolonged surge at the end of December. Concurrently the opinion prediction of the message board rockets upward. If the one day predictive relation were measured, this would be equivalent to shifting the opinion prediction graph right by one day and determining the relationship. There would still be considerable overlap between the two graphs during the end of December period, and so a significant predictive relationship would be determined. The regression statistics show this clearly. The predictive regression of price change against the previous two days' message volume, fact, and opinion yields a significant relationship with an R^2 of 0.3369. The reactive regression of price against the next two days' message volume, fact, and opinion yields a much stronger relationship with an R^2 of 0.7954. This sort of result is consistently borne out

with every message board, leading to the conclusion that message boards follow price changes rather than predicting them.

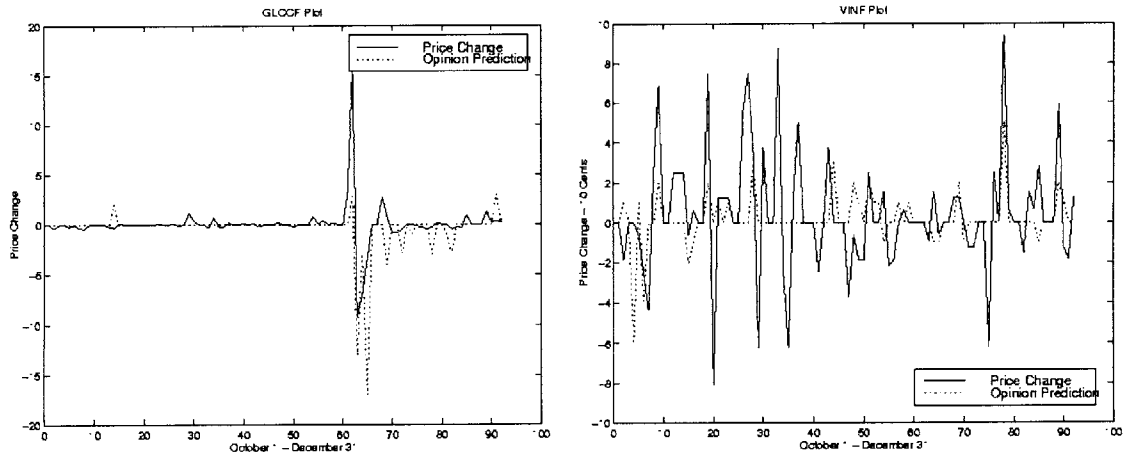
3.2.2 Hourly Regressions

The TAQ data used for hourly regressions includes records of all transactions for NYSE, NASDAQ, and AMEX stocks. INNI had to be dropped as it was not available in the TAQ data (OTC BB stocks were not available), but the remainder of the stocks were used in the hourly regression analysis. The hourly time period caused the message data to become very sparse, since the boards studied had a daily average of 3-4 messages. Since the TAQ data contains transaction level data, the last transaction in each hour period was used as the "closing" price for that hour. The price data was relatively sparse as well, since nontrading hours outnumber trading hours by more than 3 to 1.

The results from the tests were entirely insignificant. For the most part, R^2 values for the regressions were always less than 0.1, and were very often less than 0.01. Except in very few cases, individual coefficients were always insignificant. Even in those cases where significant coefficients existed, extremely low overall predictive value implied that the results were not useful. These results were not unexpected, as the data used were sparse, and many messages were not sent during trading hours. It can be concluded from the graphs that messages were often just an hour or two removed from major price changes, but with no consistent lead or lag pattern. While this detailed analysis did not work for the three month period, it had some use in examining specific incidents where message density and stock activity were high. These results are examined in Chapter 4.

3.3 Regression Results by Industry

3.3.1 Software/Internet Companies



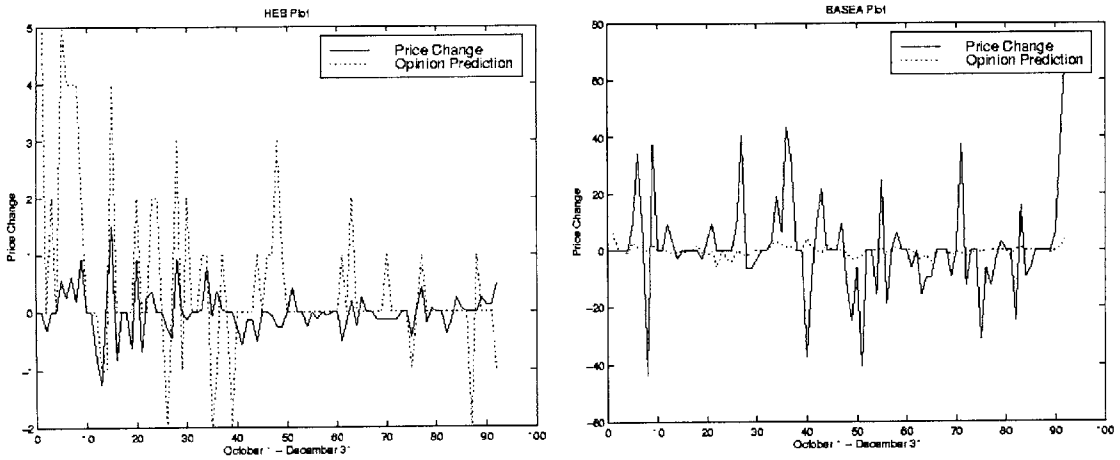
Graphs 3-5 and 3-6: GLCCF and VINF Price Change – Opinion Prediction Plots

GLCCF displayed positive significant relationships between price change, message volume, and the message predictive values. Regressing price change against message volume, prediction, and index change yielded nonzero coefficients for all regressors, with an R^2 of 0.51. Regressions testing the predictive value of messages showed significance for the previous days' message volume and fact values. Reactive regressions showed much stronger results, with an R^2 of 0.81 when the regressors were past message volume, fact, and opinion. The regression results from INNI (Graph 3-3) showed the strongest relationships among stocks studied, with all message variables positively related to price change, and R^2 values for price change regressions around 0.5. Both INNI and GLCCF had large spikes in their stock price during the period, which the message boards followed closely. These spikes appear to explain the extremely strong regression results of these message boards.

While VINF's regression results appear weak when compared GLCCF and INNI, the message prediction and opinion were still found to have significant positive

correlations with price change. In contrast with GLCCF and INNI, Graph 3-6 shows that VINF stock price gyrated wildly over the three month period. Significant relationships between price change and message prediction and opinion were observed, dispelling the notion that all results were based on a one way trend in price. Message fact was not significant for VINF. While reactive regressions for VINF produced no results, predictive regressions produced some significant results. Still, these results were not as significant as those produced by unlagged regressions.

3.3.2 Biotech/Pharmaceutical Companies



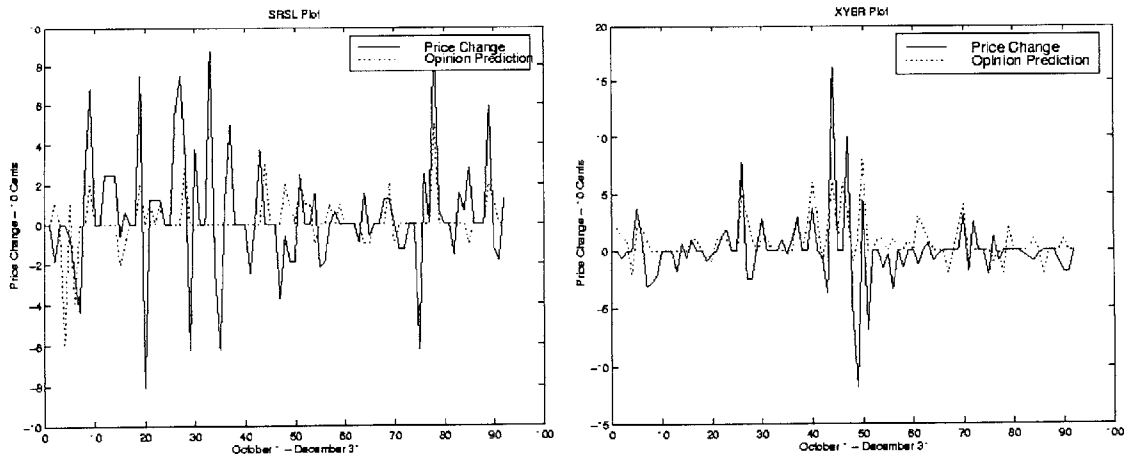
Graphs 3-7 and 3-8: HEB and BASEA Price Change – Opinion Prediction Plots

Graphs 3-7 and 3-8 plot price change and opinion prediction over time for HEB and BASEA. The opinion prediction magnitudes in the graphs are not comparable, since the y axis of the HEB plot display units of 1 while the y axis of the BASEA plot displays units of 20. While the R^2 values of the BASEA regressions were generally less than 0.2, some significant relationships were observed. Message prediction had a strong positive relationship with price change. Regressions of tomorrow's stock prices against today's

message content showed some significance, but generally less than the unlagged data. Reactive regressions showed almost no significance, indicating that the BASEA message board tracked stock movements without much time delay. HEB showed similar results, with R^2 values for price change regressions ranging as high as 0.25. Message prediction and opinion were the two strongest predictive variables, with both having positive correlations with price change. Message fact value displayed no significance for HEB, while it was significant for BASEA regressions. Neither predictive nor reactive regressions displayed any significance with the HEB message data.

TCLN displayed no significant results in any of the trials. These results with TCLN were somewhat expected given the noise observed in the initial plots. The TCLN message board contained a significant number of noise messages which had little relevance to the stock. Most of the regression R^2 values were below 0.1. Message volume was not related to either trading volume or price change. TCLN's message board was an aberration from the norm, as it was the only message board with no significant results.

3.3.3 Hi-Tech/Hardware Companies

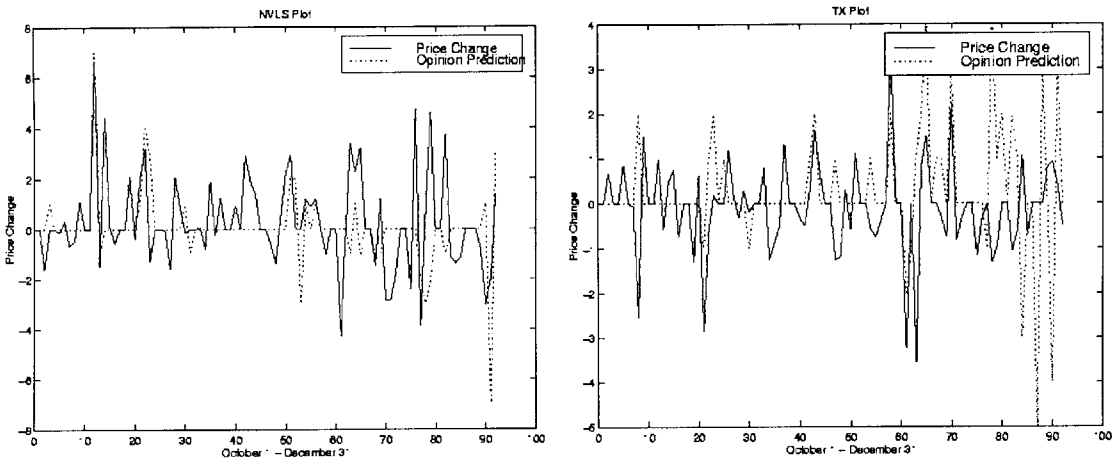


Graphs 3-9 and 3-10: SRSL and XYBR Price Change – Opinion Prediction Plots

Graphs 3-9 and 3-10 show that SRSL was more volatile than XYBR over the period. In XYBR price change regressions, message opinion values showed a strong positive correlation, and message prediction showed a weaker positive correlation. In an individual regression, message fact had a positive relation with price change, though with an R^2 of 0.06. Regressions of trading volume against message volume yielded positive correlations for message volume, with R^2 values around 0.54. The predictive and reactive regressions both showed no results for XYBR.

SRSL regression analysis showed that both message prediction and opinion had significant positive relationships with price change. Message fact had no relationship with price change, however. SRSL price regressions had a very good quality of fit since the stock price was strongly linked to the index price. This relationship did not appear in price change regressions. Reactive regressions showed that message prediction and opinion followed the previous day's price activity. SRSL stock prices fluctuated greatly over the period (see Graph 3-9), but the SRSL board had very low traffic, which decreased the strength of the statistical relationships.

3.3.4 Large Capitalization/Control Stocks



Graphs 3-11 and 3-12: NVLS and TX Price Change – Opinion Prediction Plots

Both NVLS and CPU showed extremely strong relationships between price and the Nasdaq index. As shown in Graph 3-11, NVLS rose sharply with semiconductor stocks over most of the period, while CPU dropped at the same time. In price change regressions, both stocks displayed a positive relationship between index change and price change. The NVLS message board tracked stock activity very well overall, with all message variables showing a positive relationship with price change. The CPU message board was interesting since it was the only board to show a positive relationship between fact and price change, but no other significant relationship.

The EK message board showed very strong statistical results. Price regressions were significantly negatively correlated with both the index and message volume. In price change regressions, the message prediction showed a significant positive relation with price change. Opinion values had a weaker positive relation with price change. R^2 values for the price change regressions ranged as high as 0.34 when message prediction was the key regressor. Message fact values were insignificant in all EK regressions. In TX price regressions, there was an extremely strong negative correlation with the index

value. This is due to the fact that the Nasdaq climbed over the period as oil stocks like TX slumped. When price change regressions were observed, little significance remained. The index change and message prediction variables both showed some positive relationships with price change. The R^2 values for these regressions were never higher than 0.14, however.

3.3.5 Comparative Analysis

Comparing the price change regression results across industries and other stock characteristics shows that little difference is observed across these dimensions. In terms of statistical significance, all but one board had significant price change – message relationships. In terms of quality of fit of the price change regression (R^2), Chart 3-13 shows that there was no difference across company size or industry.

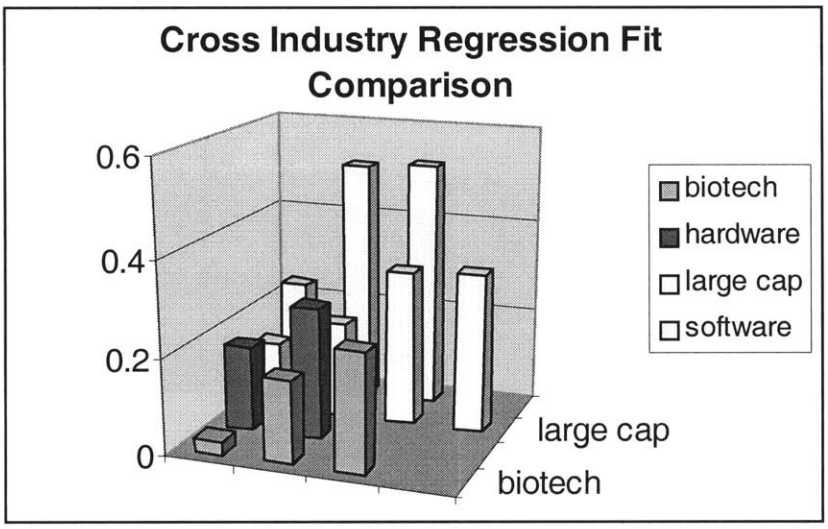


Chart 3-13

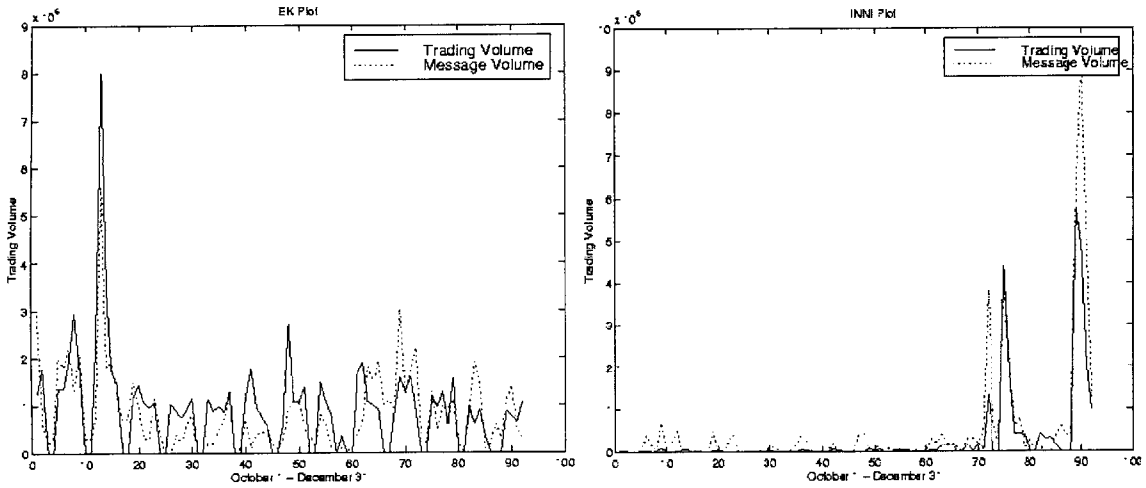
The vertical bars in Chart 3-13 display the R^2 value for each message board's price change regression. The data is taken from Table 3-2 and represented by industry here.

The bars in each row represent the companies within an industry, ordered from worst to best R^2 value. The most surprising result here is that message boards following large companies had results no different from those of small company message boards. The third row from the front displays the results for large companies. The worst of the large company message boards was the second worst performer among all boards, while the best large company board was the third best. Biotech, hardware, and software company boards all have results within the range of the large company boards. Large and small company message boards are among the most accurate (EK and INNI) and the least accurate (TX and TCLN). This shows that the stock – message board relationship was no different for large and small companies.

None of the message boards was shown to have any strong predictive value. This repudiates the hypothesis that small company message boards might consistently provide the market with predictive information on stock activity. Almost all message boards track stock activity well, regardless of company characteristics. Both I/NET, a small internet software concern, and photography giant Eastman Kodak showed relatively accurate tracking of stock activity. Across industries, biotech companies tended to fair worse overall, but when TCLN is excluded, their results are in line with those of other industries. Software companies seem to have the highest quality of fit as a result of INNI and GLCCF. Both of these stocks had enormous price spikes which caused extraordinary message board activity, however. The R^2 ranges for each industry's message boards overlapped significantly. In general, message boards' tracking of stock activity does not seem to differentiate based on characteristics of the stocks followed.

3.4 Trading Volume – Message Volume Relationship

Table 3-2 shows the quality of fit for the primary trading volume regression, where trading volume is regressed against message volume and message prediction. Every message board showed a significant positive relationship between trading volume and message volume except TCLN. In general the trading volume regression was also more statistically accurate than price or price change regressions. INNI had an almost perfect relation between trading volume and message volume, with an R^2 of 0.91 for the primary regression and a t-statistic above 6 for the message volume coefficient.



Graphs 3-14 and 3-15: EK and INNI Trading Volume – Message Volume Plots

Graphs 3-14 and 3-15 plot trading volume and message volume over time for EK and INNI, respectively (See Appendix 2 for plots of all stocks). Trading volume is in shares, and message volume is scaled so that it can be plotted on the same magnitude. The trading volume graphs have gaps in the plot since not trading occurs on weekends. These graphs show that the EK and INNI message boards both track trading volume very effectively. Both graphs also show that high message volume does not precede high

trading volume, but rather occurs simultaneously or after high trading volume. Time lag regressions confirm this analysis, showing that in general message volume tended to lag trading volume rather than lead it.

Table 3-2 shows that message boards which effectively tracked price change also effectively tracked trading volume. The message boards with the highest R^2 values for the volume regression also had the highest R^2 for the price change regression. This result was expected, since message activity has to move with trading volume in order to track stock activity. In general these results show that message board traffic volume tracks with trading volume extremely well. This is the strongest result obtained from analysis of the stock – message board relationship.

4 Trading Effects of Message Boards

The analysis of the stock price – message board relationship made no clear statement on whether message boards have affected trading patterns. Volumes of anecdotal evidence from the media and SEC cases have shown that message boards are affecting stock trading activity in a variety of ways [9] [16]. Message boards have definitely increased the speed of information dissemination. But an increasing tide of hype and touting has cluttered message boards as well. Specific incidents with companies like Leap Group, Books-a-Million, and Boston Chicken indicate that message board hype has caused spikes in these companies' stock prices [11]. Quantitative analysis is necessary to determine whether message boards have caused a systematic increase in volatility for certain types of stocks.

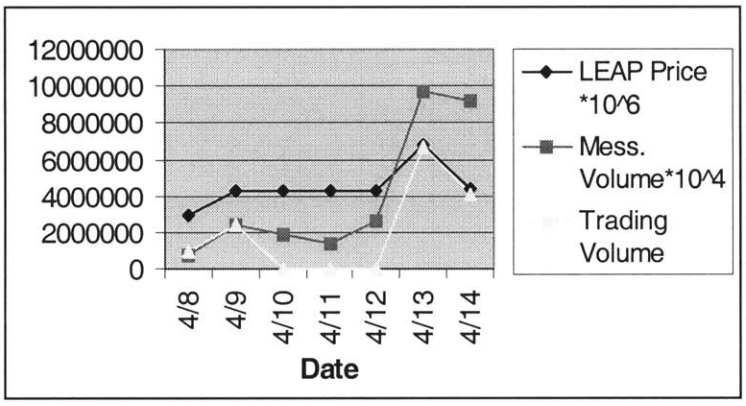
Section 4.1 describes an example scenario which shows how a company's stock price could be affected by a surge in message traffic. Section 4.2 examines the anecdotally observed effects of message boards. Section 4.3 details the results of regression analysis. These results show that message boards have caused no systematic effects on stocks, but they have contributed to volatility in stock price in certain cases.

4.1 An Example Scenario

Examining a sample scenario serves to illustrate the effects that message boards could have on stock trading. From April 8 to April 14, 1998, Chicago based advertising firm Leap Group experienced a quadrupling of its share price before a subsequent crash back down to previous price levels [5]. In early April two other advertising agencies with

at least half internet derived revenue announced favorable results. Shortly thereafter a New York Times advertising column mentioned an award garnered by a Leap group subsidiary. Then, over the week from April 8 to 14, Leap stock skyrocketed as hype that Leap would be the next internet advertising winner poured across message boards. Only 10% of Leap's revenue derived from internet advertising.

Concurrent with this sharp price spike, message board activity on the Yahoo! LEAP board soared. Starting on April 8 and continuing for several days, an investor named Sam Ko used a number of aliases, chief among them Lion_Master88, to spread hundreds of messages touting LEAP. Graph 4-1 documents the behavior of the Yahoo! LEAP message board over the time period, and Chart 4-2 shows a profile of LEAP stock. The LEAP price and message board volume plots in Graph 4-1 have been scaled in order to plot them against trading volume and show the important trends. The LEAP profile given in Chart 4-2 includes some current (May 1999) share statistics. These do not represent the exact figures at the time of the incident, but still illustrate the magnitude of company shares in trading.



LEAP Stock Profile	
Avg Vol prior to Apr. 8	27,000
Peak Volume	6,700,000
Float*	6,500,000
Shares Outstanding*	14,100,000
Market Capitalization*	34,800,000
*Current figures, very similar to those on April 8	

Graph 4-1 and Chart 4-2

An analysis of the facts shows how the message board hype played a major role in the movement of LEAP stock. Graph 4-1, which displays daily data, does not show any evidence that message board activity led stock movement. The message board traffic probably affected stock movements by the minute, however. From April 8 to April 14 a total of 2783 messages were sent on the Yahoo! LEAP message board. This does not account for the hundreds of messages Lion_Master88 and others spammed to other message boards. While most web sites were unwilling to provide statistics, Silicon Investor was able to provide information allowing a conservative estimate of 10 readers per message board post. Since LEAP is a low priced stock, the average LEAP investor could be conservatively expected to purchase 1000 shares. If 10% of the readers reading the 2783 messages (27,000 readers) each purchased 1000 shares, a total of 2.7 million shares would be purchased! This figure represents a substantial percentage of the trading volumes and float shown in Graph 4-1 and Chart 4-2. The message board thus played a significant role in the sharp spike of LEAP price. Many individuals on the Yahoo! LEAP board still blame Lion_Master88 for their losses, which indicates that people were actually compelled to invest based on message board hype.

4.2 Observed Effects

4.2.1 Effects on Information Dissemination

Considerable evidence has been put forth in the media to suggest that message boards are increasing the rate at which information on companies makes its way to investors [7]. Message boards do not compete against financial news sites to bring traditional information to the investor; rather, they increase information flow by bringing

previously private or little known information into the public. Individuals began posting messages about Iomega production increases after they noticed that the company parking lots were always full [10]. Similarly, hours after a Boeing rocket failure over the Pacific message boards were filled with information and speculation on the event [1]. Occasionally insider information is even released prematurely via message boards. An individual on a TI message board broke news of a rumored sale of TI's memory division to Micron six weeks before the deal actually took place. The Network Associates buyout of Cybermedia was also announced on message boards before its formal announcement. The anonymity of message boards has allowed individuals to post this kind of information more freely. While company insiders and other stakeholders post valuable information, the majority of content on message boards is still hype and opinion. The next section explores some of the observed effects of message board hype.

4.2.2 Hyping and Touting on Message Boards

Message boards have had an influence on market activity through both legal and illegal channels. The Securities Exchange Commission has already filed 23 cases against 44 individuals who have used the Internet to illegally tout stocks [9]. The act of publicizing a stock is not illegal, but promoters are required to disclose any resulting compensation received. In some cases individuals "pumped" and "hyped" the stock for their own personal profit, while in other cases, disgruntled ex-employees have posted false information about their previous employers. The SEC recently charged Steven A. King, president of StockstoWatch.com, for Internet fraud. According to authorities, King

urged subscribers to buy shares of a Florida-based medical firm called Surgical Safety (SURG). StockstoWatch.com predicted that the stock would rise from $0.96 to $20 within 18 months. After two days SURG traded at $3.13, so King and his firm sold their shares and profited $573,753 [9].

Stocks that tend to be highly affected by the pumping and hyping phenomenon are small penny stocks. Most of these stocks have very little media coverage, no institutional investment, and low trading volumes. The following are examples of stocks which were purportedly affecting by message board hype.

1. Leap Group (LEAP): See section 4.1 for a detailed explanation.
2. Books-A-Million (BAMM): A small Alabama bookseller unveiled a new online book retailing web site in early December and saw its price rise from less than $5 to nearly $40 before stabilizing near $10. This dizzying fluctuation was thought be linked to day traders and nearly 8000 messages sent over a one week period [2].
3. Imaginon (IMON): IMON, a small software developer in San Carlos, Calif., saw its message board activity rapidly escalate during December 1998 and January 1999. Trading volume rose from near zero to 1.84 million a postings went from a few dozen a week to a few dozen an hour. IMON CEO David Schwartz attributed increased stock volatility to the message board chatter [16].
4. AgriBioTech (ABTX): ABTX fell more than 20% after an individual made damaging comments about the company and its top executives on a Yahoo message board. The comments, posted August 25, 1998, caused record trading volume of 6.3 million shares and a 20% drop in price, according to John Francis, a company cofounder [19].

5. Boston Chicken (BOSTQ): According to Herb Greenberg, senior columnist at theStreet.com, BOSTQ experienced a spike in stock price based on unfounded message board chatter. After the restaurant chain filed for bankruptcy, message board rumors that McDonald's would acquire BOSTQ fueled a rise from $0.47 to $1 [8].

Ingrid Eisenstadter of Individual Investor Online best describes this phenomenon:

> "To get the ball rolling, all a clever person with lots of aliases needs to do is churn up enough excitement to double the stock's average daily trading volume - which immediately sets its ticker flashing across the monitors of day traders, market makers, and other momentum players who screen for fast-moving issues. When momentum traders start roiling the waters, the company's small float makes rapid price movement likely, which in turn starts the ticker flashing on even more investor's computers, until the hysteria becomes self-perpetuating. That is, until the smart players begin taking profit ... leaving ... a three-month hangover for hundreds or thousands of investors who couldn't get out fast enough..." [5]

4.2.3 Corporate Reaction

In reaction to these incidents, companies are also now playing a more active role on message boards. 75% of large corporations actively monitor the Internet's public discussion bulletin boards. eWatch, a leading provider of monitoring online public discussion, has been hired by over 600 companies and actively monitors more than 50,000 public electronic bulletin boards in the Usenet groups, America Online, Prodigy, CompuServe and Microsoft Network [18]. 5% of corporations actively post messages now as well. Chris Edgecomb, CEO of Star Telecom, actively posts on message boards, saying, "Typical shareholders aren't attending the Merrill Lynch conference... This is leveling the playing field for them."

Companies are also actively pursuing law suits and other legal means to stop libel and defamation online. Cohr, Inc., (CHRI) charged its former COO, Sandy Morford, of

posting defamatory messages [12]. According to Cohr, "defamatory, malicious" messages were posted on Yahoo! Finance "to drive away customers and investors." After Cohr launched an investigation into the matter, Yahoo's email records directed Cohr executives to Morford. In another case, a company official representing Legacy Software sued three individual investors for posting defamatory comments on a message board. One of the defendants, Dean Dumont, logged on to Silicon Investor and discovered a message containing a legal filing with his name.

4.3 Quantitative Analysis

Two different types of regression analysis were performed to examine the trading effects. Regressions of average monthly volatility against message traffic rates determined whether there was a systematic increase in volatility associated with message traffic for certain types of stocks. Hourly price change regressions were used to determine whether message boards showed a predictive relationship with price change during certain specific incidents where message board influence was likely.

4.3.1 Volatility Regression Analysis

The volatility regression analysis attempted to measure a relationship between average monthly stock volatility and monthly message traffic rates.[6] The time period for these regressions was January 1998 to March 1999. The long time period allowed for observation of any potential trends in message rates and stock volatility over time as well. Table 4-3 displays the results of the volatility regressions. The stocks in the table are

ordered by the quality of the price change regression of Table 3-2 for comparison purposes. The R^2 for volatility regression column shows the R^2 for a regression of monthly volatility against message rate and index volatility. The next two columns indicate whether message rate and index volatility were significant at the 5% level. The Nasdaq index was used for all stocks except EK, CPU, and TX, where the S&P 500 index was used.

Stock Name	R^2 for Volatility Regression	Message Rate Significant?	Index Volatility Significant?
INNI	0.7304	Yes	No
GLCCF	0.2957	No	No
EK	0.0222	No	No
NVLS	0.5099	No	Yes
XYBR	0.8368	Yes	No
HEB	0.5985	Yes	No
VINF	0.8571	No	Yes
CPU	0.038	No	No
SRSL	0.3669	No	No
BASEA	0.1861	No	No
TX	0.5571	Yes	No
TCLN	0.6035	Yes	No

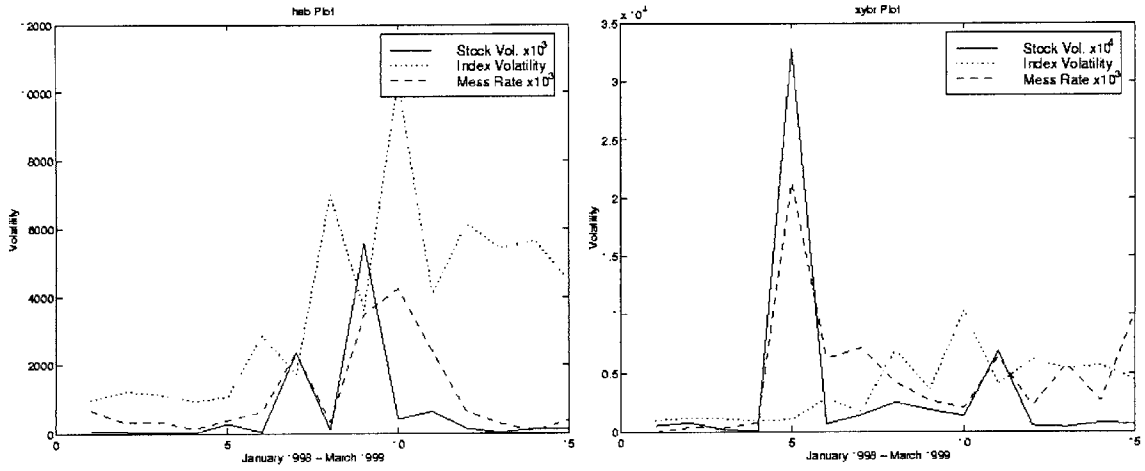
Table 4-2 Volatility Regression Results

Roughly half of the stocks examined show a significant relationship between stock volatility and message rate. There does not seem to be any correlation between these stocks and the stocks which have relatively good price – message board relationships. Contrary to expectations, message rate and volatility did not generally increase over the time period (See Appendix 3). There was no general relationship between volatility and message rate found for all stocks. No general relationships were observed by industry either. For each industry, there were some message boards which showed a relationship between volatility and message rate and others which did not. Among software

[6] Traffic rates were used instead of monthly volume to normalize for the lengths of different months.

companies, INNI had a significant relationship while GLCCF did not. Similarly, biotech companies TCLN and HEB showed significant relationships, but BASEA showed no relationship. Computer hardware company XYBR showed a significant relationship, while SRSL did not, and of the four large cap companies, only TX had a significant relationship. This demonstrates that no trends were observed across industries for the volatility – message rate relationship.

While only TX showed a significant coefficient for message rate in the regressions, half of small companies had a significant coefficient on message rate. This proportion shows that volatility and message rate have a stronger relationship for small company boards than for large company boards. The strong results for small company message boards suggest that these boards have a better relationship with stock volatility. This may be the case since simply because small cap stocks are relatively more volatile than large cap stocks. Message traffic may increase volatility in small stocks by magnifying the effects of stock movements, however. A small number of message board readers can influence these stocks, since the market capitalization for some of these companies is less than $50,000,000. For certain stocks the data does suggest message traffic amplified stock volatility through feedback.

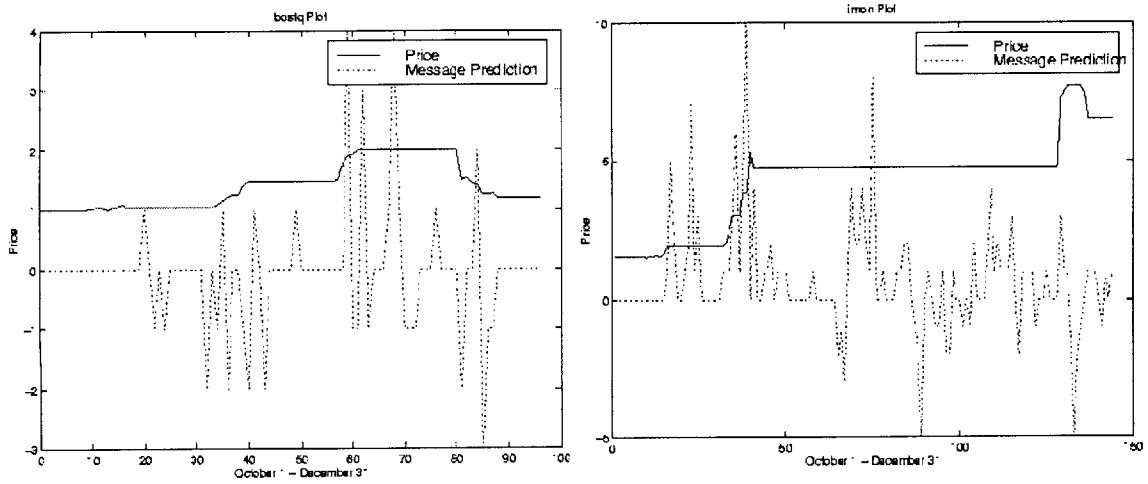


Graphs 4-3 and 4-4: HEB and XYBR Volatility – Message Rate Plots

Graphs 4-3 and 4-4 show plots of stock volatility, message rate, and index volatility over time for HEB and XYBR. The volatility of HEB stock rose sharply in the summer of 1998, and message rates tracked closely. This volatility was due in part to a strong sell recommendation by a short seller named Manuel Ansensio. Ansensio published reports on the message boards and other web sites that disparaged HEB's research into specific drugs [9]. This volatility was probably magnified by the volume of message traffic sent at the same time, particularly since the sell recommendation was issued primarily over the internet. XYBR stock price spiked after an announcement of a manufacturing deal with Sony. XYBR tripled in value, and then settled at 1 ½ times pre-announcement value, reflecting the company's improved outlook. Some of the activity in the stock spike probably occurred as a result of associated message traffic. Graphs 4-3 and 4-4 and the volatility regressions cannot verify these hypotheses, however. The hourly regressions in the next section attempt to verify these effects for a few incidents.

4.3.2 Specific Incident Hourly Regressions

BOSTQ and IMON were two of many stocks covered anecdotally in the media when their stock prices fluctuated wildly (see 4.2). The data were examined for evidence of causality in the stock price – message board relationship using hourly regressions. BOSTQ was examined over the period from November 2 to 5, 1998, and IMON was examined from January 14 to 19, 1999. Graphs 4-5 and 4-6 show plots of the behavior of these stocks over the periods. These plots display price rather than price change trends since hourly price changes are quite random, and create a noisy plot. The price plots smooth out the fluctuations and show the trend over the entire period examined.



Graphs 4-5 and 4-6: BOSTQ and IMON Price – Message Prediction Plots

The regression results for both BOSTQ and IMON were considerably less accurate than the results of the daily regressions. Unlagged price and price change regressions for both stocks showed statistical significance. Predictive price change regressions produced no significant results for either stock. Predictive price regressions did yield significant results, however. At the hourly level trading fluctuations can cover the underlying trends in price. The price regressions smooth out these fluctuations, yielding better results.

For BOSTQ and IMON both, unlagged price change regressions produced a significant positive relationship between price change and fact value, with R^2 in the 0.15 range. For both stocks predictive price regressions also showed significant relationships between price and the previous hour's message volume. For BOSTQ reactive price and price change regressions yielded much stronger results, however. The IMON results were more significant. The IMON graph shows that message predictions rose sharply prior to and during the first large price rise. Predictive price regressions showed a stronger relationship with message volume than did reactive regressions. Both predictive and reactive IMON price change regressions showed no significance. This evidence provides support for the claim that the IMON message board influenced stock activity, while it provides no support for the claim that the same occurred with BOSTQ.

5 Characteristics of Effective Message Boards

The statistical relationships between message boards and stock activity have been explored thus far. The effectiveness of message boards as an information medium must be examined as well. Differences in the features and character of different message board sites have been noted in the media [15]. Certain characteristics may make some message boards considerably more useful than others as information sources. Message boards should provide readers with an open, informed discussion on company news and stock movements. On less followed stocks, effective message boards could serve as a good source for breaking news. The statistical relationships have shown that while all boards track stock activity, they do so with varying degrees of effectiveness. A number of characteristics seem to impact a message board's effectiveness. The degree to which message board discussion is open, the presence (or lack) of a dominant figure, and the level of fact content may all affect a board's information quality.

Each of the twelve stocks message boards examined were classified according to a number of characteristics. These characterizations were then compared against the statistical relationships to determine whether certain types of boards tracked stock activity more effectively. Section 5.1 discusses the qualities effective message boards should possess, and 5.2 discusses characteristics that may identify effective boards. Section 5.33 analyzes which characteristics are most important in determining a message board's effectiveness.

5.1 Defining "Effective" Message Boards

5.1.1 Message Boards Following Large Cap Stocks

Desired Functions
Informed opinions and future predictions
Rapid aggregation of news feeds
Discussion of analyst opinions and stock activity

Message boards following large cap, well publicized companies serve a different function than boards following lesser known companies. Message boards can serve to aggregate and filter important news and analyst reports. Informed discussion of analyst opinions and news stories can also help readers better understand the wealth of information presented to them. Message board participants also provide value with informed opinions and predictions of their own. Effective message boards contain quality due diligence posted by informed investors.

5.1.2 Message Boards Following Small Stocks

Desired Functions
Direct news from insiders and company contacts
Predictions, discussion of current activity
Dissemination of public news

Message boards following less known companies serve a more important role than large company boards. Companies that have little news and no analysts have no consistent flow of information. Message boards can fill this void by aggregating available information, and by providing informed analysis and opinions. More importantly, effective boards are usually able to provide direct news from the company, through posts by company insiders and dedicated shareholders with company contacts. Effective message boards can provide this sort of breaking news consistently.

5.2 Identifying Effective Message Boards

In order to rank the quality of observed message boards, a measurable proxy for the desired features listed above was required. The quality of the stock price – message board relationships were used as this measure. Though this measure does not perfectly capture all of the ideal features of a message board, it seems to be highly correlated with these features. In identifying differences between message boards, a number of defining characteristics were observed. These characteristics seemed to have a definite impact on message board effectiveness, as measured by the quality of the statistical relationships.

The characteristics identified do not represent a comprehensive list of the characteristics that affect message board discussion. They represent an attempt to capture those message board qualities which most affected the quality of discussion. As discussed below, several of the characteristics served to distinguish between more and less effective message boards. They are thus of value in predicting how well message boards will track stock activity. Each of the characteristics and their expected impact is detailed below.

Message Board Characteristics:

1. Open vs. Closed Discussion
2. Dominant Figures
3. Well Focused vs. Noisy
4. Neutral or Polarized
5. Fact driven or Opinion driven

Open vs. Closed Discussion (Debate vs. Support Group)

In an open discussion, different message board members can freely post germane messages expressing any opinion. In a closed discussion individuals are attacked for expressing views contrary to that of the dominant group. Sometimes this silences the

contrarian, and at other times bitter disputes occur, drowning out useful discussion. Message boards can also be characterized as debate groups or support groups. Support groups exist primarily to support like minded investors. This becomes detrimental only when the support group stifles the flow of debate, creating a closed discussion.

Dominant Figures

Many message boards have dominant figures who control the flow of discussion. Dominant figures often play the role of moderator by providing much of the key news and analysis on a board. Other members look to the dominant figures for quality due diligence and analysis of breaking news. Generally the dominant figures on a board have extensive experience with the company. Some message boards have no particular dominant figures, but still possess a core group who contribute valuable information.

Well Focused vs. Noisy

Well focused boards are relatively devoid of irrelevant messages. Some message boards are cluttered by irrelevant messages about other stocks or totally unrelated topics. Message board noise is often created by disputes between members. Since most message boards are not moderated, heated arguments can occur. Message board noise is clearly negative, as it dilutes any quality content on the board.

Neutral or Polarized (bipolar, long dominated, short dominated?)

The overall opinion of a message board can impact how it responds to stock movements. Sometimes long or short dominated boards will fail to accept bad news and

will gloss over the event. With most boards it seems that the overall outlook shifts with time depending on the stock's performance and news outlook. Bipolar message boards often suffer from disputes between longs and shorts, causing noise. It seems that neutral boards might provide the most unbiased analysis.

Fact or Opinion driven

Other factors constant, a fact driven message board should be a more effective information source than an opinion driven board. With smaller companies news is generally scarce, however. Informed opinions are better than no information in this case. Since message boards are unmoderated, the quality of facts is also questionable.

5.3 Characterization of Observed Message Boards

5.3.1 Statistical Data

Table 3-2 is reproduced to show the quality rankings of the examined stocks.

Stock Name	R^2 for Price Change Regression	Prediction Significant?	Fact Sig?.	Opinion Sig.?	R^2 for Volume Regression	Message Vol. Sig.?
INNI	0.5193	Yes	Yes	Yes	0.9125	Yes
GLCCF	0.5108	Yes	Yes	Yes	0.4592	Yes
EK	0.3358	Yes	No	Yes	0.5458	Yes
NVLS	0.3273	Yes	Yes	Yes	0.2584	Yes
XYBR	0.2764	Yes	Yes	Yes	0.5308	Yes
HEB	0.2474	Yes	No	Yes	0.4335	Yes
VINF	0.2358	Yes	No	Yes	0.1476	Yes
CPU	0.2011	No	Yes	No	0.2974	Yes
SRSL	0.18	Yes	No	Yes	0.2574	Yes
BASEA	0.1729	Yes	Yes	Yes	0.1599	Yes
TX	0.1438	Yes	No	No	0.1996	Yes
TCLN	0.0285	No	No	No	0.0238	No

Appendix 4 summarizes statistics for the key stock and message data. This data was analyzed in an attempt to determine whether there were statistical trends in the means of

variables like message prediction, fact, and opinion that were related to message board effectiveness. Unfortunately, no trends were found in this data to suggest any differentiation between message boards.

5.3.2 Observational Data

The message content on the INNI, EK, HEB, BASEA, TX, and TCLN message boards is discussed in detail below. This detailed examination will clarify the heuristics used in determining the message board characterizations displayed in Table 5-1. Each column displays one of the characteristics, and indicates the value of that characteristic for each message board.

Stock Name	Open Debate	Dominant Figure	Focused Discussion	Polar	Fact Driven	SI or Yahoo
INNI	Yes	Yes	Yes	Long	Yes	SI
GLCCF	Yes	No	Yes	Short	No	SI
EK	Yes	Yes	No	Short	No	Yahoo
NVLS	Yes	No	Yes	Bipolar	Yes	SI
XYBR	No	Yes	Yes	Long	Yes	SI
HEB	No	No	No	Bipolar	No	SI
VINF	Yes	No	No	Bipolar	Yes	Yahoo
CPU	No	No	Yes	Long	Yes	SI
SRSL	Yes	No	No	Long	No	Yahoo
BASEA	Yes	No	No	Bipolar	No	Yahoo
TX	Yes	No	No	Neutral	No	Yahoo
TCLN	No	Yes	No	Neutral	No	SI

Table 5-2 Observed Message Board Characteristics

INNI experienced a very rapid rise over the course of December of 98. In general the board was very fact driven, and posters constantly looked for a factual basis for the rise. The board is dominated by one individual, Mr. Yongzhi Yang. He provided a

volume of due diligence for the board, and generally controlled the flow of discussion with his analytically based messages.

> "I called Mr. Markee [CEO] this morning to get an update.
>
> IBM engineers are finishing up the last piece of Netscape: SLL (secure transaction part) … The general release should come right after the New Year…
> He [Markee] is still in discussion with IBM on the pricing of the product. But it seems the $1000/per copy figure I used to calculate the revenue from Netscape server is way too law…"
> - Yonghzi Yang, #279, Nov. 5, 1998

> "I also thank you [Yang]. It is reassuring to have knowledgeable investors on the board sharing and educating rather than mindlessly hyping. You'd have to be crazy to take a tax loss on this one in the next two months with all we have to look forward to in the next six months."
> - Loren S., #283, Nov. 6 1998

These messages indicate the general tone of the message board, with Yang providing detailed insights, and others listening closely and echoing his analysis. The board was still open to discussion, largely because Yang and others were open to outside opinion. Very few off topic messages were observed on this board, contributing to its accuracy.

The EK board on Yahoo is an open debate forum, but is dominated primarily by an individual named PSLDigger. PSLDigger is a vehement short, and posts many messages filled with hype predicting the collapse of EK and the market in general. The following message is indicative of PSLDigger's messages:

> "Mr. T would say - "I pity the fool who buys Kodak stock" I say, you've been warned to be very careful. Equal to all profits for last nine years!!!!!! Remember Motorola."
> - PSLDigger, #1117, Oct. 6, 1998

There is noise generated on the board by discussions on the market and other stocks not specifically related to EK. Interestingly, since the board is relatively free of heated debate, the irrelevant messages don't affect the statistical relationship. Heated retorts cloud the useful information in a board as individuals send rhetoric filled messages.

The HEB message board is a constant battle between longs and shorts. Though longs outnumber shorts, both sides vehemently attack each other. Some noise is generated by insults thrown back and forth as a result. This also leads to a fair number of pure hype messages posted as retorts to the opposing opinion.

> "HEB up again! Where are the doubters now?" - Tony Hsu, #467, Oct. 12 1998
>
> "Tony, turned out you picked the top. Here's Asensio's report from attending the Boston confrence. Generally the truth stands the test of time..." - Mad2, #471, Oct. 12, 1998
>
> "To all HEB shareholders and Dr Carter, you are absolutely insane, you actually are trying to defend a market value in this Loser of a Company ... explain how HEB could possibly profit from a Drug that has been written off by every Real Company" - Anthony@Pacific, #569, Nov. 5, 1998
>
> "Well, Anthony, HEB is up 15.1% since the October 17th date ... It is all us longs that are the ones in profit positions here, and Anthony, the profits (and your losses) are getting greater every day." - Marty, #570, Nov. 5, 1998

The board serves as a support group for both longs and shorts, and it seems that open, neutral debate is rare. The BASEA message board is similar, characterized by a battle between shorts and longs with a few neutral individuals. Two of these neutral individuals post high quality factual analysis to the board. Unfortunately the neutral figures do not dominate the discussion, allowing a bipolar schism to form.

The CPU message board contains a relatively educated, news driven discussion. Unfortunately, the discussion seems somewhat closed as short opinions are not received fairly. The board members seem very intent on putting a positive spin on all news, as shown here, during a large drop in CPU share price.

> "I agree with you. Buy dips on "Downgrade". I believe CPU will be the best value stock in 1999. Best wishes..." - James Yu, #710, October 7, 1998
>
> "I really believe we have seen the bottom this time. The last time we hit 11 1/4 we shot right back up to a little over 20 and I see no reason we shouldn't have the same this

time... This is certainly a buying opportunity!" - MG1041@aol.com, #748, October 8, 1998

The board is very well focused, with very little off topic discussion. There are several very knowledgeable posters on the board, but the board is dominated by no one.

The TX message board is fairly noisy, with frequent messages on other stocks or topics completely. It is also neutral, as few posters take a strong stand on TX and its present or future value. There seem to be few highly knowledgeable individuals expressing opinions on the board. Content filled messages are posted, but many take a fairly neutral stance toward the stock.

"For about 7 months TX has been trading between 56 and 65 ... Now if the price should break upward in the next couple of months TX will have a PE of well over 25 ... Yet the TX insiders are selling heavily. So maybe this is a wonderful company whose price is going up, but it's run by dummies who are betting the wrong way on the stock?" - EasyQuanter, #583, Oct. 25, 1998

"We are a startup technology company located in Los Angeles looking for individuals with expertise in petroleum logistics ... Trading experience is also useful..." - acharvey, #617, Nov. 13, 1998

The above message typifies the abundance of off topic messages on the board. Off topic and neutral messages together lead to a relatively poor stock - message board relationship for TX. The TCLN board has even more streams of off topic discussion, primarily as a result of irrelevant side arguments. There are dominant figures on the TCLN board, but they often pull discussion towards side topics. Neither the TX nor the TCLN message board seems to convey an outlook on its stock in the end, and as a result both track stock activity relatively poorly.

5.3.3 Analysis

The rankings of the message boards indicate that INNI was the most effective, while TCLN was least effective. The statistical and observational data presented show several trends in message board characteristics. Well focused boards track their stocks more closely than noisier boards, as expected. Message boards with an open discussion fared better generally than those with a more closed discussion. While the BASEA and TX discussion boards both allowed an open discussion, BASEA was a bipolar board with much argument, and the TX board was simply neutral and very noisy. Perhaps more surprisingly, the neutral message boards were the least effective boards. In fact, boards that were either long or short dominated fared better than neutral or bipolar message boards. These boards tended to convey a consistent prediction which was more in line with stock activity than the other types. Neutral boards often did not make any consistent predictions, while bipolar boards experienced a "cancellation" effect where argumentative individuals negated each other's presence.

In all but one case the presence of a dominant figure seemed to have a positive impact on the message boards. This was expected, since dominant figures tend to curb irrelevant traffic by setting the tone of discussion. In the case of TCLN, a few dominant figures existed, but they engaged in irrelevant side debates. The presence of a dominant figure would only be beneficial if that individual set a good example for other members. It was expected that fact driven boards would be significantly more useful than hype driven boards. The statistical relationships show that while fact driven boards performed better overall, opinion driven boards can track stock activity as well in some cases. This understates the value of good factual analysis, however. Messages on hype driven boards

like EK and GLCCF changed opinions quickly based on day to stock activity, whereas fact driven boards like INNI conveyed a consistent message based on due diligence.

6 Conclusions and Future Research

This thesis explored several questions surrounding the relationship between stock message boards and equities trading. These included determining whether a relationship existed between message activity and stock price changes, and analyzing the effects of message activity on stock volatility. These relationships were observed across a range of industries in order to identify broader trends. Message boards were examined to determine which features distinguished effective message boards from ineffective boards. The research yielded strong results in answer to these questions, and also provided insights into the design of future research on stock message boards. Section 6.1 describes the research results and future challenges in regression analysis on stock message boards. Section 6.2 summarizes the research and insights on message board characteristics, and Section 6.3 describes potential work on message board design and business models for message boards. Finally, Section 6.4 summarizes the key conclusions of this research.

6.1 Regression Analysis

The principal regression results produced were that message board content tracks stock activity well, and that message board traffic volume tracks trading volume closely. No predictive relationship was observed between message boards and stock activity, however. Examinations into message boards' effect on trading yielded the result that message boards cause no general effects on market trading. Message hype has increased stock volatility in specific cases, however. Small company message boards were more sensitive to stock volatility in the regressions. This, together with the results of

investigations into two specific cases, indicated that message boards increased short term volatility for some small cap stocks. Regression analysis provided feedback on the message grading systems used as well. The message prediction grade was found to track stock activity as accurately as the message fact and opinion grades together. This showed that simple prediction grading was equivalent to the more complex system. Future research can use a simple grading system with the knowledge that it is effective.

Further research could explore several avenues of stock message board regression analysis. The message fact and opinion values were weighted equally in all regressions. Message fact and opinion could be weighted differently, or even used in an interaction variable (fact*opinion), to determine whether any of these regression models would produce stronger results. The regression analysis examined a total of only fourteen stocks and message boards. Broader samples of stocks and message boards should be examined in order to strengthen the results of comparative analysis across different stocks, industries, and types of message boards. Multiple message boards should be examined for each stock, when possible. The research into specific incidents analyzed only two stocks in detail. This sample must increase greatly in order to find stronger evidence on whether message boards are responsible for the increased volatility in small cap stocks. Throughout the research, time series data was used to construct a specific relationship between a single stock and message board. Panel data and cross sectional analysis should be used in an attempt to uncover general relationships that hold for groups of stocks.

6.2 Message Board Characteristics

Several message board characteristics were shown to have a direct relationship with message board effectiveness. Well focused boards track their stocks more closely than noisier boards. Fact driven message boards were more effective than opinion driven message boards. Message boards with dominant figures were also more effective, since dominant figures curbed irrelevant traffic by setting the tone of discussion. Message boards with an open discussion fared better generally than those with a more closed discussion. Collectively, these characteristics can be used to predict the effectiveness of message boards.

While the message board characteristics proved stable over the period examined, future research should examine trends in these characteristics for message boards. Future research should also attempt to measure some of the message board characteristics quantitatively. Message board noise could be measured by using the fraction of "noisy" (irrelevant messages with grades of all 0's) messages. Fact driven message boards could be identified using appropriately normalized measures of average fact values. Finally, other characteristics which affect message board effectiveness should be explored. The five characteristics examined in this research may not include all characteristics which determine message board effectiveness.

6.3 Message Board Design and Business Model

The research done on characterizing effective message boards can be used in research on message board design. Message board design features like user fees and

access limitations, degree of moderation of the board, spam filtering, and the user interface should be examined in the design of more effective message boards. The value propositions of message boards were not explored in this research. What is the value proposition of message boards for a market brokerage like ETrade or Schwab? To date message boards have enjoyed success through a mixture of community and information [1]. Would smaller, potentially more focused and less noisy brokerage firm forums add value for clients? The number of stock message board web sites is climbing rapidly, making research into effective design and good business strategy more crucial.

6.4 Conclusion

This research shows that message boards track stock activity and news as effectively as other internet news sources. They offer no more than this, and do not represent a superior information source for individual investors. The most powerful reason for stock message board success is the development of a community, as with message boards serving other topics. With stock message boards, however, the community of investors may hold great power over stock activity. The volatility increases observed in a number of small cap stocks cannot be taken lightly. The SEC continues to examine and redefine its definitions of illegal stock "touting." Message boards continue to evolve, and the next several years may tell whether they become more meaningful as information sources or more dangerous as sources of hype and stock volatility. Continued research into the design of more effective message boards will enhance the usefulness of this new information medium.

Appendix 1: Regressions

The following is a list of all regressions run in analysis of price, price change, and trading volume relationships. Regressions were performed using Stata, and the regressions appear as Stata code.

Variable Key:

reg	Stata regression command. First variable in each list is the regressand; regressors follow.
price	Stock price
mvol	Message volume
pred	Message prediction
ind	Index price
change	Stock price change, measured as difference in price from last period to current period.
indch	Index price change
fact	Message fact
op	Message opinion
vol	Trading volume

A preceding "y" indicates a variable that has been lagged one time period. Variable yop is the previous period's opinion. A preceding "t" indicates a variable that has been moved ahead one time period. Variable tfact is the next period's fact. If a numeral is appended to these variable names, that increases the degree of the time shift. For example, tfact3 is the message fact three periods in the future.

```
reg price mvol pred ind
reg price mvol ind
reg price mvol pred
reg price ind

reg change mvol pred indch
reg change mvol fact op indch
*reg change mvol fact op pred
reg change fact op
reg change fact op indch
reg change pred
reg change mvol fact
reg change mvol op
reg change fact
reg change op
```

*This regression is the only regression including fact, op, and pred. It was not actually used, since multicollinearity problems were found when pred was used in any regression including fact or op.

```
reg vol mvol indch pred
reg vol mvol fact op
reg vol mvol

reg price ymvol ypred ind
reg price ymvol ymvol2
reg price ypred ypred2
```

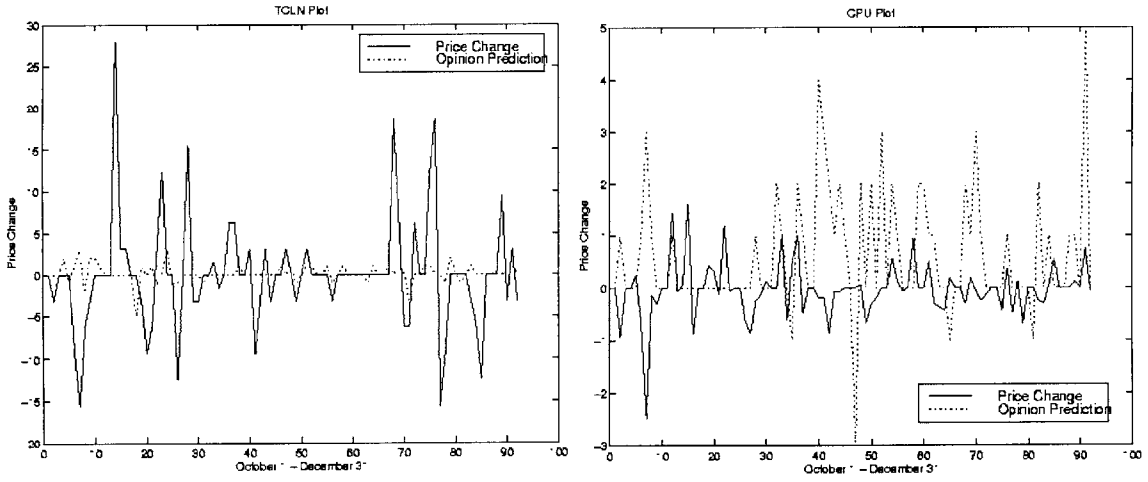
```
reg change ymvol ypred indch
reg change ymvol yfact yop indch
reg change ymvol ymvol2 yfact yfact2 yop yop2
reg change ymvol ymvol2 ypred ypred2 indch
reg change yfact yfact2 yop yop2 indch
reg change ypred ypred2
reg change ymvol ymvol2
reg change yfact yfact2
reg change yop yop2

reg vol ymvol ypred
reg vol ymvol ymvol2

reg price tmvol tpred ind

reg change tmvol tmvol2 tfact tfact2 top top2
reg change tmvol tmvol2 tpred tpred2 indch
reg change tfact tfact2 top top2 indch

reg change tpred tpred2
reg change top top2

reg vol tmvol tmvol2
```

Additional Regressions Run for Hourly Time Units

```
reg price ymvol ymvol2 ymvol3 ypred ypred2 ypred3
reg price ymvol ymvol2 ymvol3
reg price ypred ypred2 ypred3

reg change ymvol ymvol2 ymvol3 ypred ypred2 ypred3
reg change ymvol ymvol2 ymvol3 yfact yfact2 yfact3 yop yop2 yop3
reg change ymvol ymvol2 ymvol3 ypred ypred2 ypred3
reg change yfact yfact2 yfact3 yop yop2 yop3

reg change ypred ypred2 ypred3
reg change ymvol ymvol2 ymvol3
reg change yfact yfact2 yfact3
reg change yop yop2 yop3

reg vol ymvol ymvol2 ymvol3 ypred ypred2 ypred3
reg vol ymvol ymvol2 ymvol3

reg change tmvol tmvol2 tmvol3 tfact tfact2 tfact3 top top2 top3
```

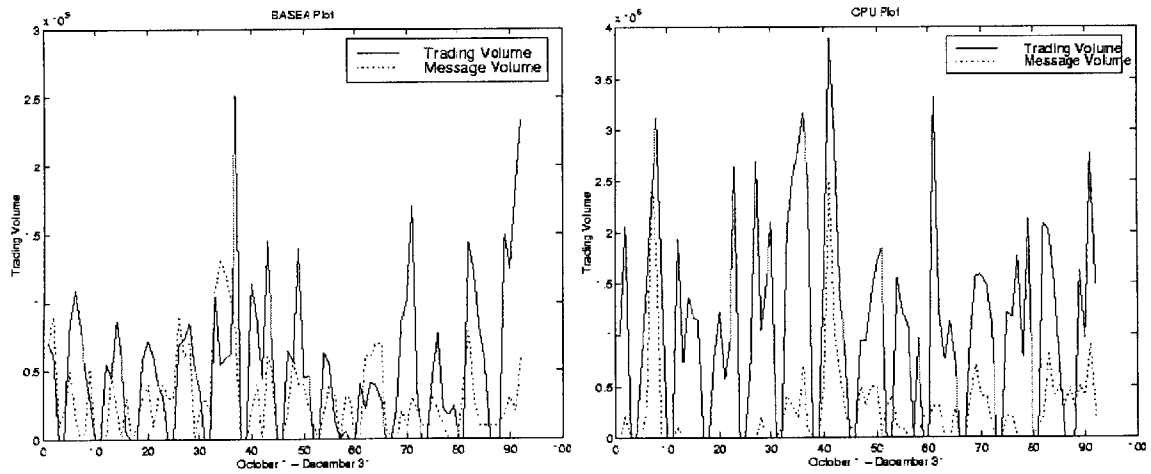
Appendix 2: Graphs of Stock – Message Board Relationship

Chapter 3 includes plots of price change and opinion prediction for all but two stocks, TCLN and CPU. The graphs for these stocks are shown below.

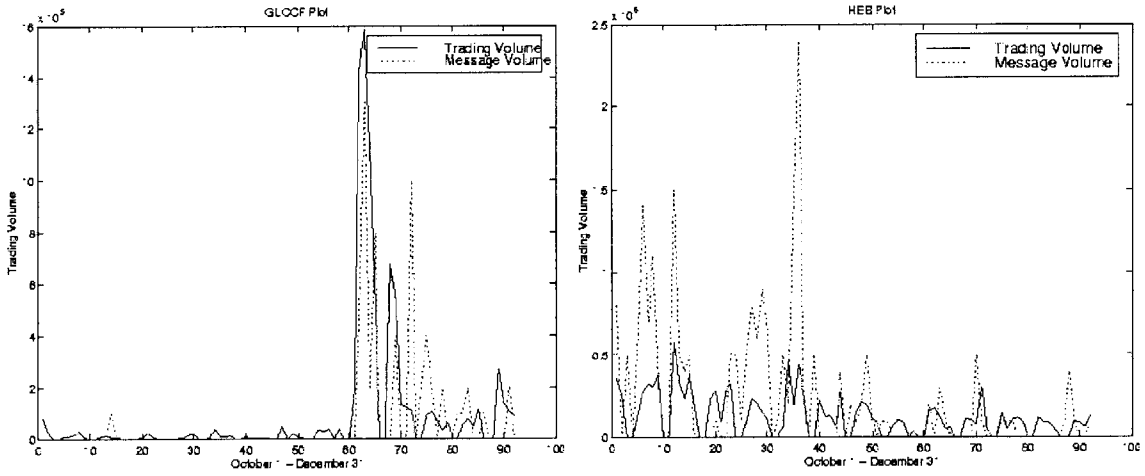


Graphs A1 and A2: TCLN and CPU Price Change – Opinion Prediction Plots

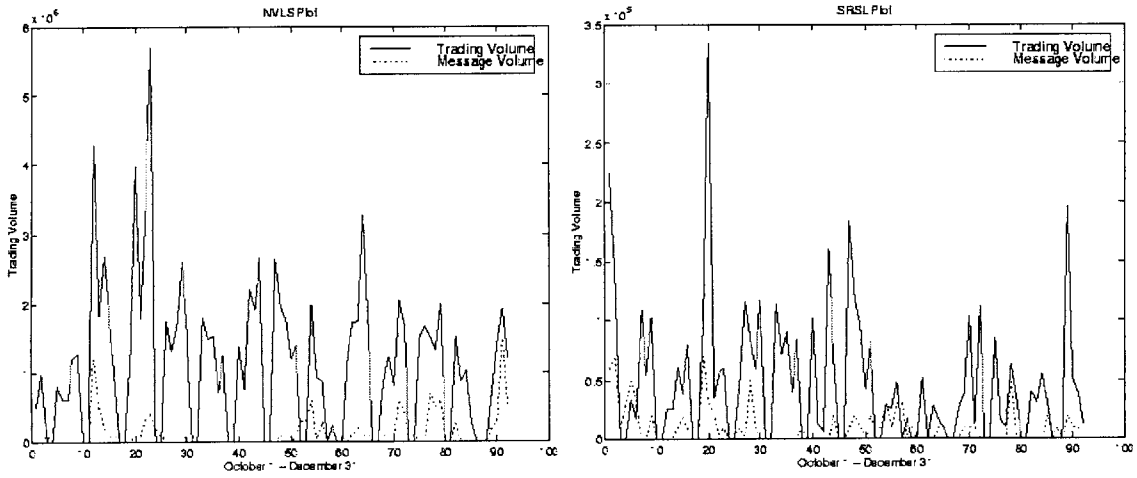
Chapter 3 also includes plots of trading volume and message volume against time for EK and INNI. Similar plots for the remaining stocks follow below.



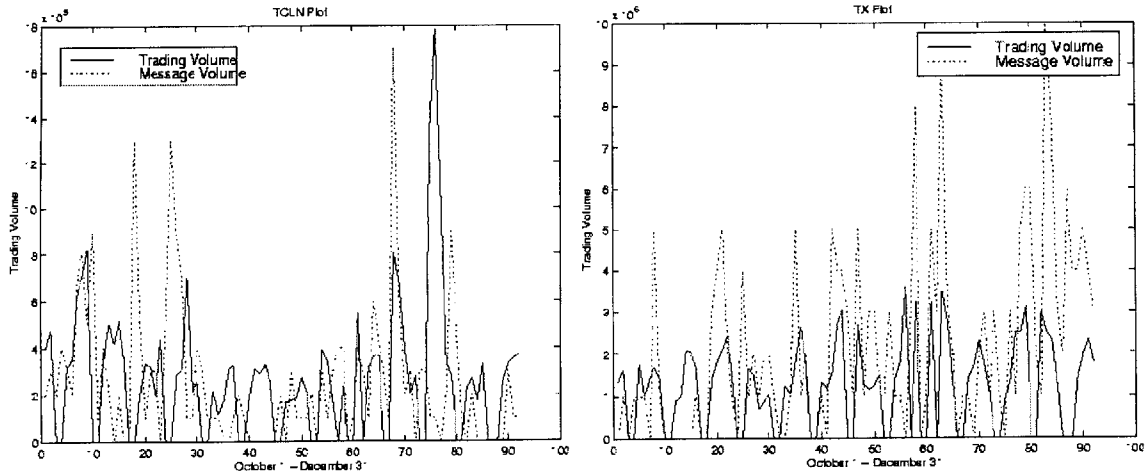
Graphs A3 and A4: BASEA and CPU Trading Volume – Message Volume Plots



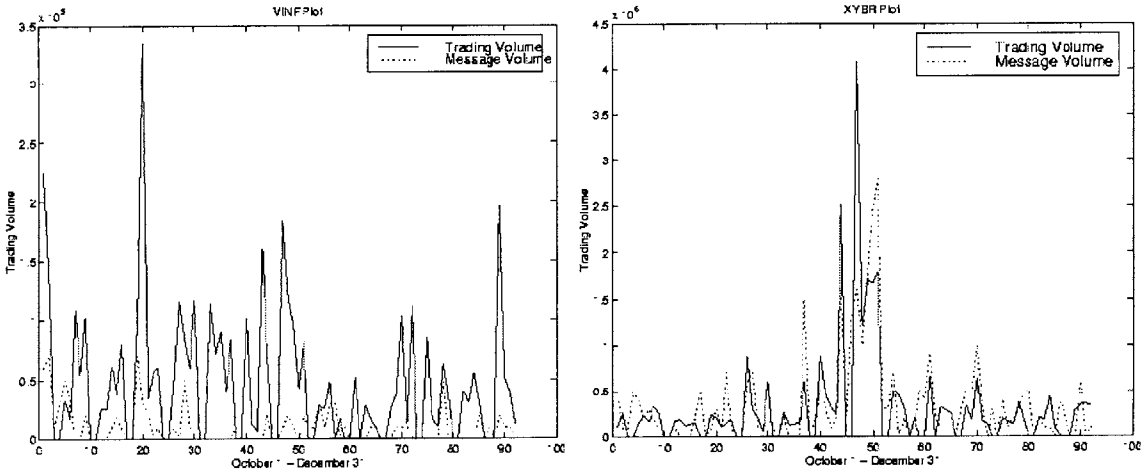
Graphs A5 and A6: GLCCF and HEB Trading Volume – Message Volume Plots



Graphs A7 and A8: NVLS and SRSL Trading Volume – Message Volume Plots



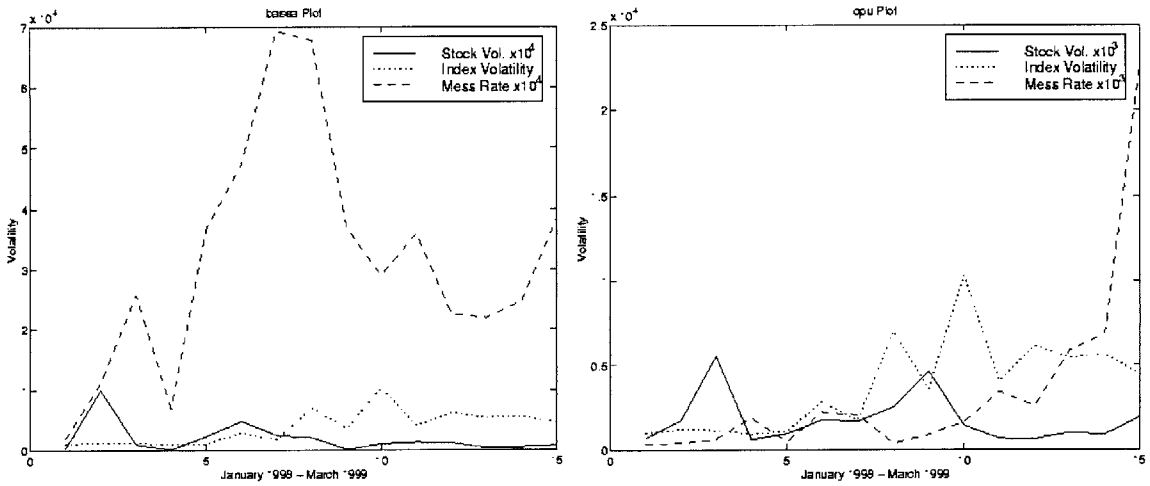
Graphs A9 and A10: TCLN and TX Trading Volume – Message Volume Plots



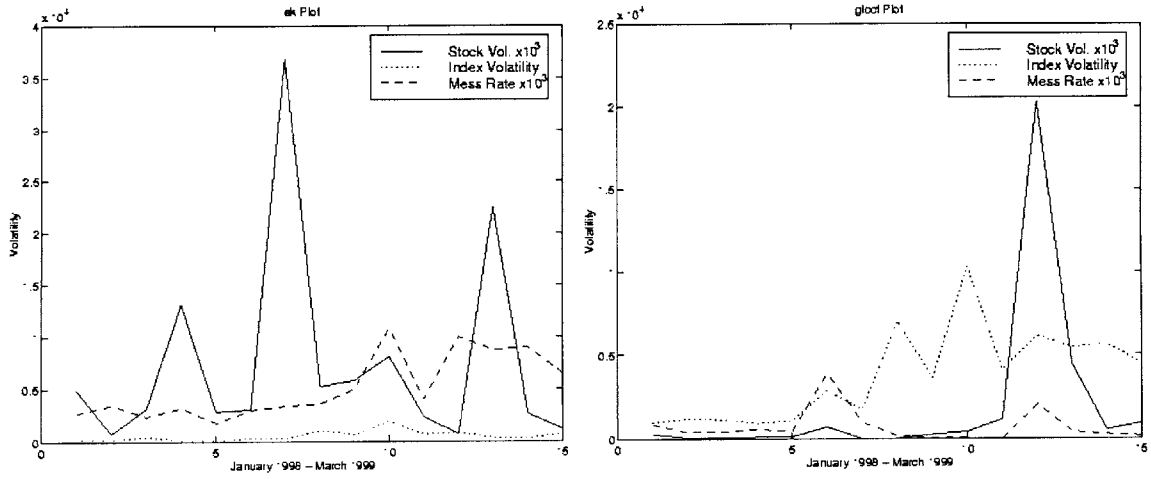
Graphs A11 and A12: VINF and XYBR Trading Volume – Message Volume Plots

Appendix 3: Volatility Regression Analysis

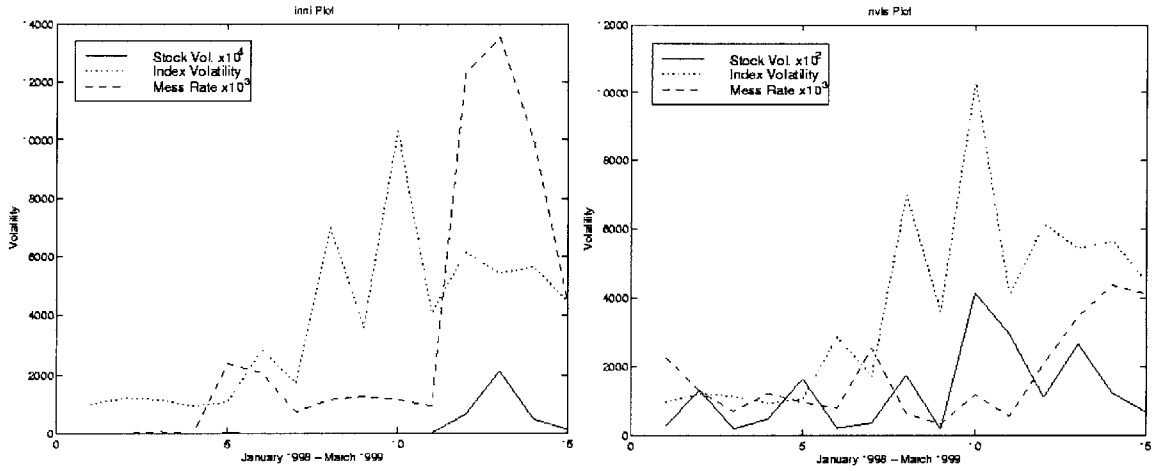
Graphs of stock volatility and message rate over time were provided for XYBR and HEB in Chapter 4. Here similar graphs are reproduced for the remaining stocks. From these graphs it is evident that stock volatility and message rate did not generally trend upward over the fifteen months observed.



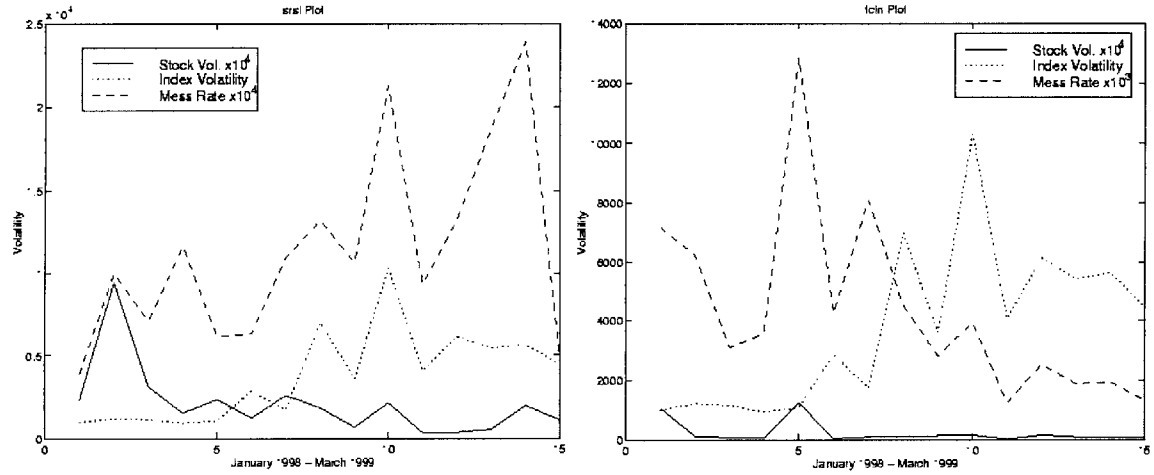
Graphs A13 and A14: BASEA and CPU Volatility – Message Rate Plots



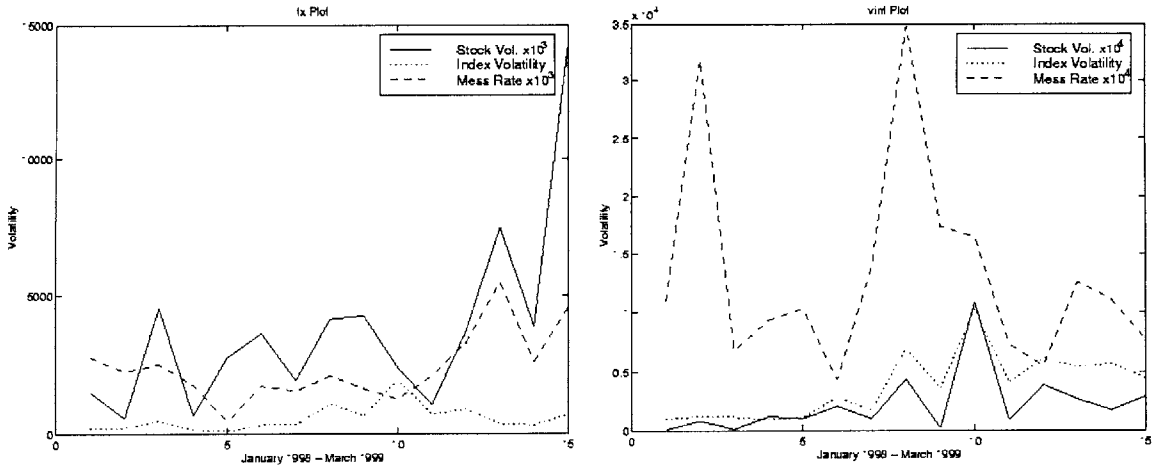
Graphs A15 and A16: EK and GLCCF Volatility – Message Rate Plots



Graphs A17 and A18: INNI and NVLS Volatility – Message Rate Plots



Graphs A19 and A20: SRSL and TCLN Volatility – Message Rate Plots



Graphs A21 and A22: TX and VINF Volatility – Message Rate Plots

Appendix 4: Summary Message Data Statistics

The following is a table of message data statistics which summarizes the mean, standard deviation, minimum, and maximum for the message predictive variables of each message board. "Pred" indicate prediction, "fact" indicates fact value, and "op" indicates opinion value. These variables represent totals for each day. "#Messages" indicates the total number of messages examined for each board.

Variable	#Messages	Mean	Std. Dev.	Min	Max
BASEA	268				
pred		-.0326087	1.10395	-3	4
fact		.1086957	.7479866	-2	3
op		-.2173913	1.721398	-6	6
CPU	235				
pred		.3369565	.8929169	-2	4
fact		.0978261	.6300145	-2	2
op		.5978261	1.139306	-3	5
EK	763				
pred		-.6304348	1.99294	-12	4
fact		-1.130435	2.734466	-12	5
op		-.9673913	2.561126	-13	4
GLCCF	64				
pred		-.2608696	1.405402	-10	2
fact		-.1521739	.9937694	-8	2
op		-.4347826	2.382649	-17	3
HEB	223				
pred		.4021739	1.090012	-2	5
fact		.2282609	.4939916	0	2
op		.4673913	1.378414	-2	5
INNI	447				
pred		1.608696	4.640852	-1	32
fact		.6413043	1.913409	-1	11
op		1.793478	5.155895	-1	35
NVLS	119				
pred		.0869565	1.289451	-7	6
fact		.0326087	.5231159	-2	2
op		.0434783	1.405742	-7	7
SRSL	135				
pred		.2173913	1.04646	-3	6
fact		.0978261	.6123237	-2	3
op		.1847826	1.554358	-6	7
TCLN	237				
pred		.076087	.7447058	-3	2
fact		.1304348	.9968896	-6	4
op		.1304348	1.040049	-5	3
TX	204				
pred		.25	1.001373	-3	4
fact		.2173913	1.097711	-4	4
op		.25	1.280668	-5	4
VINF	90				
pred		.076087	.9046121	-3	3
fact		.0978261	.5148302	-1	3
op		.1630435	1.224989	-6	5
XYBR	323				
pred		.8152174	1.382221	-1	7
fact		.576087	.9046121	-1	4
op		.7934783	1.732292	-2	8

Appendix 5: Message Grading Examples

Several sample messages and their grades are shown below. These examples help illustrate how message content was graded. Message prediction was graded 1, 0, or –1. Message fact and opinion were each graded from positive 3 to negative 3. Positive values indicate that the message content has a positive outlook for the stock.

The following messages are from the CPU Silicon Investor message board:

"<< So CPU will get more profit per unit sold <<

Higher prices translate to fewer units sold at smaller margin, I believe Scotman. Not good.

I do agree with what you said yesterday, though. Eventually, computing will be like watching TV and cost about the same. High tech will move on to an equally exciting territory in currently unchartered waters.

I think high profits in the internet will last a bit longer than "boxes". That might be why Intel and DELL are throwing in with those sorts of companies at this stage."
- gmccon, #740, Oct. 8, 1998

Prediction: -1 Fact: -1 Opinion: -1

"Picked up some CPU today, 500 shares at 14 5/8, and 500 at 14 1/2.I believe that CPU will be able use it's established retailing base. As a very strong platform and name recognition to be a player in the computer/ internet business for many years to come. Now I have 3100 shares of CPU and have made CPU my #1 holding.
My # 1 holdings have yet to disappoint."
- Michael, #857, Dec. 1, 1998

Prediction: 1 Fact: 1 Opinion: 1

"To all:

Briefing.Com has Q2 earnings being released on Jan.4 after market closes.This seems strange,because CPU usually releases toward the end of Jan.Comments?"
- Friend, #893, Dec. 22, 1998

Prediction: 0 Fact: 0 Opinion: 0

The following messages come from the Yahoo BASEA message board:

"Since the W-L decision to go with ME instead of FS several things have occurred. The plant getting the system first has been very happy with the new software. Other parts/ companies contained within Warner Lambert have looked at the evaluation done between FS and ME and looked at what is occuring at this test site. More orders will be coming soon from within other parts of W-L."
- guruofcode, #962, Oct. 9, 1998

Prediction: 1 Fact: 2 Opinion: 1

"This stock is dead

Don't worry about this thing rising,they probably will be out of business in two years maybe sooner with Gardner at the helm ,especialy with Alex at his side.When you have to do business with outfits like Andrew Garrett you know you have hit bottom and the end is near."
- basesucks, #988, Oct. 22, 1998

Prediction: -1 Fact: 0 Opinion: -3

"Your first post said:

<<The most impressive thing from the CC was the 24 million in backlog TG spoke of>>

As you correctly stated in your second posting, there is only $4MM in backlog, but $24MM in activity in the pipeline. Now, as anyone who ever did time in sales knows, a pipeline and a quarter will...well you know the rest. Here's a quick analysis:

Given an optimistic close rate of 50% and a 12 month selling cycle, the pipeline only represents a $12MM annual run rate in sales. I would expect closer to $50MM in identified opportunities if these guys are going to ever be profitable.

I don't doubt the size of the *potential* market for BASEA's (and POMS') wares, but I do doubt the size of the *active* market who is shopping for products like this today, with purchase order in hand. I think it is quite small."
- RBullota, #1061, Nov. 4, 1998

Prediction: -1 Fact: 0 Opinion: -3

References

[1] Batsell, Jake. "Message Boards on Investing can be Insightful or Irrelevant." The Seattle Times. November 6, 1998. Pp. C17

[2] Buckman, Rebecca. "The madness of Online Crowds: How 'Day Traders,' Computers Are Raising a Ruckus with Internet Stocks." Wall Street Journal. Dec 3, 1998. Sec C, Pp. 2

[3] Carbonara, Peter. "Smokin' Joe." Money, April 1999.

[4] J. Durbin and G. S. Watson, "Testing for Serial Correlation in Least Squares Regression," Biometrika, vol. 38, pp. 159-177, 1951.

[5] Eisenstadter, Ingrid. "Look Before You Leap: Where's the fine line between Internet fun and fraud?" Source: http://www.iionline.com/mag/default.asp?-sec=openbell&artnum=1387&month=October1998

[6] S. M. Goldfeld and R. E. Quandt, "Some Tests for Homoscedasticity," Journal of the American Statistical Society, vol. 60, pp. 539-547, 1965.

[7] Goldstein, Alan. "Internet Message Boards full of Fact and Fantasy on Stocks." Dallas Morning News. November 10, 1998. Pp. 1B

[8] Greenberg, Herb. "The danger lurking in investor message boards." Fortune, New York, Nov. 23, 1998.

[9] Hannon, Kerry. "SEC Files actions against Internet stock Hucksters". USA Today. October 29, 1998. Pp. 1B.

[10] Iacovetta, Linda. "Stock Talk Is (Not Always) Cheap: Web chats and message boards can be valuable investing tools or misleading rumor mills." Source: http://www.zdnet.com/zdtv/moneymachine/investing/ story/0,3666,2113569,00.html

[11] Leder, Michelle. "Stemming the Tide of Touts on Those Stock Message Boards." New York Times. February 21, 1999. Sec 3, Pp. 9

[12] Macavinta, Courtney. "Yahoo message board users sued." CNET News.com. September 9, 1998. Source: http://www.news.com/News/Item/0,4,26134,00.html? st.cn.nws.rl.ne

[13] Nickel, Joe. "Part 4: The SEC Tries to Learn New Tricks." Source: http://www.wired.com/news/news/business/story/6984.html

[14] Pindyck, Robert, and Daniel Rubinfeld. "Econometric Models and Economic Forecasts." Irwin McGraw-Hill. 4th ed. Boston, MA, 1998.

[15] Roberts-Witt, Sarah L. "Financial Discussion: Internet/Web/Online Service Information." PC Magazine. March 23, 1999. Pp.162

[16] Wirth, Gregg. "Battle of the Boards." Source: http://www.thestandard.net/articles/display/0,1449,2492,00.html

[17] "Financial message Boards Tackle Stock Hyping." Web Finance. December 14, 1998.

[18] "Yahoo! Halts the Pumping and Hyping." The Toronto Sun. October 21, 1998. Business Section, Pp. 54

[19] "AgriBioTech fall on Net chat." Reuters. Cnet News.com, August 27, 1998.